Osvaldo Gervasi et al. (Eds.)

Computational Science and Its Applications – ICCSA 2005

LNCS 3482

International Conference
Singapore, May 2005
Proceedings, Part III

3 Part III

Springer

Lecture Notes in Computer Science 3482

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

Osvaldo Gervasi Marina L. Gavrilova
Vipin Kumar Antonio Laganà
Heow Pueh Lee Youngsong Mun
David Taniar Chih Jeng Kenneth Tan (Eds.)

Computational Science and Its Applications – ICCSA 2005

International Conference
Singapore, May 9-12, 2005
Proceedings, Part III

Springer

Volume Editors

Osvaldo Gervasi
University of Perugia
E-mail: ogervasi@computer.org

Marina L. Gavrilova
University of Calgary
E-mail: marina@cpsc.ucalgary.ca

Vipin Kumar
University of Minnesota
E-mail: kumar@cs.umn.edu

Antonio Laganà
University of Perugia
E-mail: lag@dyn.unipg.it

Heow Pueh Lee
Institute of High Performance Computing, IHPC
E-mail: hplee@ihpc.a-star.edu.sg

Youngsong Mun
Soongsil University
E-mail: mu@computing.soongsil.ac.kr

David Taniar
Monash University
E-mail: David.Taniar@infotech.monash.edu.au

Chih Jeng Kenneth Tan
Queen's University Belfast
E-mail: cjtan@optimanumerics.com

Library of Congress Control Number: Applied for

CR Subject Classification (1998): D, F, G, H, I, J, C.2.3

ISSN 0302-9743
ISBN-10 3-540-25862-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-25862-9 Springer Berlin Heidelberg New York

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11424857 06/3142 5 4 3 2 1 0

Preface

The four volume set assembled following *The 2005 International Conference on Computational Science and its Applications*, ICCSA 2005, held in Suntec International Convention and Exhibition Centre, Singapore, from 9 May 2005 till 12 May 2005, represents the fine collection of 540 refereed papers selected from nearly 2,700 submissions.

Computational Science has firmly established itself as a vital part of many scientific investigations, affecting researchers and practitioners in areas ranging from applications such as aerospace and automotive, to emerging technologies such as bioinformatics and nanotechnologies, to core disciplines such as mathematics, physics, and chemistry. Due to the shear size of many challenges in computational science, the use of supercomputing, parallel processing, and sophisticated algorithms is inevitable and becomes a part of fundamental theoretical research as well as endeavors in emerging fields. Together, these far reaching scientific areas contribute to shape this Conference in the realms of state-of-the-art computational science research and applications, encompassing the facilitating theoretical foundations and the innovative applications of such results in other areas.

The topics of the refereed papers span all the traditional as well as emerging Computational Science realms, and are structured according to six main conference themes:

- Computational Methods and Applications
- High Performance Computing, Networks and Optimisation
- Information Systems and Information Technologies
- Scientific Visualisation, Graphics and Image Processing
- Computational Science Education
- Advanced and Emerging Applications

In addition, papers from 27 Workshops and Technical Sessions on specific topics of interest, including information security, mobile communication, grid computing, modeling, optimization, computational geometry, virtual reality, symbolic computations, molecular structures, web systems and intelligence, spatial analysis, bioinformatics and geocomputations, to name a few, complete this comprehensive collection.

The warm reception of the great number of researchers to present high quality papers in ICCSA 2005 has taken the Conference to record new heights. The continuous support of Computational Science researchers has helped build ICCSA to be a firmly established forum in this area. We look forward to building on this symbiotic relationship together to grow ICCSA further.

We recognize the contribution of the International Steering Committee and we deeply thank the International Program Committee for their tremendous support in putting this Conference together, nearly nine hundred referees for

their diligent work, and the Institute of High Performance Computing, Singapore for their generous assistance in hosting the event.

We also thank our sponsors for their continuous support without which this Conference would not be possible.

Finally, we thank all authors for their submissions and all Invited Speakers and Conference attendants for making the ICCSA Conference truly one of the premium events in the scientific community, facilitating exchange of ideas, fostering new collaborations, and shaping the future of the Computational Science.

May 2005

Marina L. Gavrilova
Osvaldo Gervasi

on behalf of the co-editors:

Vipin Kumar
Antonio Laganà
Heow Pueh Lee
Youngsong Mun
David Taniar
Chih Jeng Kenneth Tan

Organization

ICCSA 2005 was organized by the Institute of High Performance Computing (Singapore), the University of Minnesota (Minneapolis, MN, USA), the University of Calgary (Calgary, CA) and the University of Perugia (Italy).

Conference Chairs

Vipin Kumar (Army High Performance Computing Center and University of Minnesota, USA), Honorary Chair
Marina L. Gavrilova (University of Calgary, Canada), Conference co-Chair, Scientific
Osvaldo Gervasi (University of Perugia, Italy), Conference co-Chair, Program
Jerry Lim (Institute of High Performance Computing, Singapore), Conference co-Chair, Organizational

International Steering Committee

Alexander V. Bogdanov (Institute for High Performance Computing and Information Systems, Russia)
Marina L. Gavrilova (University of Calgary, Canada)
Osvaldo Gervasi (University of Perugia, Italy)
Kurichi Kumar (Institute of High Performance Computing, Singapore)
Vipin Kumar (Army High Performance Computing Center and University of Minnesota, USA)
Andres Iglesias (University de Cantabria, Spain)
Antonio Laganà (University of Perugia, Italy)
Heow Pueh Lee (Institute of High Performance Computing, Singapore)
Youngsong Mun (Soongsil University, Korea)
Chih Jeng Kenneth Tan (OptimaNumerics Ltd, and The Queen's University of Belfast, UK)
David Taniar (Monash University, Australia)

Local Organizing Committee

Kurichi Kumar (Institute of High Performance Computing, Singapore)
Heow Pueh Lee (Institute of High Performance Computing, Singapore)

Workshop Organizers

Approaches or Methods of Security Engineering

Haeng Kon Kim (Catholic University of Daegu, Korea)
Tai-hoon Kim (Korea Information Security Agency, Korea)

Authentication, Authorization and Accounting

Eui-Nam John Huh (Seoul Women's University, Korea)

Component Based Software Engineering and Software Process Model

Haeng Kon Kim (Catholic University of Daegu, Korea)

Computational Geometry and Applications (CGA'05)

Marina Gavrilova (University of Calgary, Calgary, Canada)

Computer Graphics and Geometric Modeling (TSCG'2005)

Andres Iglesias (University of Cantabria, Santander, Spain)
Deok-Soo Kim (Hanyang University, Seoul, Korea)

Computer Graphics and Rendering

Jiawan Zhang (Tianjin University, China)

Data Mining and Bioinformatics

Xiaohua Hu (Drexel University, USA)
David Taniar (Monash University, Australia)

Digital Device for Ubiquitous Computing

Hong Joo Lee (Daewoo Electronics Corp, Korea)

Grid Computing and Peer-to-Peer (P2P) Systems

Jemal H. Abawajy(Deakin University, Australia)
Maria S. Perez (Universitad Politecnica de Madrid, Spain)

Information and Communication Technology (ICT) Education

Woochun Jun (Seoul National University, Korea)

Information Security & Hiding, ISH 2005

Raphael C.W. Phan (Swinburne University of Technology, Malaysia)

Intelligent Multimedia Services and Synchronization in Mobile Multimedia Networks

Dong Chun Lee (Howon University, Korea)
Kuinam J Kim (Kyonggi University, Korea)

Information Systems Information Technologies (ISIT)

Youngsong Mun (Soongsil University, Korea)

Internet Comunications Security (WICS)

Josè Sierra-Camara (University Carlos III of Madrid, Spain)
Julio Hernandez-Castro (University Carlos III of Madrid, Spain)
Antonio Izquierdo (University Carlos III of Madrid, Spain)
Joaquin Torres (University Carlos III of Madrid, Spain)

Methodology of Information Engineering

Sangkyun Kim(Somansa Co. Ltd, Korea)

Mobile Communications

Hyunseung Choo (Sungkyunkwan University, Korea)

Modelling Complex Systems

Heather J. Ruskin (Dublin City University, Ireland)
Ruili Wang (Massey University, New Zealand)

Modeling of Location Management in Mobile Information Systems

Dong Chun Lee (Howon University, Korea)

Molecular Structures and Processes

Antonio Laganà (University of Perugia, Perugia, Italy)

Optimization: Theories and Applications (OTA) 2005

In-Jae Jeong (Hanyang University, Korea)
Dong-Ho Lee (Hanyang University, Korea)
Deok-Soo Kim (Hanyang University, Korea)

Parallel and Distributed Computing

Jiawan Zhang (Tianjin University, Tianjin, China)

Pattern Recognition & Ubiquitous Computing

Woongjae Lee (Seoul Women's University, Korea)

Spatial Analysis and GIS: Local or Global?

Stefania Bertazzon (University of Calgary, Calgary, Canada)
Borruso Giuseppe (University of Trieste, Trieste, Italy)
Falk Huettmann(Institute of Arctic Biology, USA)

Specific Aspects of Computational Physics for Modeling Suddenly-emerging Phenomena

Paul E. Sterian (Politehnica University, Romania)
Cristian Toma (Titu Maiorescu University, Romania)

Symbolic Computation, SC 2005

Andres Iglesias (University of Cantabria, Spain)
Akemi Galvez (University of Cantabria, Spain)

Ubiquitous Web Systems and Intelligence

David Taniar (Monash University, Australia)
Wenny Rahayu (La Trobe University, Australia)

Virtual Reality in Scientific Applications and Learning, VRSAL 2005

Osvaldo Gervasi (University of Perugia, Perugia, Italy)
Antonio Riganelli (University of Perugia, Perugia, Italy)

Program Committee

Jemal Abawajy (Deakin University, Australia)
Kenny Adamson (EZ-DSP, UK)
Srinivas Aluru (Iowa State University, USA)
Frank Baetke (Hewlett Packard, USA)
Mark Baker (Portsmouth University, UK)
Young-Cheol Bang (Korea Polytech University, Korea)
David Bell (The Queen's University of Belfast, UK)
Stefania Bertazzon (University of Calgary, Canada)
Sergei Bespamyatnikh (Duke University, USA)
J. A. Rod Blais (University of Calgary, Canada)
Alexander V. Bogdanov (Institute for High Performance Computing and Information Systems, Russia)
Richard P. Brent(University of Oxford, UK)
Peter Brezany (University of Vienna, Austria)
Herve Bronnimann (Polytechnic University, NY, USA)
John Brooke (The University of Manchester, UK)
Martin Buecker (Aachen University, Germany)
Rajkumar Buyya (University of Melbourne, Australia)
YoungSik Choi (University of Missouri, USA)
Hyunseung Choo (Sungkyunkwan University, Korea)

Bastien Chopard (University of Geneva, Switzerland)
Min Young Chung (Sungkyunkwan University, Korea)
Toni Cortes (Universidad de Catalunya, Barcelona, Spain)
Yiannis Cotronis (University of Athens, Greece)
Danny Crookes (The Queen's University of Belfast, UK)
Josè C. Cunha (New University of Lisbon, Portugal)
Brian J. d'Auriol (University of Texas at El Paso, USA)
Alexander Degtyarev (Institute for High Performance Computing and Data Bases, Russia)
Frédéric Desprez (INRIA, France)
Tom Dhaene (University of Antwerp, Belgium)
Beniamino Di Martino (Second University of Naples, Italy)
Hassan Diab(American University of Beirut, Lebanon)
Ivan Dimov (Bulgarian Academy of Sciences, Bulgaria)
Iain Duff (Rutherford Appleton Laboratory, UK and CERFACS, France)
Thom Dunning (NCSA, USA)
Fabrizio Gagliardi (CERN, Switzerland)
Marina L. Gavrilova (University of Calgary, Canada)
Michael Gerndt (Technical University of Munich, Germany)
Osvaldo Gervasi (University of Perugia, Italy)
Bob Gingold (Australian National University, Australia)
James Glimm (SUNY Stony Brook, USA)
Christopher Gold (Hong Kong Polytechnic University, Hong Kong)
Yuriy Gorbachev (Institute of High Performance Computing and Information Systems, Russia)
Andrzej Goscinski (Deakin University, Australia)
Jin Hai (Huazhong University of Science and Technology, China)
Ladislav Hlucky (Slovak Academy of Science, Slovakia)
Shen Hong (Japan Advanced Institute of Science and Technology, Japan)
Paul Hovland (Argonne National Laboratory, USA)
Xiaohua Hu (Drexel University, USA)
Eui-Nam John Huh (Seoul Women's University, Korea)
Terence Hung (Institute of High Performance Computing, Singapore)
Andres Iglesias (University de Cantabria, Spain)
In-Jae Jeong (Hanyang University, Korea)
Elisabeth Jessup (University of Colorado, USA)
Peter K. Jimack (University of Leeds, UK)
Christopher Johnson (University of Utah, USA)
Benjoe A. Juliano (California State University, Chico, USA)
Peter Kacsuk (MTA SZTAKI Research Instiute, Hungary)
Kyung Woo Kang (KAIST, Korea)
Carl Kesselman (University of Southern California, USA)
Daniel Kidger (Quadrics, UK)
Deok-Soo Kim (Hanyang University, Korea)
Haeng Kon Kim (Catholic University of Daegu, Korea)

Jin Suk Kim (KAIST, Korea)
Tai-hoon Kim (Korea Information Security Agency, Korea)
Yoonhee Kim (Syracuse University, USA)
Mike Kirby (University of Utah, USA)
Jacek Kitowski (AGH University of Science and Technology, Poland)
Dieter Kranzlmueller (Johannes Kepler University Linz, Austria)
Kurichi Kumar (Institute of High Performance Computing, Singapore)
Vipin Kumar (Army High Performance Computing Center and University of Minnesota, USA)
Domenico Laforenza (Italian National Research Council, Italy)
Antonio Laganà (University of Perugia, Italy)
Joseph Landman (Scalable Informatics LLC, USA)
Francis Lau (The University of Hong Kong, Hong Kong)
Bong Hwan Lee (Texas A&M University, USA)
Dong Chun Lee (Howon University, Korea)
Dong-Ho Lee (Hanyang University, Korea)
Heow Pueh Lee (Institute of High Performance Computing, Singapore)
Sang Yoon Lee (Georgia Institute of Technology, USA)
Tae Jin Lee (Sungkyunkwan University, Korea)
Bogdan Lesyng (ICM Warszawa, Poland)
Zhongze Li (Chinese Academy of Sciences, China)
Laurence Liew (Scalable Systems Pte, Singapore)
David Lombard (Intel Corporation, USA)
Emilio Luque (University Autonoma of Barcelona, Spain)
Michael Mascagni (Florida State University, USA)
Graham Megson (University of Reading, UK)
John G Michopoulos (US Naval Research Laboratory, USA)
Edward Moreno (Euripides Foundation of Marilia, Brazil)
Youngsong Mun (Soongsil University, Korea)
Jiri Nedoma (Academy of Sciences of the Czech Republic, Czech Republic)
Genri Norman (Russian Academy of Sciences, Russia)
Stephan Olariu (Old Dominion University, USA)
Salvatore Orlando (University of Venice, Italy)
Robert Panoff (Shodor Education Foundation, USA)
Marcin Paprzycki (Oklahoma State University, USA)
Gyung-Leen Park (University of Texas, USA)
Ron Perrott (The Queen's University of Belfast, UK)
Dimitri Plemenos (University of Limoges, France)
Richard Ramaroson (ONERA, France)
Rosemary Renaut (Arizona State University, USA)
Alexey S. Rodionov (Russia Academy of Science, Russia)
Paul Roe (Queensland University of Technology, Australia)
Reneé S. Renner (California State University at Chico, USA)
Heather J. Ruskin (Dublin City University, Ireland)
Ole Saastad (Scali, Norway)

Muhammad Sarfraz (King Fahd University of Petroleum and Minerals, Saudi Arabia)
Edward Seidel (Louisiana State University, USA and Albert Einstein Institut, Germany)
Josè Sierra-Camara (University Carlos III of Madrid, Spain)
Dale Shires (US Army Research Laboratory, USA)
Vaclav Skala (University of West Bohemia, Czech Republic)
Burton Smith (Cray, USA)
Masha Sosonkina (University of Minnesota, USA)
Alexei Sourin (Nanyang Technological University, Singapore)
Elena Stankova (Institute for High Performance Computing and Data Bases, Russia)
Gunther Stuer (University of Antwerp, Belgium)
Kokichi Sugihara (University of Tokyo, Japan)
Boleslaw Szymanski (Rensselaer Polytechnic Institute, USA)
Ryszard Tadeusiewicz (AGH University of Science and Technology, Poland)
Chih Jeng Kenneth Tan (OptimaNumerics, UK and The Queen's University of Belfast, UK)
David Taniar (Monash University, Australia)
John Taylor (Quadrics, UK)
Ruppa K. Thulasiram (University of Manitoba, Canada)
Pavel Tvrdik (Czech Technical University, Czech Republic)
Putchong Uthayopas (Kasetsart University, Thailand)
Mario Valle (Visualization Group, Swiss National Supercomputing Centre, Switzerland)
Marco Vanneschi (University of Pisa, Italy)
Piero Giorgio Verdini (University of Pisa and Istituto Nazionale di Fisica Nucleare, Italy)
Jesus Vigo-Aguiar (University of Salamanca, Spain)
Jens Volkert (University of Linz, Austria)
Koichi Wada (University of Tsukuba, Japan)
Kevin Wadleigh (Hewlett Packard, USA)
Jerzy Wasniewski (Technical University of Denmark, Denmark)
Paul Watson (University of upon Tyne)
Jan Weglarz (Poznan University of Technology, Poland)
Tim Wilkens (Advanced Micro Devices, USA)
Roman Wyrzykowski (Technical University of Czestochowa, Poland)
Jinchao Xu (Pennsylvania State University, USA)
Chee Yap (New York University, USA)
Osman Yasar (SUNY at Brockport, USA)
George Yee (National Research Council and Carleton University, Canada)
Yong Xue (Chinese Academy of Sciences, China)
Igor Zacharov (SGI Europe, Switzerland)
Xiaodong Zhang (College of William and Mary, USA)
Aledander Zhmakin (SoftImpact, Russia)

Krzysztof Zielinski (ICS UST / CYFRONET, Poland)
Albert Zomaya (University of Sydney, Australia)

Sponsoring Organizations

University of Perugia, Perugia, Italy
University of Calgary, Calgary, Canada
University of Minnesota, Minneapolis, USA
The Queen's University of Belfast, UK
Society for Industrial and Applied Mathematics, USA
The Institution of Electrical Engineers, UK
OptimaNumerics Ltd, UK
MASTER-UP, Italy
The Institute of High Performance Computing, Singapore

Table of Contents – Part III

Grid Computing and Peer-to-Peer (P2P) Systems Workshop

Spatial Analysis and GIS: Local or Global? Workshop

Computer Graphics and Rendering Workshop

Data Mining and Bioinformatics Workshop

Parallel and Distribuited Computing Workshop

Symbolic Computation, SC 2005 Workshop

Specific Aspects of Computational Physics for Modeling Suddenly-Emerging Phenomena Workshop

Internet Comunications Security (WICS) Workshop

Component Based Software Engineering and Software Process Model Workshop

Computer Graphics and Geometric Modeling (TSCG 2005) Workshop

Methodology of Information Engineering Workshop

Table of Contents – Part IV

Information and Communication Technology (ICT) Education Workshop

Digital Device for Ubiquitous Computing Workshop

Optimization: Theories and Applications (OTA) 2005 Workshop

General

Tracks

Table of Contents – Part I

Information Systems and Information Technologies (ISIT) Workshop

Mobile Communications (Mobicomm) Workshop

Authentication Authorization Accounting (AAA) Workshop

Computational Geometry and Applications (CGA '05) Workshop

Virtual Reality in Scientific Applications and Learning (VRSAL 2005) Workshop

Molecular Structures and Processes Workshop

Pattern Recognition and Ubiquitous Computing Workshop

Table of Contents – Part II

Approaches or Methods of Security Engineering Workshop

Information Security and Hiding (ISH 2005) Workshop

Modeling of Location Management in Mobile Information Systems Workshop

Intelligent Multimedia Services and Synchronization in Mobile Multimedia Networks Workshop

Ubiquitous Web Systems and Intelligence Workshop

Modelling Complex Systems Workshop

Resource and Service Discovery in the iGrid Information Service

Giovanni Aloisio, Massimo Cafaro, Italo Epicoco, Sandro Fiore, Daniele Lezzi, Maria Mirto, and Silvia Mocavero

Center for Advanced Computational Technologies,
University of Lecce/ISUFI, Italy
{giovanni.aloisio, massimo.cafaro, italo.epicoco, sandro.fiore, daniele.lezzi, maria.mirto, silvia.mocavero}@unile.it

Abstract. In this paper we describe resource and service discovery mechanisms available in iGrid, a novel Grid Information Service based on the relational model. iGrid is developed within the GridLab project by the ISUFI Center for Advanced Computational Technologies (CACT) at the University of Lecce, Italy and it is deployed on the European GridLab testbed. The GridLab Information Service provides fast and secure access to both static and dynamic information through a GSI enabled web service. Besides publishing system information, iGrid also allow publication of user's or service supplied information. The adoption of the relational model provides a flexible model for data, and the hierarchical distributed architecture provides scalability and fault tolerance.

1 Introduction

The grid computing paradigm is widely regarded as a new field, distinguished from traditional distributed computing because of its main focus on large-scale resource sharing and innovative high-performance applications [1]. The grid infrastructure can be seen as an ensemble of Virtual Organizations (VOs), reflecting dynamic collections of individuals, institutions and computational resources [2].

In this context, achieving flexible, secure and coordinated resource sharing among participating VOs requires the availability of an information rich environment to support resource and service discovery, and thus decision making processes. Indeed, we think of distributed computational resources, services and VOs as sources and/or potential sinks of information. The data produced can be static or dynamic in nature, or even dynamic to some extent. Depending on the actual degree of dynamism, information is better handled by a Grid Information Service (static or quasi-static information) or by a Monitoring Service (highly dynamic information).

In this context, information plays a key role, therefore in turn Grid Information Services are a fundamental building block of grid infrastructure/middleware. Indeed, high performance execution in grid environments is virtually impossible

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1–9, 2005.

without timely access to accurate and up-to-date information related to distributed resources and services. Our experience has shown that using manual of default configurations hinders application performance, and the lack of information about the execution environment prevents design and implementation of grid-aware applications. As a matter of fact, an application can react to changes in its environment only if these changes are advertised. Therefore, self-adjusting, adaptive applications are natural consumers of information produced in grid environments. Neverthless, making relevant information available on-demand to consumer applications is actually nontrivial, since information can be (i) diverse in scope, (ii) dynamic and (iii) distributed across one or more VOs. Moreover, obtaining information about the structure and state of grid resources, services, networks etc. can be challenging in large scale grid environments.

The rest of the paper is organized as follows. We recall related work in Section 2, and discuss resource and service discovery mechanisms available in iGrid in Section 3. Finally, we conclude the paper in Section 4.

2 Related Work

2.1 Resource Discovery

Historically, flooding protocols [3] have been used for resource discovery in networks; later, gossiping protocols such as [4] have been shown to be more efficient for information dissemination. Recently, the problem of efficient resource and service discovery received a great deal of attention. Essentially, the current focus is on structured and/or unstructured peer-to-peer systems. Iamnitchi and Foster [5] propose some heuristic solutions for decentralized distributed resource discovery, and define a taxonomy based on membership protocol, overlay construction, preprocessing and request processing. Butt et al. [6] describe the use of Condor pools coupled with the Pastry [7] distributed hash table algorithm, and utilize the peer-to-peer overlay network to attain scalable resource discovery.

Other approaches are based on the notion of non uniform information dissemination, such as the work of Iyengar et al. [8]. The key idea is to update replicated information services using non uniform instead of full dissemination of information. This leads to reduced information replication overhead, maintaining anyway accurate information at locations where it is most likely to be needed. The authors observe that grid resources share some properties of sensor networks, and thus tend to be of more interest to nearby users, because the overhead in starting a job and transferring related data and results increases with the distance to the resources. The dissemination protocols proposed work by propagating resource state information more aggressively and in more detail to nearer information repositories than they do to farther ones. This leads to the Change Sensitive Protocol and Prioritized Dissemination Protocol. Maheswaran et. al. [9] share the same concept, introducing in their work the notion of grid potential, which is used to weight a grid resource capability with its distance from the application launch point.

The majority of the works deals with variation of distributed hash table algorithms for building scalable infrastructures. We recall here CAN [10], Chord [12], Kademlia [13], Tapestry [11], Viceroy [14]. The CONE data structure by Bhagwan et al. [15], and the peer-to-peer information retrieval system designed by Schmidt et al. [16] using Hilbert Space Filling Curve and the Chord overlay network topology, specifically address distributed resource discovery. We conclude this subsection citing the Globus Toolkit Metadata Directory Service (MDS v2) [17], an LDAP based approach characterized by mechanisms and protocols for storing information and building scalable distributed collections of servers. No specific organization strategies, overlay topologies, or dissemination protocols have been specified. This service has been deployed and used on many grid projects, both for production and development. However, the performances of the MDS Information Service were not satisfactory enough, and the Globus project provided users with a new version of this service, available in the Globus Toolkit version 3.x and called Monitoring and Discovery Service (again, MDS v3). Version 3.x of the Globus Toolkit has been based on the Open Grid Service Infrastructure (OGSI) and the Open Grid Service Architecture [19] specifications. The MDS has been developed using Java and the Grid Services framework [18], so that, again, its performances are not satisfactory enough. The MDS service will change once again in the Toolkit version 4, and it will be based on the emerging Web Service Resource Framework (WSRF) specification [20].

It is important to understand that while many versions of the Globus Toolkit exist, only version 2.x and version 3.x pre-OGSI have been widely deployed for use in production grids and testbeds. Indeed, version 3.x based on OGSI has been targeted only by developers, since this version was known to be replaced next year by the new WRSF based Toolkit, which will also offer the possibility to develop grid services and clients using C/C++ besides Java.

The GridLab project started adopting and extending initially the Globus MDS [32], but we decided to move from LDAP to the relational model, in order to overcome the Globus MDS shortcomings: a data model better suited for frequent reading, not able to cope with frequent updates, a weak query language missing key operations on data (e.g., there is no *join*), and performances. The relational data model allows us to model information and store data in tabular form, relationships among tables are possible, typical relational DBMS are strongly optimized for both fast reading and writing, and finally a good query language, namely SQL, allows complex operations on data, including joins.

2.2 Service Discovery

Historically, service discovery has been initially addressed by the ANSA project [21], an early industry effort to define trading services to advertise and discover relevant services. As an evolution, we recall here the CORBA trading service [22]. Recently, yet another industry standard is emerging in the context of web services, UDDI [23]. This is a centralized registry storing WSDL description of business oriented web services. A grid oriented approach, called Web Services Inspection Language (WSIL) [24] has been proposed by IBM specifically for grid services, as envisioned by the OGSA architecture.

The Web Service Discovery Architecture (WSDA) proposed by Hoschek [25], specifies communication primitives useful for discovery, service identification, service description retrieval, data publication as well as query support. The individual primitives when combined and plugged together by clients and services yield a wide range of behaviors. The author also introduces an hyper registry, a centralized database node for discovery of dynamic distributed content. This registry supports XQueries over a tuple set from a dynamic XML data model. The architecture includes a Unified Peer-to-Peer Database Framework (UPDF) and a corresponding Peer Database Protocol (PDP).

The Jini Lookup Service [26] is a centralized service approach which exploits multicast and soft state protocols. However, it is worth noting here that the protocols rely on Java object serialization, therefore the service usefulness is limited to Java clients. The Java JXTA peer-to-peer network is a notable effort; we recall here its stateless, best-effort protocols for ad hoc, pervasive, and multi-hop P2P computing. JXTA includes, among others, the following protocols. The Endpoint Routing Protocol allows discovering routes from one peer to another peer, given the destination peer ID. The Rendezvous Protocol provides publish and subscribe functionalities within a peer group. The Peer Resolver Protocol and Peer Discovery Protocol allow publishing advertisements and simple queries that are unreliable, stateless, non-pipelined, and non-transactional.

The Service Location Protocol (SLP) [27] can be used to advertise and query the location, type and attributes of a given service; it uses multicast and soft state protocols. A recent extension, called Mesh Enhanced Service Location Protocol (mSLP) [28], increases scalability through the use of multiple cooperating directory agents. The Service Discovery Service (SDS) [29] supports XML based exact match queries, and is based on multicast and soft state protocols too.

3 Resource and Service Discovery in iGrid

iGrid is a novel Grid Information Service developed within the European GridLab project [30] by the ISUFI Center for Advanced Computational Technologies (CACT) at the University of Lecce, Italy. An overview of the iGrid Information Service can be found in [33]; here we delve into details related specifically to resource and service discovery mechanisms available.

iGrid distributed architecture is shown in Fig. 1. The architecture is based on iServe and iStore GSI [31] enabled web services. An iServe collects information related to the computational resource it is installed on, while iStore gathers information coming from trusted, registered iServes. The current architecture resembles the one adopted by the Globus Toolkit MDS, therefore iStores are allowed to register themselves to other iStores, creating arbitrarily complex distributed hierarchies. Even though this architecture proved to be effective to build scalable distributed collections of servers, neverthless we are already investigating peer-to-peer overlay networks based on current state of the art distributed hash table algorithms in order to improve iGrid scalability. The implementation includes system information providers outputting XML, while trusted users

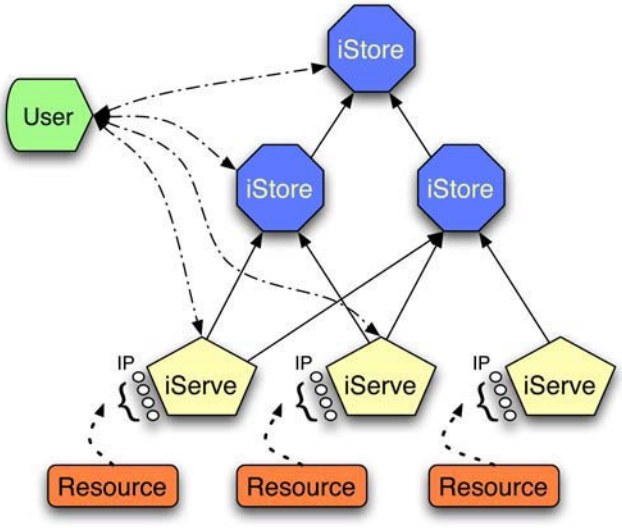


Fig. 1. iGrid hierarchical architecture

and/or services can publish information simply calling a web service registration method.

Resource discovery using the iGrid Information Service is based on the availability of the following information (not exaustive):

System operating system, release version, machine architecture etc;

CPU for CPUs, static information such as model, vendor, version, clock speed is extracted; the system also provides dynamic information such as idle time, nice time, user time, system time and load;

Memory static information such as RAM amount and swap space is available. Dynamic information related to available memory and swap space is published too;

File Systems static as well dynamic information is extracted, such as file system type, mount point, access rights, size and available space;

Network Interfaces network interface names, network addresses and network masks;

Local Resource Manager the information belonging to this category can be further classified as belonging to three different subclasses: information about queues, jobs and static information about Local resource Management System (LRMS). Some examples of extracted information are: LRMS type and name; queue name and status, number of CPU assigned to the queue, maximum number of jobs that can be queued, number of queued jobs, etc; job name, identifier, owner, status, submission time etc. Currently information providers for OpenPBS and Globus Gatekepeer are available, with LSF planned;

Certification Authorities certificate subject name, serial number, expiration date, issuer, public key algorithm etc.

Virtual Organization information related to VO can be used to automatically discover which resources belong to a given VO; we have VO name, resource type, help desk phone number, help desk URL, job manager, etc.

Of course, this set of information is not meant to be static, the iGrid schema will continue to evolve and will be extended to support additional information as required by the GridLab project or iGrid users.

One of the most important requirements for grid computing scenarios is the ability to discover services and web/grid services dynamically. Services in this context refers to traditional unix servers. The iGrid system provides users and developers with the following functionalities: register, unregister, update and lookup. More than one instance for each service or web service can be registered. The following information is available for services: logical name, instance name, service description, default port, access URL, distinguished name of the service publisher, timestamps related to date of creation an date of expiration of the published information.

For web services, relevant information includes logical name, web service description, WSDL location (URL), web service access URL, distinguished name of publisher and timestamps related to date of creation and date of expiration of the published information.

Information related to firewalls is strictly related to service information. As a matter of fact, before registering a service, developers will query iGrid to retrieve the range of open ports available on a specified computational resource. This is required in order to chose an open port, allowing other people/services to connect to a registered service. The information available includes firewall hostname, open ports, time frame during which each port (or a range of ports) is open, the protocol (TCP/UDP) used to connect to these ports, the distinguished name of the firewall administrator, and timestamps related to date of creation and date of expiration of the published information.

iGrid uses a push model for data exchange. Indeed, system information (useful for resource discovery) extracted from resources is stored on the local database, and periodically sent to registered iStores, while user and/or service supplied information (useful for service discovery) is stored on the local database and immediately sent to registered iStores. Thus, an iStore has always fresh, updated information related to services, and almost fresh information related to resources; it does not need to ask iServes for information. The frequency of system information forwarding is based on the information itself, but we also allow defining a per information specific policy. Currently, system information forwarding is based on the rate of change of the information itself. As an example, information that does not change frequently or change slowly (e.g. the amount of RAM installed) does not require a narrow update interval. Interestingly, this is true even for the opposite extreme, i.e., for information changing rapidly (e.g., CPU load), since it is extremely unlikely that continuous forwarding of this kind of

information can be valuable for users, due to information becoming quickly inaccurate. Finally, information whose rate of change is moderate is forwarded using narrow update intervals. Allowing user's defined policies provides a forwarding protocol similar to the Prioritized Dissemination Protocol, whilst our update frequencies closely resemble the way the Change Sensitive Protocol works.

We have found that the push model works much better than the corresponding pull model (adopted, for instance, by the Globus Toolkit MDS) in grid environments. This is due to the small network traffic volume generated from iServe to iStore servers: on average, no more than one kilobyte of data must be sent. Moreover, we tag information with a time to live attribute that allows iGrid to safely removes stale information from the database when needed. For instance, when users search for data, a clean-up operation is performed before returning to the client the requested information, and during iGrid system startup, the entire database is cleaned up. Therefore the user will never see stale information.

Finally, it is worth recalling here that the performances of iGrid are extremely good, as reported in [33] .

4 Conclusion

We have described iGrid, a novel Grid Information Service based on the relational model, and its mechanisms for resource and service discovery. The GridLab Information Service provides fast and secure access to both static and dynamic information through a GSI enabled web service. Besides publishing system information, iGrid also allows publication of user's or service supplied information. The adoption of the relational model provides a flexible model for data, and the hierarchical distributed architecture provides scalability and fault tolerance. The software, which is open source, is freely available and can be downloaded from the GridLab project web site [34].

Acknowledgements

We gratefully acknowledge support of the European Commission 5th Framework program, grant IST-2001-32133, which is the primary source of funding for the GridLab project.

References

1. Foster, I., Kesselman C.: The Grid: Blueprint for a new Computing Infrastructure, Morgan Kaufmann (1998)
2. Foster, I., Kesselmann, C., Tuecke, S.: The Anatomy of the Grid: Enabling Scalable Virtual Organizations. Intl. Journal of High Performance Computing Applications **15**(3) (2001) 200–222. http://www.globus.org/- research/papers/anatomy.pdf.
3. Harcol-Balter, M., Leighton, P., and Lewin, D.: Resource discovery in distributed networks. In Proc. of ACM Symp. on Principles of Distributed Computing (1999) 229–238.

4. Li, L., Halpern, J., and Haas, Z.: Gossip-based ad hoc routing. In IEEE Infocom (2002) 1707–1716.
5. Iamnitchi, A., and Foster. I.: A Peer-to-Peer Approach to Resource Location in Grid Environments. Proceedings of the 11th Symposium on High Performance Distributed Computing (HPDC'02) (2002) 419.
6. Butt, A.R., Zhang, R., and Hu, Y.C.: A Self-Organizing Flock of Condors. In Proceedings of IEEE/ACM SC2003 (2003).
7. Rowstron, A., Druschel, P.: Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In Proceedings of the 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware) (2001) 329–350. http://freepastry.rice.edu
8. Iyengar, V., Tilak, S., Lewis, M.J., and Abu-Ghazaleh, N.B.: Non-Uniform Information Dissemination for Dynamic Grid Resource Discovery. IEEE NCA 2004 (2004) 97–106.
9. Maheswaran, M., and Krauter, K.: A parameter-based approach to resource discovery in grid computing system. In Proceedings of the First IEEE/ACM International Workshop on Grid Computing (2000) 181–190.
10. Ratnasamy, S., Francis, P., Handley, M., Karp, R., Shenker, S.: A scalable content-addressable network. in Proceedings of ACM SIGCOMM MANCA VOLUME(2001) PAGINE.
11. Hildrum, K., Kubiatowicz, J., Rao, S., Zhao, B.: Distributed Object Location in a Dynamic Network. In Proceedings of 14th ACM Symp. on Parallel Algorithms and Architectures (SPAA'02) (2002) 41–52.
12. Stoica, I. et al: Chord: A scalable peer-to-peer lookup service for Internet applications. In Proceedings of ACM SIGCOMM (2001) 149–160. http://www.pdos.lcs.mit.edu/chord.
13. Maymounkov, P., Mazieres, D.: Kademlia: A peer-to-peer information system based on the XOR metric. In Proceedings of the 1st International Workshop on Peer-to-Peer Systems, Springer-Verlag (2002). http://kademia.scs.cs.nyu.edu/
14. Malkhi, D., Naor, M., Ratajczak, D.: Viceroy: A scalable and dynamic emulation of the butterfly. In Proceedings of ACM Principles of Distributed Computing(PODC'02) (2002) 183–192.
15. Bhagwan, R., Varghese, G., and Voelker, G.M.: Cone: Augmenting DHTs to Support Distributed Resource Discovery. In Proceedings of CS2003 (2003).
16. Schmidt, C., and Parashar, M.: Flexible Information Discovery in Decentralized Distributed Systems. In Proceedings of HPDC (2003) 226–235.
17. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing, IEEE Press (2001) 181.
18. Foster, I., Kesselmann, C., Nick, J., Tuecke, S.: Grid Services for Distributed System Integration. Computer. **35**(6) (2002) 37–46.
19. Foster, I., Kesselmann, C., Nick, J., Tuecke, S.: The Physiology of the Grid: An Open Grid Services Architecture for Distributed System Integration. Technical Report for the Globus project. http://www.globus.org/research/papers/ogsa.pdf
20. The WSRF specification. http://www.globus.org/wsrf/specs/ws-wsrf.pdf
21. Beitz, A., Bearman, M., and Vogel, A.: Service Location in an Open Distributed Environment. In Proc. of the 2nd International Workshop on Services in Distributed and Networked Environments (1995) 28.
22. Object Management Group. Trading Object Service. OMG RPF5 Submission, May 1996.

23. UDDI Consortium. UDDI: Universal Description, Discovery and Integration. http://www.uddi.org.
24. IBM. Web Services Inspection Language (WS-Inspection) 1.0 specification. http://www-106.ibm.com/developerworks/webservices/library/ws-wsilspec.html.
25. Hoschek, W.: Peer-to-Peer Grid Databases for Web Service Discovery. Concurrency: Pract. Exper. (2002) 1–7.
26. Waldo, J.: The Jini architecture for network-centric computing. Communications of the ACM **42**(7) (1999) 76–82.
27. Guttman, E.: Service Location Protocol: Automatic Discovery of IP Network Services. IEEE Internet Computing Journal **3**(4) (1999) 71–80.
28. Zhao, W., Schulzrinne, H., and Guttman, E.: mSLP - Mesh Enhanced Service Location Protocol. In Proc. of the IEEE Int. Conf. on Computer Communications and Networks (ICCCN'00) (2000).
29. Czerwinski, S.E., Zhao, B.Y., Hodes, T., Joseph, A.D., and Katz, R.: An Architecture for a Secure Service Discovery Service. In Fifth Annual Int. Conf. on Mobile Computing and Networks (MobiCOM'99) (1999) 24–35.
30. The GridLab project. http://www.gridlab.org
31. Foster, I., Kesselmann, C., Tsudik G., Tuecke, S.: A security Architecture for Computational Grids. Proceedings of 5th ACM Conference on Computer and Communications Security Conference (1998) 83–92.
32. Aloisio G., Cafaro M., Epicoco I., Lezzi D., Mirto M., Mocavero S.: The Design and Implementation of the GridLab Information Service. Proceedings of The Second International Workshop on Grid and Cooperative Computing (GCC 2003)Lecture Notes in Computer Science, Springer-Verlag **3032** (2004) 131–138.
33. Aloisio, G., Cafaro, M., Epicoco, I., Fiore, S., Lezzi, D., Mirto, M. and Mocavero, S.: iGrid, a Novel Grid Information Service. To appear in Proceedings of the first European Grid Conference (EGC) 2005.
34. The iGrid Web Site. http://www.gridlab.org/WorkPackages/wp-10.

A Comparison of Spread Methods in Unstructured P2P Networks

Zhaoqing Jia, Bingzhen Pei, Minglu Li, and Jinyuan You

Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200030, China
jiazhaoqing@hotmail.com

Abstract. In recent years, unstructured Peer-to-Peer (P2P) applications are very popular on the Internet. Unstructured P2P topology has power-law characteristic in the link distribution, containing a few nodes that have a very high degree and many with low degree. This reflects the presence of central nodes which interact with many others and play a key role in relaying information. The system performance can be improved by replicating file location information to the high degree nodes in unstructured P2P. In this paper, we present an overview of several spread mechanisms for unstructured P2P and analyze the performance of them relative to the number of information replica and the bandwidth consumption. The optimal spread mechanism is high degree spread method. Simulation results empirically evaluate their behavior in direct comparison and verify our analysis.

1 Introduction

In the last few years, unstructured P2P applications are very popular, such as Gnutella [1], Kazaa [2]. There is no precise control over the network topology or file placement in these systems. They are designed for sharing files among the peers in the networks. In general, they employ flooding scheme for searching object, and waste a lot of bandwidth [1]. Today, bandwidth consumption attributed to these applications amounts to a considerable fraction (up to 60%) of the total Internet traffic [3]. It is of great importance to reduce the total traffic of them for the user and the broad Internet community.

An efficient technique for improving performance of search method is to replicate file location information in P2P. In [4□8], they study the relationship between number of copies of a file and its popularity. They examine uniform and proportional replication strategy, and present that optimal replication strategy is between them. They also point out that Square-root allocation can be achieved by path replication and random replication is better than path replication. In this paper, we focus on that where file location information copies are placed is more effective for search in unstructured P2P and how to replicate information copies to those places without global knowledge.

Unstructured P2P have power-law link distribution, such as Gnutella, containing a few nodes that have a very high degree and many with low degree [5]. This power law in the link distribution reflects the presence of central routers which interact with

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 10 – 18, 2005.

many others on a daily basis and play a key role in relaying information. Power-law networks are extremely vulnerable to targeted attack to high degree nodes and enhancing security of the high degree nodes can effectively improve the robustness of the network. The presence of the high degree nodes facilitates the spread of computer viruses, and increasing the anti-virus capacity of the high degree nodes is very helpful for controlling the spread of computer viruses [6,7].

Hence, these high degree nodes play very important role in unstructured P2P. The performance of search methods can be improved by replicating information copies to high degree nodes in unstructured P2P. In most case, ones have not global knowledge about the networks and how to replicate information to the important nodes with local knowledge. A simple method is flooding mechanism, but it wastes a lot bandwidth. A good spread mechanism can replicate information to the important nodes of the networks with low cost.

In this paper, we present that the performance of search method can be effectively improved by replicating file location information copies to high degree nodes in unstructured P2P. We present an overview of several spread mechanisms: flooding spread, percolation-based spread, random walk spread and high degree walk spread, and analyze their performance. We present a framework for evaluating the performance of a spread mechanism, simulate the presented spread mechanisms and present a direct comparison of their performance.

2 Our Framework

Unstructured P2P networks consist of many peers and links among these peers. Peers that are direct linked in P2P overlay are neighbors. Peers communicate when they forward messages. Messages include file location information and query message. Each message is assigned an identifier which enables peers to make distinction between new messages and duplicate ones received due to a cycle.

Each peer keeps a local collection of files, while it makes request for those it wishes to obtain. The files are stored at various nodes across the network. Peers and files are assumed to have unique identifier.

We defined the replica rate and the cost of a spread method. They are used to characterize the performance of the spread method. The ratio of cached file information on one node in the network to total spread file information is the replica rate of this node. The replica rate of the highest degree node in the network is used as the replica rate of the spread method. The cost of the spread method is given by the messages which each spread produces.

If more file information can be replicated to high degree nodes with one spread mechanism than with another spread mechanism at same cost, then the performance of the former mechanism is better than one of the latter mechanism.

Flooding search and random walk search are two typical search techniques, and are commonly used in many fields. Flooding search is very simple and each node forwards query messages to its all neighbors [1]. Random walk is a well-known technique, which forwards a query message to a randomly chosen neighbor at each step until the object is found [8]. This message is known as “random walker”. A search for a file in a network is successful if it discovers at least one replica of requested file

information. The ratio of successful to total searches made is the success rate of the algorithm. The cost of a search method is given by the messages that each search produces.

In order to examine the effectives of the spread mechanisms and verify our analysis, flooding search and random walk search are used to locate files in networks and we focus on the success rates and the costs of two search methods.

A global TTL parameter represents the maximum hop-distance a query message or file location information can reach before it gets discarded.

3 Spread Mechanisms

3.1 Flooding Spread Method

Flooding spread method is used to improve search in power-law network in Ref. [9]. When a node starts a spread, it sends file location information to its all neighbors with TTL=r, and each intermediate node forwards this message to its all neighbors. This message will be spread to those nodes within a certain radius r.

When the value TTL is enough high, file location information can spread through the whole network. It is sure that file location information is also stored on the high degree nodes. But it wastes a lot of bandwidth. If radius r is smaller, location information of files on some low degree nodes can not arrive at the high degree nodes.

3.2 Percolation-Based Spread Method

This approach is a variation of flooding spread method. In Ref. [10], the percolation-based method is employed for spreading query message in network. When a node starts a spread, it only sends file location information to a neighbor node with probability q.

This method certainly reduces average messages production compared to the flooding spread method. But it still contacts a large number of nodes. The most serious disadvantage of this method is that it is unfair to the files on the low degree nodes. For a node with degree k, the probability that the location information of files on the node is replicate to its neighbor nodes is given by $1-(1-q)^k$. It is obvious that the degrees of high degree nodes are high, and the location information of files on these nodes is easily spread over the network. But it is difficult to replicate the location information of files on the low degree nodes to the high degree nodes. In many cases, its performance is worse than flooding spread's.

3.3 Random Walk Spread Method

In ref [11], random walk is employed for spreading file location information. In this method, the file information is forwarded to a randomly chosen neighbor node at each step.

In networks, a random edge arrives at a node with probability proportional to the degree of the node, i.e., p1(k)~kp(k), where k is the degree of the node and p(k) is the probability that a randomly chosen node has degree k. It is obvious that the probability that a random walker walks through the high degree node is high. Hence, the file

location information can be easily replicate to the high degree nods through random walk spread method.

Message production is largely reduced compared to the standard flooding spread mechanism, and independent on the network size. Its disadvantage is that its performance is largely dependent on the degrees of the high degree nodes.

3.4 High Degree Walk Spread Method

In the Ref. [12], high degree walk spread method is used to spread file location information for improving search in unstructured peer-to-peer networks. This method is a variation of random walk spread. It is reasonable that one know the knowledge of its neighbors. In this method, when a node starts a spread, it sends file location information to the highest degree node of its neighbors. The file location information is always forwarded to the highest degree neighbor at each step.

Because the high degree node is intentionally chose at each step, the probability that high degree walker walks through the high degree nodes is higher than the one that random walker does. By high degree walk, more file location information copies is stored on the high degree nodes. Its cost is equal to the cost of the random walk spread method.

Because the probability that high degree walker arrives at a node relates the order of the degree of this node, its performance has little dependence on the degrees of the high degree nodes and it overcomes the shortage of the random walk spread method.

4 Simulations

4.1 Simulation Setup

In this section, we simulate the above four spread methods: Flooding Spread Method (FSM), Percolation-Based Spread Method (PBSM), Random Walk Spread Method (RWSM) and High Degree Walk Spread Method (HDWSM). For many real world networks have power-law exponent between 2 and 3, in simulation we used two power-law graph models: graph I12000 with τ=2.3, which is produced by Innet-3.0 [13], and graph P12000 with τ=2.1, which is produced by Pajek [14]. The node is randomly chose which wants to spread file information over the network and location information of 1000 files are spread in all simulations.

4.2 Simulation Results of Four Spread Methods

Fig.1 displays the simulation results of the above four spread methods. This figure shows that high degree walk spread method achieves the highest replica rate in all cases and it is the optimal spread method in unstructured P2P. With high degree walk spread method, file location information can be replicate to the high degree nodes at low cost. With percolation-based spread, the replica rate is the lowest in all cases, for location information of files on some low degree nodes can not spread over the network. When TTL is low, the replica rate of the random walk spread and flooding spread is also low. But when TTL is high, random walk spread and flooding spread

Table 1. Two Power-Law Graph Models

Graphs	Graph Size (N)	Maximum Degree (m)	Average Degree (<k>)	Exponent (τ)
I12000	12000	3159	5.26	2.3
P12000	12000	441	5.84	2.1

have better performance. The replica rate of random walk spread is little higher than the one of flooding spread at same cost.

For flooding spread method, as shown in Fig. 1(a) and Fig. 1(b), when TTL is low, location information of the files on the low degree nodes can not arrive at high degree nodes and its performance is poor. However, when TTL is high, it achieves high replica rate (Fig. 1(c) and Fig. 1(d)). Comparing Fig. 1(c) and Fig. 1(d), the replica rate of flooding spread method on graph I12000 is higher than one of it on the graph P12000, and this shows that it has more dependence on the degrees of high degree nodes. When the degrees of the high degree nodes are very high, the high degree nodes have much more neighbor nodes, and location information of more files can easily arrive at the high degree nodes.

In percolation-based spread, the location information of the files on the high degree nodes can easily spread over the whole network, but that on the low degree nodes is difficult. So, its performance is poor and its curves are close to horizontal line in all cases.

Comparing Fig. 1(c) with Fig. 1(d), when the degrees of the high degree nodes are very high, the random walk spread method has same replica rate as the high degree walk spread method. But, when the degrees of the high degree nodes are relative low, high degree walk spread has higher replicate rate. This shows that the performance of the random walk spread has more dependence on the degrees of the high degree nodes. But the high degree walk spread exhibits similar behavior on the two graphs, and this shows that its performance has little dependence on the degrees of the high degree nodes.

4.3 Simulation Results of Random Walk Search and Flooding Search

In this section, random walk search and flooding search are used to locate files on the graph I12000 and the graph P12000. For the average path length of a random graph is proportional to ln(N)/ln(<k>), where <k> is the average degree of the graph, TTL is set at ln(N). All files is equally popular and be accessed at same probability. Based on the simulation results of the Fig. 1(c) and Fig. 1(d), we simulate random walk search and flooding search on the two graphs.

Fig. 2 and Table 2 show that random walk search achieves the highest success rates with high degree walk spread method and the cost is the lowest, and it is the best choice for improving performance of random walk search. With the percolation-based spread method, the success rates are lowest on the two graphs and the cost is highest. The success rates with random walk spread is higher than ones with flooding spread and the cost is lower. The success rate with random walk spread method is almost same to ones with high degree walk spread method on the graph I12000, but it is

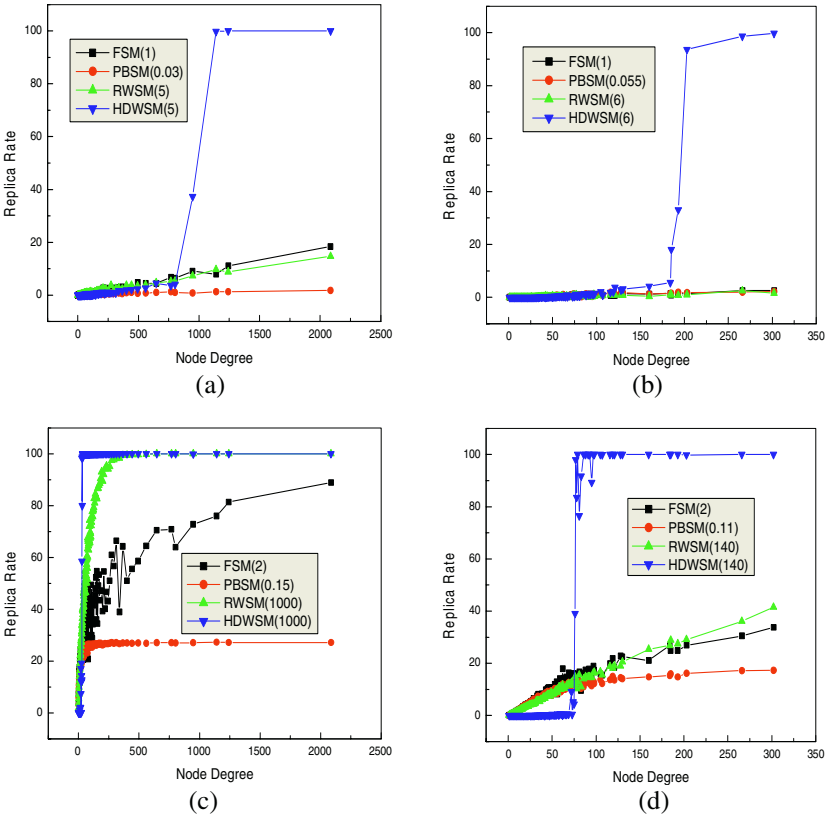


Fig. 1. Replica Rate vs Node Degree. (a) replica rate curves of four spread methods on graph I12000 at same cost (about 5 messages per spread). (b) replica rate curves on graph P12000 (about 6 messages per spread). (c) replica rate curves on graph I12000 (about 1000 messages per spread). (d) replica rate curves on graph P12000 (about 140 messages per spread)

Table 2. Cost of Random Walk Search at Different Spread Methods (Messages/Query)

Graphs	FSM	PBSM	RWSM	HDWSM
P12000	16.36	18.75	14.28	11.58
I12000	6.77	16.02	4.26	3.84

lower on the graph P12000. This shows that the performance of random walk spread method is dependent on the degrees of the high degree nodes.

The success rates on the graph I12000 are higher than ones on the graph P12000. The reason has twofold factors: the degrees of the high degree nodes of graph I12000 are much higher than ones of graph P12000 and the probability that random walk search walks through high degree nodes is largely dependent on the degrees of the high degree nodes.

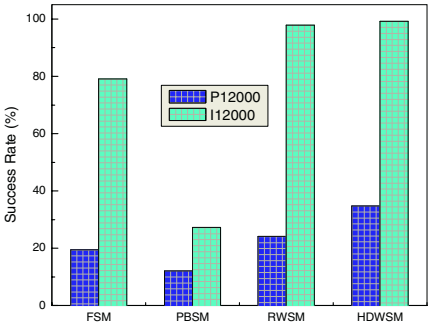


Fig. 2. Success Rates of Random Walk Search at Different Spread Methods

Fig. 3 illustrates the success rats with flooding search and shows that all success rates are almost equal. When TTL is set at ln(N), the query messages can almost spread through the whole network, so that almost all searches are successful with any spread method, and the success rates are almost 100%. Table 3 depicts the cost of flooding search with four different spread methods. The cost with percolation-based spread mechanism is much higher than one with anyone of the other spread methods. With random walk spread method, the cost of flooding search is similar to one with high degree walk spread on the graph I12000, but it is higher on the graph P12000. The cost with high degree walk spread is the least for each graph. Hence, the high degree walk spread method is the optimal spread mechanism for improving the performance of flooding search.

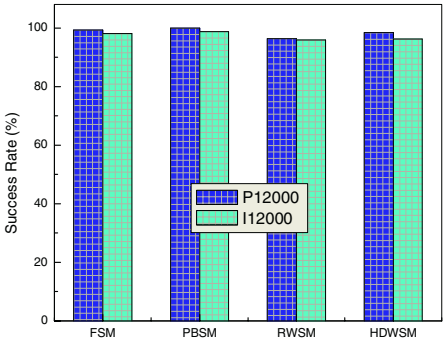


Fig. 3. Success Rates of Flooding Search at Different Spread Methods

Table 3. Cost of Flooding Search at Different Spread Methods (Messages/Query)

Graphs	FSM	PBSM	RWSM	HDWSM
P12000	902.3	18135.7	24.6	15.6
I12000	45.0	8381.0	2.1	2.0

In this section, the simulation results display that the performance of two search methods with high degree walk spread mechanism is best. The conclusion is consistent with the results of Sec. 4.2, and further validates our analysis.

5 Conclusion

Unstructured P2P have power-law in link distribution, and the high degree nodes play a key role in relaying information. Replicating file location information to the high degree nodes can effectively improve the performance of search methods in unstructured P2P.

We present an overview of several spread methods and analyze their performance. We present two metrics, replicate rate and cost, to characterize the performance of a spread method and the simulation results show they are useful for evaluating the performance of the spread method in unstructured P2P.

Simulation results of the four spread methods on the two graphs are consistent with our analysis. The performance of the percolation-based spread method is the worst, and its replicate rates are the lowest at all cases. The performance of random walk spread method is little better than one of flooding spread method, and they are largely dependent on the degrees of the high degree nodes. The high degree walk spread achieves great results in all cases, and can replicate more file location information to the high degree nodes at same cost. So, it is the optimal spread mechanism.

Finally, the simulation results of flooding search and random walk search show that the performance of search methods can be effectively improved by replicating the file location information to the high degree nodes and the high degree walk spread method is the best spread mechanism for improving search in unstructured P2P.

The high degree walk spread method is easily implemented in realistic settings. The conclusion is not only useful for unstructured P2P, but also for other power-law networks.

References

1. Gnutella website:http://gnutella.wego.com.
2. Kazaa website:http://www.kazaa.com.
3. Sandvine Inc. An Industry White Paper: The Impact of File Sharing on Service Provider Networks. Dec. 2002.
4. E. Cohen and S. Shenker. Replication Strategies in Unstructured Peer-to-Peer Networks. SIGCOMM'02. Aug. 2002.
5. Mihajlo A. Jovanovic. Modeling Large-scale Peer-to-Peer Networks and a Case Study of Gnutella. Master Thesis. Apr. 2001.
6. R.P-Satorras and A. Vespignani. Epidemic Spreading in Scale-Free Networks. Phys. Rev. Letters. Vol. 86, No. 14. Apr. 2001.
7. R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. Nature no. 406, pp. 378–382, 2000.
8. Q. Lv, P. Cao, E. Cohen, K. Li and S. Shenker. Search and Replication in Unstructured Peer-to-Peer Networks. ICS'02. Jun.2002. Pages: 84-95.

9. L. Adamic, R.Lukose, A. Puniyani, and B. Huberman. Search in Power-Law Networks. Phys. Rev. E64(2001), 046135.
10. D. A.Menasce and L. Kanchanapalli. Probabilistic Scalable P2P Resource Location Services. ACM SIGMETRICS Performance Evaluation Review, Volume 30, Issue 2. Sep. 2002.
11. N. Sarshar, V.P. Roychowdury, and P. Oscar Boykin. Percolation-Based Search on Unstructured Peer-to-Peer Networks. Sunmission to IPTPS 2003.
12. Z. Jia, X. Tang and J. You. Search in Unstructured Peer-to-Peer Networks. WISE 2004. 2004.
13. C. Jin, Q. Chen, and S. Jamin. Inet: Internet Topology Generator. Technical Report CSE-TR443-00 , Department of EECS, University of Michigan,2000.
14. Pajek website: http://vlado.fmf.uni-lj.si/pub/networks/pajek/

A New Service Discovery Scheme Adapting to User Behavior for Ubiquitous Computing[1]

Yeo Bong Yoon and Hee Yong Youn

School of Information and Communications Engineering,
Sungkyunkwan University, 440-746, Suwon, Korea
{yuny,youn}@ece.skku.ac.kr

Abstract. Discovering primary services requested by users is a very difficult but crucial task in networked system, especially in ubiquitous computing system. The service discovery also has to be an accurate, fast, and stable operation. The typical lookup services such as SLP, Jini, and UPnP focus on either wired networks or particular networks. In this paper, thus, we propose a new service discovery algorithm using service agents adapted to user behavior. We compare the proposed algorithm with the DEAPspace algorithm using real experiment. It reveals that the proposed scheme offers services to the users seamlessly and accurately as much as DEAPspace, while the number of packets used is about half of it.

Keywords: Distributed service, service discovery, ubiquitous computing, user behavior, and wireless network.

1 Introduction

Various computing and networking paradigms have been evolving continuously in the past decades. Nowadays, the users no more need to adjust to the new computing systems because they were designed to learn the pattern of user behavior and adapt themselves to it. Ubiquitous computing system sets up quite efficient environment based on the habits of users. There also exist numerous kinds of small devices such as cellular phone, notebook, mobile devices, and PDA which are being deployed radically in the field nowadays. Advent of such mobile devices allows people to get any service in any place. However, having mobile devices is not enough to construct efficient ubiquitous computing environment. Here it is essential to provide the services not only efficiently but also promptly.

The ubiquitous computing system requires systematic registration and management of services for allowing seamless service in huge network environment. Furthermore, small devices have to rely on battery to operate in all kinds of mobile networks. Therefore, the key issue is how to offer stable and prompt services to the users using small battery power. For this, the ubiquitous computing system has to employ energy-

[1] This research was supported by the Ubiquitous Autonomic Computing and Network Project, 21st Century Frontier R&D Program in Korea and the Brain Korea 21 Project in 2004. Corresponding author: Hee Yong Youn.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 19 – 28, 2005.

efficient, seamless, and user-adaptable service discovery technology. A variety of services have been developed for users in mobile internet environment. Because of this, service discovery technology becomes very important in mobile network.

There exist several solutions for the issue mentioned above, which have been proposed and adopted in the real world. The Service Location Protocol (SLP) [1] is a standard protocol which offers scalable framework automatically discovering resources in the IP network. Lookup service was defined by Sun Microsystems to discover and manage services in Jini [2] middleware system based on CORBA [3]. Jini Lookup Service provides mainly naming and trading service in the network. The Service Discovery Protocol (SDP) [4] was addressed by Bluetooth Forum to compose a Bluetooth piconet and user services. UPnP [5] of Microsoft defines an architecture for pervasive peer-to-peer network connectivity in home network.

Service discoveries with SLP and Jini are very stable and powerful in terms of offering services to the users. However, they do not focus on the mobile network and their operations are restricted to low power mobile devices. In this paper, thus, we propose a new service discovery algorithm using service agents adapted to user behavior. In the proposed algorithm each node has a number of requested services. Network traffic is substantially reduced in the proposed scheme by keeping a list of services maintained according to the frequency requested. If a service infrequently used is requested, it is served by the on-demand model. The proposed algorithm keeps the latest services which are frequently requested, and sends infrequently requested service to reduce network traffic. We compare the proposed algorithm with the DEAPspace algorithm [6-8] using real experiment. It reveals that the proposed scheme offers services to the users seamlessly and accurately as much as DEAPspace, while the number of packets used is about half of it. Note that reducing packet transmission is a very important issue for small mobile devices of limited energy.

The rest of the paper is organized as follows. Section 2 briefly reviews the existing discovery services. Section 3 proposes a new service discovery algorithm adapted to user behavior. Section 4 presents the experiment results. Finally, we conclude the paper in Section 5.

2 Related Works

Numerous approaches have been proposed to discover various services in a network. Here we briefly review two systems among them, DEAPspace and Konark [9]. DEAPspace furnishes an algorithm that discovers services in a single hop adhoc network, while Konark is a middleware offering service discovery in multi-hop adhoc network. DEAPspace and Konark are the representative system in terms of performance and effectiveness for service discovery in single hop and multi-hop adhoc network, respectively, and thus they are mostly adopted as references.

2.1 DEAPspace

IBM has developed DEAPspace that solves the problem of service discovery in wireless single-hop adhoc networks. It restricts itself to a small network by assuming a single-hop adhoc network and broadcasting fit in a message. Its final proposal is for

fast concentration of available service information in the network. Maintenance of a centralized node storing all the service information in the network is difficult and complicated. DEAPspace thus selects a distributed approach with the push model in which servers send unsolicited service advertisements to the clients.

In the DEAPSpace, each node has a world view which stores a list of all services offered in the network. Furthermore, each node periodically broadcasts its own world view to other nodes using the adaptive backoff mechanism. If a node receives a service advertisement, it checks the service id and expiration time of its own world view, and decides whether to advertise a service message or update the cache data. Thereafter, the node increases the rate of broadcasting. If new service advertisements offered in the network have longer lifetime values than those in the internal cache of a node, it adds them into its own internal cache [10,11].

2.2 Kornark

Konark is a middleware designed specifically for discovery and delivery of services in multi-hop adhoc networks. It thus aims at wireless adhoc network that is usually larger than the network DEAPspace supports. The Konark architecture mimics a typical operating system and it is programming language independent. It provides a framework in which services are described and delivered using open standard XML technology over IP network connectivity.

Konark supports both the push and pull model. When a node receives a service message of the multicast address of the Konark, it multicasts different world view of relevant services and other services contained in the received message. When a client sends service requests, the servers replies unsolicited service advertisements. Each node adopts Konark SDP Manager which is responsible for discovery of the requested services, and registration and advertisement of its local services. To discover services in the network, clients use a discovery process known as active pull mechanism. Each node joins a locally-scoped multicast group.

The SDP Manager of each node maintains a cache, called a service registry. The service registry is a structure that enables devices to store their local services. It also allows them to maintain information on the services that they might have discovered or received via advertisements.

3 The Proposed Scheme

In general, if a server frequently advertises the service list, the clients can quickly discover the requested service. Here discovery service must not take too much traffic while offering services in a timely fashion. The two conflicting goals, less traffic and fast response, should be well balanced when the services are available to use. In addition to satisfying these goals, a service discovery which is also user-adaptable is proposed.

3.1 The Overview

For ubiquitous computing it will be efficient to use service discovery for both the server and client such as Jini lookup service. The lookup service is located in the

middle of the clients to manage the requested services. This kind of service discovery approach based on agent system can provide fast, efficient, and seamless services to the one needing them. However, the server and client system can be effective only in the networked environment. Mobile devices are not fixed at specific locations inside the mobile network. This is one of the reasons that a server system is not set in mobile network. Mobile devices need to frequently send a broadcast message to get service information with respect to service location, service time, and so on. They have their own repository to keep the latest service information.

Note that most people want to repeatedly use only some specific services from the network. They are just interested in the services which are related to their job or hobby. It is thus better to offer an efficient service mechanism only for a confined set of services instead of inefficient wide spectrum of diverse services. If a user is not interested in a specific service the user may demand it infrequently or never does that.

The basic idea of the proposed scheme is to substantially reduce the traffic due to infrequently requested services as identified above. Providing service list adapted to user behavior and preference will significantly reduce network traffic and offer fast services. The proposed scheme employs an approach in which each device broadcasts such adapted service list for the distributed system. Moreover, when a user wants to use services, the user sends a request message like an on-demand service model for reducing network traffic.

3.2 The Flow of Operation

Figure 1 shows how the proposed service discovery operation works, which consists of the following three steps for example.

Step 1. If Node-2 needs a service, it checks the service cache in its own repository.
Step 2. If the service is not in the service cache, it sends a request to other devices.
Step 3. If other devices have the service requested by Node-2, the device broadcasts the service list to other devices.

In the proposed scheme each node has a service cache (SC) storing the adaptable service list. Figure 2 shows the SC consisting of service elements (SE). Each SE has three fields; Service ID, Expiry, and Hit. Service ID is used to distinguish the service and Expiry indicates how long the service can be used by the users. Hit indicates how many times the user requested the service. When a user requests a service, its Hit value is increased by 3 points.

Each Hit value is decreased by 1 point if not requested in every 20 minutes. As a result, the Hit value of frequently requested service can maintain relatively higher Hit value than that of uninterested services decreasing continuously. All the SEs in the SC are sorted by the Hit value. We assume that ten most frequently requested service elements in the SC belong to a set of special interests, while others are plain service elements probably unrelated to the users. The ten SEs are called ‘hot elements’, and the number of hot elements can be varied for each specific implementation.

The hot elements are kept in the latest service list of the SC, and the message containing them is broadcast more frequently than other messages to reduce network traffic.

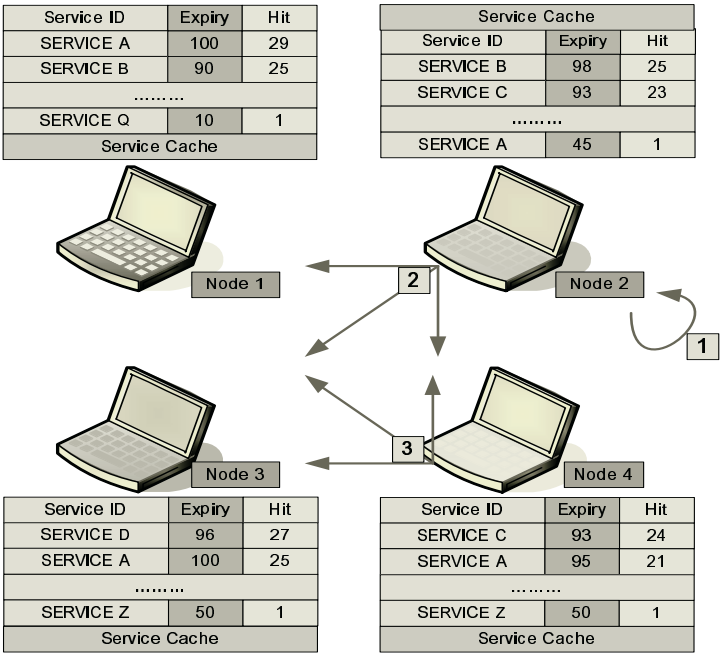


Fig. 1. The operation flow in the proposed scheme

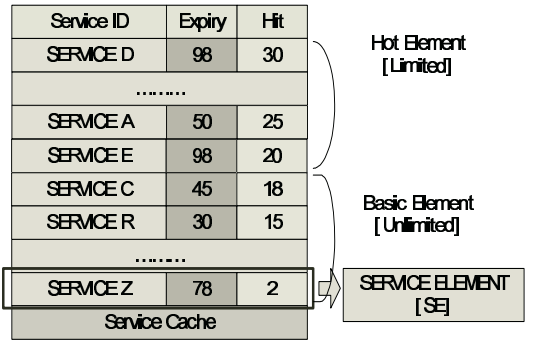


Fig. 2. The structure of service cache

3.3 Algorithm

The proposed algorithm consists of three main functions; Local Service Function, Receive Element Function, and Receive Service Function. Figure 3 shows a pseudo code of Local Service Function checking whether the requesting service is in the SC or not. Figure 4 shows how Receive Element Function works when a device receives a message requesting a specific service from it. Finally, Receive Service Function, defined in Figure 5, settles other's SC receiving a message from a remote node.

3.3.1 Local Service Function

Line 3: When a user requests a particular service, the hit count of the SE is increased.
Line 4: The SC is sorted by the hit counts of the SEs to maintain the latest data reflecting the usage.
Line 5: Local Service Function checks if the requested service exists in the SC. In Figures 3, MINE in line 3 is its own SC.
Line 7: Check_expiry function has two kinds of operations. If the requested service id belongs to hot elements, check_function checks the expiration time of all the hot elements. When the requested service element is just a basic element, check_function checks the expiration time of just the service element.
Line 9: If the hot element and requested service have a smaller value than minimum value of expiration time, local service function broadcasts a query message. One of the user's favorite services or just requested service will soon be shut down.
Line 14: If the own SC does not have the service element, Local Service Function also broadcasts the query message.

```
1   LOCAL SERVICE (SE d.ID)
2   {
3         Increase (MINE.d.HIT)
4         Sort (MINE)
5         IF (d.ID ∈ MINE.ID)
6         {
7           IF (Check_expiry())
8              {
9                  BROADCAST ELEMENT(d.ID)
10             }
11        }
12        ELSE
13        {
14          BROADCAST ELEMENT (d.ID)
15        }
16  }
```

Fig. 3. The pseudo-code of the Local Service Function

3.3.2 Receive Element Function

When a node receives a broadcast message, the Receive Element Function checks whether the SE is in its own SC or not. If it exists in its own SC and the expiration time is greater than the minimum value, the Receive Element Function broadcasts all the SEs in the SC. It is unnecessary for a node to request the SE even though the node does not have the SE. This is because the node will receive the SE from other nodes sooner or later through the Receive Service Function message. Note that the service lists are synchronized through the distributed service in the network.

```
1   RECEIVE ELEMENT (c.ID)
2   {
3         IF (c.ID ∈ MINE.ID)
4         {
5             IF(MINE.ID.EXPIRY > MIN)
6             {
7                 BROADCAST SERVICE (MINE)
8             }
9         }
10   }
```

Fig. 4. The pseudo-code of the Receive Element Function

3.3.3 Receive Service Function

Line 4-7: When a node receives a new SE from a remote node, the node inserts that in its own SC. The service is stored at bottom of the SC.

Line 10-11: If the SE is already in the SC, it checks the expiry time. If the received SE has the latest date, the Receive Service Function updates the expiry time of it.

```
1   RECEIVE SERVICE (REMOTE)
2   {
3         For each r □ REMOTE
4         IF (r □ MINE)
5         {
6           Insert(r, MINE)
7         }
8         ELSE
9         {
10          IF(r.expiry > MINE.expiry)
11          {  Update(r, MINE) }
12        }
13   }
```

Fig. 5. The pseudo-code of the Receive Service Function

4 Performance Evaluation

This section describes the experiment environment and the results of experiment.

4.1 Experiment Environment

In our experiment we use four 2.4GHz Pentium IV machines with 1GB of main memory each. We assumed the number of services in a network is 50, and each ser-

vice has a random expiration time from 1 to 100. Expiration time is decremented in every two minutes. Also, every service updates its own expiry time to a random number between 50 and 100 in every twenty minutes. Each node has 30 service elements at the starting point. When we set the value of time-out as 3'00'', the average time-out with the DEAPspace algorithm turned out to be 4'25''. Thus, we set a random time value between 2'00'' and 4'00'' as the time a user requests a service in the experiment with the proposed algorithm.

4.2 Experiment Result

Figure 6 shows the average expiry time value of the proposed scheme and DEAPspace. We can see that the difference between the two is very small, which means that the proposed scheme offers seamless and accurate service to the users as much as DEAPspace does. Notice that each peak point in the plots represents service update time in every twenty minutes as we preset.

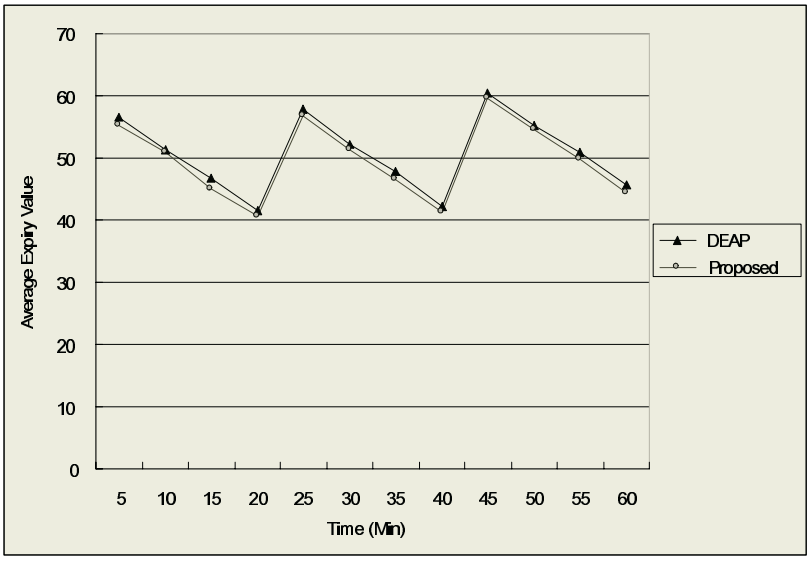


Fig. 6. Comparison of average expiration time.

Figure 7 shows the amount of packets transmitted in every ten minutes. Recall that when a user requests a service belonging to uninterested service elements (non-hot elements), the proposed scheme broadcasts a query message to get the information of others. Proposed-1 is the case that services in the SC are requested equally likely. Meanwhile, Proposed-2 is the case that user requests some service elements more frequently than the others. When this happens, the Local Service Function checks the expiration time of the hot elements. If one of the hot elements has a minimum expiration time, the node broadcasts a query message until it gets the latest data of the service at every request time. As a result, Proposed-2 uses more packets than

Proposed-1. Notice from Figure 6 that Proposed-1 and Proposed-2 use fewer packets than DEAPspace. Also, notice from the figure that the amount of packets of the proposed scheme and DEAPspace linearly increases as time passes. In other words, the average number of packets transmitted per minute is about 300 and 600 for the proposed scheme and DEAPspace, respectively. The proposed scheme uses about half of the packets used by DEAPspace, which is a quite substantial reduction.

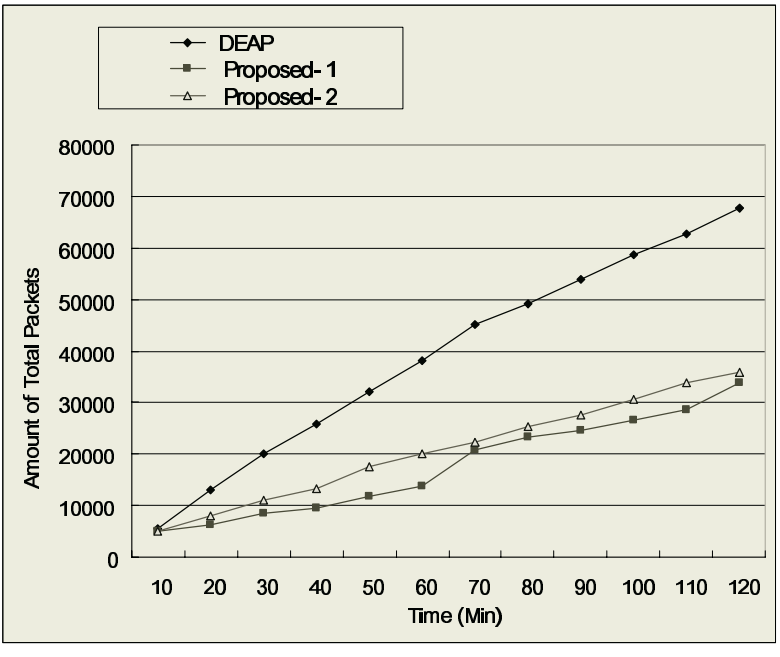


Fig. 7. Comparison of total number of packets transmitted

5 Conclusion

We have addressed a new computing paradigm in which the users demand individual service. The ubiquitous system requires registration and management of systematic control of services for each user in mobile network. For this reason, service discovery technology becomes very important in mobile network. We have also reviewed some present solutions such as Jini, SLP, SDP, and UPnP with more attention to DEAPspace and Konark. Most of the present service discoveries are focusing on networked system or particular network such as Bluetooth or Piconet.

In this paper we have proposed an energy-efficient service discovery algorithm using a list of services adapted to the frequency of requests. Comparing with DEAPspace, we have shown that the proposed scheme allows the same service quality using much less network traffic. Experiment with four machines displayed that the proposed scheme uses about half of the amount of packets used by DEAPspace. We will ex-

pand the experiment considering various factors such as the type of the services requested and duration of usage.

When a hot service element cannot be offered to a user, several packets need to be transmitted to get the service information. We will carry out research to get over the problem using effective service registration.

References

[1] SUN Microsystems: Jini Architecture Specification, Version 1.2. http://wwws.sun.com/software/jini/specs/jini1.2html/jini-title.html (2001)

[2] Guttman, E.: Service location protocol. Automatic discovery of IP network services. in Internet Computing, IEEE (1999) Volume 3, Issue 4, 71-80

[3] Object Management Group: The Common Object Request Broker Architecture. Core Specification, Version 3.0.3, Editorial changes formal/04-03-12 (2004)

[4] Bluetooth, Specification of the Bluetooth System, Specification Volume 1, 2001, Specification Volume 2, 2001. http://www.bluetooth.com

[5] UPnP FORUM, Universal Plug and Play Device Architecture Version 1.0.1, http://www.upnp.org, 2003

[6] M. Nidd : Service Discovery in DEAPspace, IEEE Personal Communications, August 2001.

[7] R. Hermann, D. Husemann, M. Moser, M. Nidd, C. Rohner, A. Schade : DEAPspace - Transient Ad-Hoc Networking of Pervasive Devices, Computer Networks, vol. 35, pp. 411-428, 2001.

[8] M. Nidd, "Timeliness of Service Discovery in DEAPspace," 29th Int'l. Conf. Parallel-Processing," 29th Int'l. Conf. Parallel Proc. Workshop. Pervasive Comp., Aug. 2000, pp. 73–80.

[9] S. Helal, N. Desai, V. Verma, C. Lee : Konark - A Service Discovery and Delivery Protocol for Ad-hoc Networks, Proceedings of the Third IEEE Conference on Wireless Communication Networks (WCNC), New Orleans, March 2003.

[10] Honghui Luo, Michel Barbeau : Performance Evaluation of Service Discovery Strategies in Ad Hoc Networks, 2nd Annual Conference on Communication Networks and Services Research (CNSR 2004) Fredericton, N.B., Canada, May 19-21, 2004

[11] M. Barbeau, E. Kranakis, Modeling and Performance Analysis of Service Discovery Strategies in Ad Hoc Networks, Proceedings of International Conference on Wireless Networks (ICWN), Las Vegas, Nevada, 2003.

The Design and Prototype of RUDA, a Distributed Grid Accounting System

M.L. Chen, A. Geist, D.E. Bernholdt, K. Chanchio, and D.L. Million

Oak Ridge National Laboratory, Oak Ridge, TN, USA
{chenml, gst, bernholdtde, chanchiok, milliondl}@ornl.gov

Abstract. The Grid environment contains a large and growing number of widely distributed sites with heterogeneous resources. It is a great challenge to dynamically manage and account for usage data of Grid resources, such as computational, network, and storage resources. A distributed Resource Usage Data management and Accounting system (RUDA) is designed to perform accounting in the Grid environment. RUDA utilizes fully decentralized design to enhance scalability and supports heterogeneous resources with no significant impact on local systems. It can easily be integrated into Grid infrastructures and maintains the integrity of the Grid security features.

1 Introduction

Modern science and technology are increasingly collaborative and often demand huge computing and storage resources, which individual research organizations may not possess. A distributed infrastructure - Grids - was created in early 2000. Grid technology, such as the protocols and services developed by Globus [1], enables flexible, controlled resource sharing on a large scale. Driven by the demand and attracted by the promising future of Grids, Grid applications multiply swiftly and in turn drive the rapid development of Grid technology. While many Grid services are maturing, Grid accounting still remains a research issue. The Grid environment, which contains a large and growing number of widely distributed sites with heterogeneous resources, poses great challenges to Grid accounting. Rapid evolution of Grid software infrastructures and increasing security concerns add additional complications. Therefore, a Grid accounting system needs to be scalable, flexible, and secure [2]. Though many methods and tools have been successfully used in individual sites for resource usage management and accounting [3-4], they are local, centralized systems, of which the scalability is limited by the load capability and data size of the centralized server and database. These accounting methods and tools do not satisfy the requirements of Grid accounting.

The Science Grid (SG) project is a collaboration involving researchers at Lawrence Berkeley (LBNL/NERSC), Pacific Northwest (PNNL), Argonne (ANL), and Oak Ridge (ORNL) National Laboratories, sponsored by the U.S. Department of Energy (DOE), to explore the Grid environment for use with the large-scale computer and storage systems available at DOE sites. The ability to properly account for resource

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 29 – 38, 2005.

usage across such a Grid is crucial to its acceptance and use. To the best of our knowledge at the time of this project started, there was no Grid accounting system that could meet the challenges of the Grid environment.

We have designed and prototyped a Grid Resource Usage Data management and Accounting system (RUDA). RUDA is designed in a fully distributed manner for scalability. It employs customizable interfaces to communicate with local systems to support widely diversified resources without significant impact on them. RUDA leverages security features embedded in Globus for secure wide-area communication, and accommodates current economic models of the Grid with respect to valuation of resources. RUDA is an essentially self-contained system that can be easily integrated into the Grid environment.

The following three sections present the system architecture, the accounting process, the feature design targets and approaches. In later sections, the implementation and experiment of the prototype are discussed.

2 System Architecture

In the Grid environment for which RUDA is designed, resources are provided by computer centers or divisions of individual sites. The users are organized in research projects and each project has charge-account(s) used by the providers to credit/charge the allocations of the project. The allocation in a charge-account may be distributed to the individual users who share this account.

The basic building-block of RUDA is a client/server software package. The multithreaded server consists of core-software and interfaces. The latter collects resource usage data from local accounting systems and the former, isolated from the local system, provides data management, accounting, and web interface services. A RUDA server can be configured in two running modes, called *basic-server* and *head-server* respectively. The client-process provides the services for users/administrators to access/control the server and for other servers in RUDA to communicate with this server. The client-process resides on the same machine with its server. The Globus Toolkit's GRAM [5] commands are employed to remotely run the client-process to communicate with the server from any computer on the Grid.

In RUDA, a *basic-server* daemon runs on each participating resource and periodically pulls resource usage information of Grid users from the local accounting system. It manages and accounts for the resource usage data, and stores the records in a local MySQL database. *Head-servers* are used to aggregate accounting and usage information for organizations, projects, and other interested entities; they can be deployed on any computer on the Grid. The *head-server* is configured to select the "member" *head-* or *basic-servers*, which contain the resource usage data of interest, and the *head-server* queries desired data segments from each member server by running the client-process of the member remotely with criteria for the specified projects, charge-accounts, and users. The member server writes the requested data set into a RUDA standard data file. Transferring the data back by GridFTP [5], the *head-*

server manages and accounts for the data collectively. Since a *head-server* can collect data from both *basic-servers* and other *head-server*s, a RUDA system with flexible hierarchical structures can be built for large organizations to perform accounting.

RUDA also provides allocation services. In the DOE community, allocation on major resources requires pre-approval of authorized administrators. RUDA server provides web interfaces for project Principle Investigators (PIs) to check available resources and to apply for allocation.

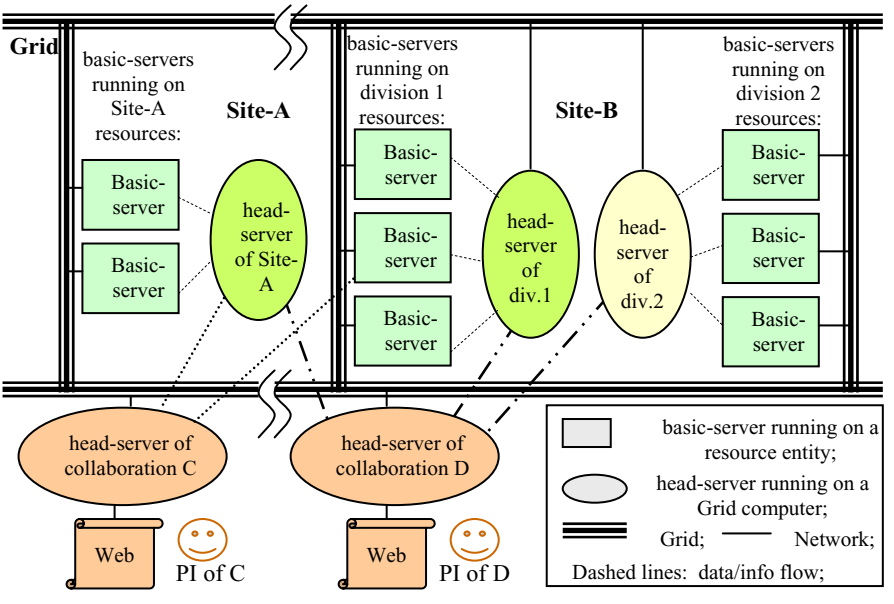


Fig. 1. A simple example of RUDA system

Figure 1 shows a much simplified RUDA system example. Independent and geographically separated sites A and B provide computer resources to the Grid. A *basic-server* on each individual computer collects data and accounts for the Grid resource usage at the levels of job, user, and charge-account. The resource providers, Site-A and two divisions at Site-B, employ *head-server*s to manage and account for their resource usage. Each *head-server* collects data from its member servers and performs accounting collectively at the levels of user and charge-account. These *head-server*s need not to collect detailed data at job level unless desired. Projects C and D use computer resources at both Site-A and Site-B. Each PI runs a *head-server* and configures relevant servers as its members. This also provides a simple example of the hierarchical structure of the system.

3 Accounting Process

3.1 Standard Resource Usage Records and Global User Identifications

Grid accounting systems must manage and exchange resource usage data globally, however there are currently no standards for this type of data. For example, a survey [6] shows that more than 15 technical terms are used for CPU usage records by various local accounting systems in 5 national laboratories, and each of them has its own definition and data type. To perform Grid accounting, a standard set of usage record fields has to be defined and a mechanism to convert the local fields into the standard fields needs to be developed. A standard set of resource usage fields has been defined for RUDA based on the Global Grid Forum (GGF) suggestions, a survey of DOE laboratories [6,8], and a more detailed examination of local accounting systems used at ORNL, PNNL, and LBNL/NERSC. A data structure called *data-cargo* accommodates the standard fields within RUDA servers and a corresponding file format is used when exchanging records within a distributed RUDA deployment. For each new type of local accounting system, a customized reference table needs to be setup to allow the standardization routine of RUDA server interface to convert the local system's input records into a standard data set and store them in a *data-cargo* structure.

Grid-wide user identification is also critical for Gird accounting. One user may use different user identifications (IDs) on different computers. To uniquely identify users Grid-wide, rules to create global user IDs have to be defined.

RUDA utilizes the Grid user distinguished name (GUdn) as its global user ID. GUdn is defined by Grid Security Infrastructure (GSI) [5] of Globus. For authentication purpose, GSI grants each user a user-certificate, which contains a globally unique GUdn. For access control purpose, each resource has a grid-map file to map each GUdn with local user ID (LUid). Using information extracted from the grid-map file, a RUDA *basic-server* can pair up LUids with GUdns. The standard data record contains both GUdn and LUid for user identification.

When the customized interface of *basic-server* reads data in, it calls the standardization routine to covert the local data into a standard data set, pairs up each GUdn with LUid, and fills the standard data set into a *data-cargo* structure which is ready to be transferred to RUDA server for usage data normalization and accounting.

3.2 Usage Data Normalization and Accounting

Various Grid resources carry different qualities of service and different capabilities. The ability to normalize accounting data across diverse Grid resources is important to the Grid "economy" [2, 9], especially if user resource allocations are fungible across multiple resources. RUDA provides the flexibility to support economy-based pricing of Grid resources. RUDA defines a "RUDA-allocation-unit (RAU$)" as the standard charge unit and each server performs accounting for each job according to a customized formula. The formula currently used in the prototype is

$$total\_charge = \left[\sum_{i=1}^{n} (Usage\_amount(i)) \times (weight(i)) \right] \times \prod_{j=1}^{m} (global\_weight(j))$$

Here, n is the total number of resource categories of the standard resource usage data set and i spanning from 1 to n represents category index. *Weight(i)* presents the weight factor of resource category i. m is the total number of the global weight factors. The *global_weight(j)* is j'th global weight factor, such as the priority or quality of service, usage timing (peak, off-peak), or any factors which effect the overall cost of a job.

Referring the "Demand and supply model" suggested by [2], the weight factors of each resource can be configured by its provider independently according to their site policy and user demand, though the Grid administrator may publish a list of weight factors for a number of commonly used resources as reference. Grid economy research shows that this model would regulate the resource demand and supply automatically, give each site maximum control over the price, and eliminate the necessity of centralized resource evaluation. The formula and the weight factor information are stored in the server's database and available for users upon request.

The *basic-server* on each resource records the usage data and charge of each running job in a dynamic data structure. Upon the completion of a job, its data are moved into a local database. Based on the data of both current and completed jobs, the server calculates the individual category resource usage and total charges accumulated since the beginning of current accounting period for each user, charge-account, and project respectively. By means of *head-server*(s), PIs or users collect the relevant data from the resources they use and account aggregately for their dynamical Grid resource usage and total charges.

4 Feature Design Targets and Approaches

4.1 Scalability

The large and growing scale of Grid environments poses great challenges to the accounting system design. Instead of technically enhancing the capability of the centralized server and database, RUDA takes a distributed approach that focuses on maximally decoupling the loads of the server and database from the scale of Grid. The data collection and storage are performed locally by a *basic-server* on each resource entity. Therefore the *basic-server* load and its database size are independent of the Grid scale. *Head-server*s perform usage data management and accounting for projects or providers. A *head-server* for a provider manages the data on the resources they provided. A *head-server* for a PI only sees the data relevant to his/her project or group. The data collected/managed by a *head-server* and stored in its database depends on the size of the project/group or the resource provider's environment, not on the scale of the Grid. This decoupling strategy eases the scalability limitation caused by database size and server load without requiring breakthrough technologies. The flexible *head-server* architecture also enhances scalability, since servers can be deployed as needed on a per-project, per-provider, or other basis. It is only necessary

for a *head-server* to know which other servers to poll to obtain the desired resource usage information. In the DOE laboratory environment, this is fairly straightforward, since accounts and allocations are typically tracked and already known to the providers/PIs.

4.2 Security

Significant efforts in Grid software development have gone into secure network communications and remote interactions. The Globus Grid Security Infrastructure (GSI) is based on existing standard protocols and APIs [5] and extends these standards to single sign-on and delegation, forming a comprehensive security system [10]. RUDA employs the Globus toolkit, and GSI in particular, to enable secure authentication and communication over the open network. For instance, to collect data from a remote *basic-server*, a *head-server* issues a GRAM command to call the *basic-server*'s client-process remotely for data query. The Globus servers on both ends then handle mutual authentication and secure remote communication for the RUDA servers. During the data transfer procedure, GridFTP manages the mutual authentication and insures the communication integrity. In this way, RUDA is based upon standard Globus security features.

The authorization for data access is performed by the RUDA server in two layers. One layer is applied to authorize a *basic-server* collecting data from the local system. Each *basic-server* owns a user/project map provided by the local administrator through a configuration file. The map lists the users and projects of which the data is allowed for RUDA to access, and the server interface limits queries to the authorized data only. The other layer is control of the access to the RUDA server's database. The database contains the user attributes originally obtained from the local system. When a remote user queries for data, the server performs the data access control according to the user status. For example, an ordinary user can only access his/her own data, while a PI can access the data of their entire project/group.

4.3 Fault Tolerance

We have chosen a fail-over mechanism as a short-term solution to fault tolerance. We run backup daemons on backup machines, which periodically probe the existence of running daemons. A backup daemon of *basic-server* also keeps a copy of the current resource usage data of the running daemon. Once a failure is detected, the backup daemon can conclude that the original daemon or machine has died and inform the other daemons that are interacting with the original daemon that it has taken over. Simultaneously, the backup daemon informs the system administrators that the fail-over has occurred. The backup daemon of a *basic-server* only provides the latest accounting data upon request. The administrator is expected to troubleshoot and recover the original server or machine as soon as receiving the failure information.

Although the above mechanism is easy to implement and sufficient for a small group of RUDA daemons, it can not distinguish network partition from machine failure, handle simultaneous failures of the original and backup daemons, or deal with

the complications in a large distributed environment. We have chosen the group membership management mechanisms as the long-term approach to fault tolerance [11, 12]. In such a system, the RUDA daemons will monitor one another and operate self-healing mechanisms once they agree that a group member has failed.

4.4 Flexibility and Manageability

RUDA's design includes interfaces that can be customized to communicate with local accounting systems, allowing RUDA to support heterogeneous resources with various local accounting systems with a minimum local impact. To avoid the complications caused by the modification of Grid software infrastructure, RUDA software is built to be essentially self-contained. The utilization of Grid infrastructure is limited to its high level APIs and the modifications behind the APIs have no effect on RUDA. Furthermore, the APIs are called in a couple of customization routines. By modifying the customization routines, RUDA can utilize various versions of Globus or other Grid infrastructures.

The RUDA server is fully configurable, such as its user/project map, data polling period, data backup method, and so on, and supports runtime reconfiguration. By means of GRAM and RUDA command line interfaces, the administrator can configure and control the server remotely from any computer on the Grid.

5 Prototype Implementation

A prototype of RUDA has been developed on an SG testbed. It contains most of the major components of RUDA to test the feasibility of RUDA design, but the fail-over mechanism is not implemented due to lack of backup machines. The functions of prototype's client/server package are briefly described in the following.

The server chooses one of the three data collection interfaces (Fig. 2) according to its configuration. The interface shown on the left side of Fig. 2 accepts RUDA standard data files transferred between RUDA servers. The one in the middle is the major data input port of *basic-server*. By means of the customized routine, the interface inputs data from a local accounting system through CLIs/APIs provided by the system, converts them into a standard data set, and loads the server's *data-cargo*. The interface on the right side is designed for situations where the additional security requirements are imposed. To secure sensitive information, some sites do not allow foreign access to the local systems and sensitive information must be filtered out before resource usage data reach the Grid. This interface provides an independent daemon process run by authorized site administrators. To periodically pull data from the local system, the daemon calls a customized routine, which filters out the sensitive information and coverts the local data into the standard data set. The daemon then loads the data into its MySQL database, which RUDA server is authorized to access. On receiving a *new-data-ready* signal from the daemon, the server reads in the data from the daemon database.

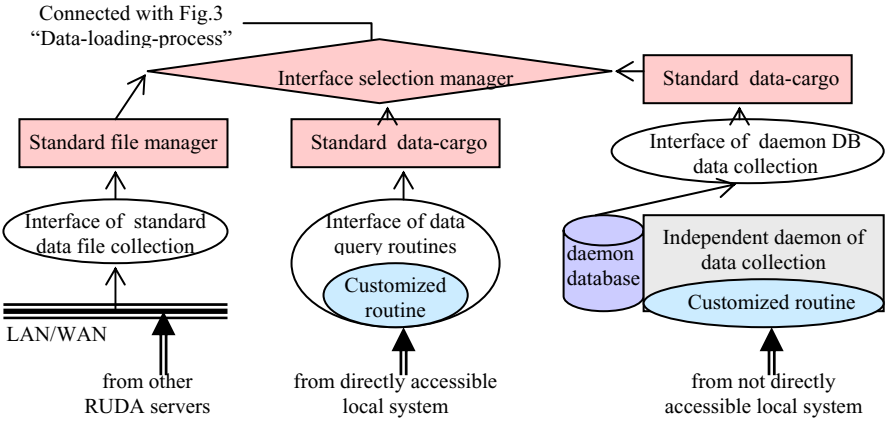


Fig. 2. RUDA server block diagram part 1: Server interfaces

Upon receiving new data, the server core-software (Fig. 3) utilizes the customized formula shown in Section 3.2 to calculate the charge of individual jobs, and summarizes the usage data and charges from both current jobs and the jobs completed within the accounting period for users, charge-accounts, and projects. The updated current data are stored in server's data structure and also copied into a local database called mirror-site, which can be used to recover the current data in case the need arises.

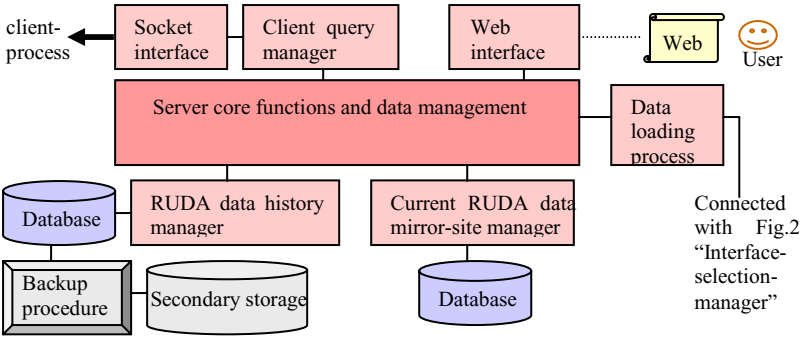


Fig. 3. RUDA server block diagram part 2: Server core-software

The client-process only communicates with the server through local socket connections. The Globus toolkit GRAM is employed to run the client-process from a computer on the Grid to perform remote communications with the server.

The prototype provides web interfaces for users to conveniently access their resource usage data and for PIs to check resource availability and apply allocations.

The command line interfaces are also provided for users query resource usage data, and for administrators to configure and manage the server locally or remotely.

6 Experimental Results

The prototype system has been experimented on the Earth System Grid (ESG). The ESG project uses Grid technology to support the climate research community to discover, access, and analyze large-scale global climate model simulation results in a distributed and heterogeneous computational environment. The RUDA prototype has been deployed on ESG computer resources at ORNL in Tennessee and National Center for Atmospheric Research in Colorado. The computer platforms include IBM AIX, Sun Solaris, and Linux redhat, on which the RUDA software is portable. Globus is the software infrastructure of ESG. GRAM and GridFtp of Globus Toolkit (version 2.2.4) with embedded GSI are enabled on these machines, and MySQL and Apache web server are available.

A total of five *basic-server* daemons ran in this experiment on a mixture of IBM AIX, Sun Solaris, and Redhat Linux systems, and a *head-server* daemon running on an IBM AIX machine configured them as its members. The average number of accounted jobs running simultaneously on each machine was at a level of a few tens to a hundred at any moment, with a duration varying from a few minutes to several days. Though a full set of standard resource usage fields had been defined in the prototype, only CPU time, wall time, and the memory usage (requested or high-water-mark) were captured in this experiment. Artificial weight factors (see section 3.2) were assigned to simplify the process of checking accounting results.

The experiment first ran for four weeks and concentrated on checking server functions. During this period, an error was reported by the *head-server* on data transfer from one of its member servers and it was caused by the downtime of that remote member computer. As mentioned, the fail-over mechanism has not been implemented. When a member server is down, the *head-server* uses the latest data set collected from that server as current data (with original data collection timestamp) until the member server is up and running again. The memory and CPU usage of the servers were also measured. According to the SG project's survey of resource providers and users, updating resource usage data and accounting records every hour would fully satisfy the requirement of dynamic accounting. With this configuration, the CPU time was less than 190 seconds per day for the *head-server*, and 10 seconds for each *basic-server*, respectively. The memory usage (high-water-mark) of all individual servers is less than 3 MB.

The experiment then continuously ran for 12 more weeks. All servers were configured to collect data every 4 minutes for experimental purposes. The snapshots of resource usage data and accounting results were taken and checked on a daily basis. All servers ran smoothly through the whole period and no accounting error was found. The RUDA web interfaces were used daily to monitor the system. The allocation application functions were also tested through the web interface, though the responses of administrators were performed automatically by a piece of simulation software.

7 Conclusion and Acknowledgement

A Grid Resource Usage Data management and Accounting system, RUDA, has been designed for the DOE Science Grid project. A prototype RUDA system has been implemented and tested to demonstrate the feasibility of the design. With more customized data collection routines, RUDA can be deployed onto computers with various local accounting systems and onto other types of resources such as mass storage systems. We are planning to perform further experiments on a larger RUDA system, with fault tolerance features implemented, in the near future.

We thank Scott Jackson and David Cowley at PNNL, the ORNL local accounting system development group at ORNL and East Tennessee State University, and Francesca Verdier at LBNL/NERSC for their support of our local accounting system investigation. The support from ESG project on the RUDA experimental deployment is also very much appreciated.

Science Grid project is sponsored by the U.S. DOE Office of Science under the auspices of the Scientific Discovery through Advanced Computing program (SciDAC). ORNL is managed by UT-Battelle, LLC for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

1. Foster, I. & Kesselman, C. Globus, “A Toolkit-Based Grid Architecture”, ibid, p. 278, Morgan-Kaufmann (1999).
2. Bill Thigpen & Tom Hacker, “Distributed Accounting on the GRID”, Global Grid Forum (GGF), 4-7 March 2001.
3. http://www.emsl.pnl.gov/docs/mscf/qbank
4. http://www.nersc.gov/nusers/accounts/nim
5. http://www-unix.globus.org/toolkit
6. Mi young Koo, “Usage Record Fields – Survey Results and Proposed Minimum Set”, GGF Oct., 2002
7. http://www.gridforum.org
8. Laura F. McGinnis, “Resource Accounting – Current Practices”, GGF Feb., 2001.
9. Rajkumar Buyya, David Abramson, and Jonathan Giddy, “A Case for Economy Grid Architecture for Service Oriented Grid Computing”, 10th Heterogeneous Computing Workshop, Apr. 23-27, 2001
10. http://www.globus.org/security
11. M. Franceschiti and J. Bruck, “A Group Membership Algorithm with a Practical Specification,” IEEE Transactions on Parallel and Distributed Systems, vol 12, no 11, 2001
12. K. Chanchio, A. Geist, M.L. Chen, “A Leader-based Group Membership Protocol for Fault-tolerant Distributed Computing”, submitted to the 14th International Heterogeneous Computing Workshop, 2005

An Adaptive Routing Mechanism for Efficient Resource Discovery in Unstructured P2P Networks

Luca Gatani[1], Giuseppe Lo Re[1,2], and Salvatore Gaglio[1,2]

[1] Dip. di Ingegneria Informatica, Università di Palermo,
[2] Istituto di Calcolo e Reti ad Alte Prestazioni, C.N.R.
Viale delle Scienze, 90128 Palermo, Italy
gatani@csai.unipa.it
{lore, gaglio}@unipa.it

Abstract. The widespread adoption of large-scale decentralized peer-to-peer (P2P) systems imposes huge challenges on distributed search and routing. Decentralized and unstructured P2P networks are very attractive because they require neither centralized directories, nor precise control over network topology or data placement. However their search mechanisms are extremely unscalable, generating large loads on the network participants. In this paper, to address this major limitation, we propose and evaluate the adoption of an innovative algorithm for routing user queries. The proposed approach aims at dynamically adapting the network topology to peer interests, on the basis of query interactions among users. Preliminaries evaluations show that the approach is able to dynamically group peer nodes in clusters containing peers with shared interests and organized into a *small world* topology.

1 Introduction

In the past few years, the peer-to-peer (P2P) paradigm has emerged, mainly by file sharing systems such as Napster and Gnutella. In the research community there has been an intense interest in designing and studying such systems. Due to the decentralization, these systems promise improved robustness and scalability, and therefore they open a new view on data integration solutions. However, several design and technical challenges arise in building scalable systems. While active, each peer maintains neighbor relationships with a small set of peers and it participates in the application protocol on the P2P network. These neighbor relationships, logical links between peers, define the topology of the P2P network. Therefore, the P2P topology forms an overlay on the IP-level connectivity of the Internet. One of the most challenging problems related to data-sharing P2P systems, is the content location. Content location determines whether the system resources can be efficiently used or not. Moreover, it greatly affects the scalability of P2P systems and their other potential advantages. Currently, there are two kinds of searching schemes for decentralized P2P systems [1]: structured searching scheme and unstructured searching scheme. Although structured systems such as Chord [2], Pastry [3], and CAN [4] scale well and perform efficiently,

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 39–48, 2005.

they have many limitations. First, they have high requirements on data placement and network topology, which are not applicable to the typical Internet environment, where users are widely distributed, extremely transient, and come from non-cooperating organizations. Second, they efficiently support search by identifiers, but not general search facilities. Unstructured P2P systems like Gnutella [5] do not have the problems mentioned above and they are the most suitable in the current Internet. However, they use an overlay network in that the topology and file placement are largely unconstrained, so no clue emerges as to where content is located, and queries have to be flooded through the whole network to get results. Flooding is robust and reliable but highly redundant [6], producing many duplicated messages and much network traffic. Several researches are currently carried out on improvements in flooding-based message routing scheme: most notable are random walk, multiple parallel random walks, iterative deepening, and local indexes. An interesting paper by Adamic *et al.* [7] studies random walk search strategies in power-law topology, but although otherwise effective, this strategy places most of the burden on the high degree nodes and thus potentially creates additional bottlenecks and points of failure, reducing the scalability of the search algorithm. Lv *et al.* [1] propose a k-walker random walk algorithm that reduces the load generated by each query. They also propose the adoption of uniform random graph, and study query efficiency under different query and replication models. Liu *et al.* [8] introduce a location-aware topology matching technique, building an efficient overlay by disconnecting low productive connections and choosing physically closer nodes as logical neighbors. Another interesting query routing technique is the iterative deepening, suggested by Yang *et al.* [9], where unsatisfied queries are repeated several times, each time increasing their search depth. In [10], the authors propose that nodes maintain metadata that can provide "hints" such as which nodes contain data that can answer the current query. Query messages are routed by nodes making local decisions based on these hints. In this paper, we focus on the unstructured architectural model and, to address searching inefficiencies and scalability limitations, we propose an adaptive routing algorithm whose aim is to suppress flooding. The routing algorithm adopts a simple Reinforcement Learning scheme (driven by query interactions among neighbors), in order to dynamically change the topology of the peer network based on commonality of interests among users.

The remainder of the paper is structured as follows. Section 2 provides background on the *small world* network paradigm. Section 3 illustrates the adaptive routing protocol proposed. The simulation setup and performance measurements are presented in Section 4. Finally, Section 5 concludes the paper.

2 Small World Networks

The *small world* phenomenon was first observed by Milgram [11], who discovered the interesting "six degrees of separation" in a social network. Although the notion of *small world* phenomenon originates from social science research, it has been observed that the *small world* phenomenon is pervasive in a wide range of

settings such as social communities, biological environments, and data communication networks. For example, recent studies (e.g., [12]) have shown that P2P networks such as Freenet may exhibit *small world* properties. Roughly speaking, a *small world* network can be viewed as a connected graph characterized by low *characteristic path length*(i.e., similar to the average path length in random networks) and high *clustering coefficient* (i.e., much greater than that of random networks). To mathematically define the two properties, let $G = (V, E)$ denote a connected graph modeling a *small world* network, $N = |V|$ the cardinality of the set of vertices, and $D(i, j)$ the length (in hops) of the shortest path between two vertices i and j in V. The characteristic path length $L(G)$ is defined as the number of edges in the shortest path between two vertices, averaged over all pairs of vertices. To define the clustering coefficient $C(G)$, suppose that a vertex $v \in V$ has k_v neighbors; then at most $k_v(k_v - 1)/2$ edges can exist between them (this occurs when every neighbor of v is connected at every other neighbor of v). Let C_v, the local clustering coefficient of v, denote the fraction of these allowable edges that actually exist. Define the clustering coefficient as the average of C_v over all v. While L measures the typical separation between two vertices in the graph (a global property), C measures the cliquishness (degree of compactness) of a typical neighborhood (a local property). A low average hop distance implies that one can locate information stored at any random node by only a small number of link traversals (low latency object lookup), while a high clustering coefficient implies the network can effectively provide contents even under heavy demands. Other works, in addition to our proposal of constructing a *small world* network, discuss the adoption of a *small world* topological structure in order to efficiently perform searching in P2P networks. Iamnitchi *et al.* [13] propose a solution for locating data in decentralized, scientific, data-sharing environments that exploits the small-worlds topology. Manku *et al.* [14] propose to build a one-dimensional *small world* network, by the adoption of a simple protocol for managing a distributed hash table in a dynamic peer network. The broadcast problem for communication in a *small world* network is considered in [15]. In [12], the authors propose a scheme for storing data in an unstructured P2P network such as Freenet, such that the P2P network may exhibit some of the *small world* properties.

3 Adaptive Routing Protocol

The key problem addressed in our work is the efficient and scalable localization of shared resources. The underlying idea is that an intelligent collaboration between the peers can lead to an emergent clustered topology, in which peers with shared interests and domains, tend to form strongly connected communities of peers. To this aim, we adopt an approach that dynamically selects the neighbors to which a query has to be sent or forwarded. The selection process is driven by an adaptive learning algorithm by which each peer exploits the results of previous interactions with its neighbors to build and refine a model (*profile*) of the other peers, describing their interests and contents. When an agent must

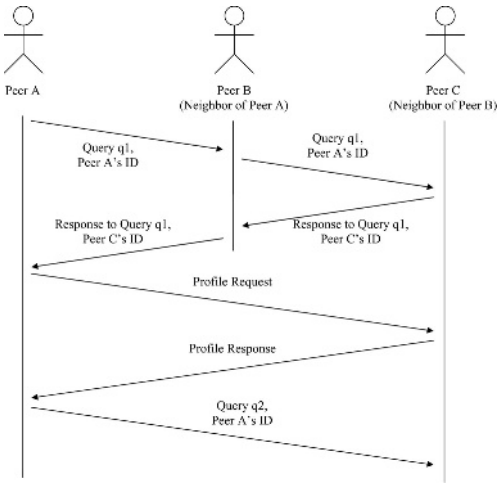


Fig. 1. The process of neighbor discovery

forward a query, it compares the query with its known profiles, in order to rank all known peers and to select the best suited to return good response. The network topology (i.e., the actual set of peers that are neighbors in the overlay) is then dynamically modified on the basis of the learned contexts and the current information needs, and the query is consequently routed according to the predicted match with other peers' resources. Since our goal is to allow peers to form communities in a fully distributed way, they need to find new peers and to evaluate their quality in relation to their own interests. Initially the network has a random, unstructured topology (each peer is assigned N_s neighbors randomly chosen), and queries are forwarded as in the scoped flood model. The peers can then discover other peers through known peers, during the normal handling of queries and responses. Each peer has a fixed number, N_m, of slots for known peers. This number can vary among peers depending on their available memory. Here, we assume that the N_m value is the same for each agent. For each known peer, a profile concisely describing the shared resources is stored. The actual set of N_a neighbors, i.e. those to whom queries are sent, is selected dynamically for each query at time step t among all the $N_k(t)$ known peers. In particular, when a peer receives a query locally generated, it compares the query with its stored profiles, applying a simple ranking algorithm for dynamically selecting peers to which it sends the query. The maximum number of selected peers, N_a, depends on peer bandwidth and computational power. In our test networks, we assume that N_a is fixed and equal for each agent. The system currently adopts Bloom filters [16] to build peer profiles and supports a basic query language where a query string is interpreted as a conjunction of keys. When presented with a query, the system searches the information in its profile database in order to obtain a list of candidate peers that might have data matching the query. When a peer receives a query by another peer, if it shares resources that match the request,

1. `Profile acquisition.` When a peer is first discovered, a profile is requested. After the peer's profile is acquired, a local peer description is initialized with the information stored in the Bloom filter.
2. `Profile updating.` When a response from a neighbor (or from a neighbors' neighbor) arrives, it is evaluated and used to update the description of known peers, adding the query keywords to the peer profile. Moreover, new peers that respond to issued queries are added to the list of known peers.
3. `Peer ranking.` When a new query has to be sent, all N_k known peers are ranked by similarity between the query and the peer descriptions, exploiting the membership information provided by Bloom filters.
4. `Query sending.` The new query is sent to the top N_a ranked peers. Then step 1 is newly considered.

Fig. 2. Basic steps of the algorithm proposed

it can directly respond. Moreover, it can forward the query to that neighbors, whose profiles match the query. To this aim, the peer uses the same selection algorithm applied to locally generated queries In order to prevent potential DoS attacks which exploit the response system, we impose that a peer replies to a forwarded query sending the response to the neighbor that has forwarded the query, and not directly to the originating peer. To limit congestion and loops in the network, queries contain a Time-To-Live (TTL), which is decreased at each forward, and queries will not be forwarded when TTL reaches 0. When a peer receives the responses for a locally generated query, it can start the actual resource downloading. Moreover, if a peer that has sent a response is not yet included in the list of known peers, a profile request is generated. The two peers contact each other directly (see Fig. 1). When the message containing the profile will arrive, the new peer will be inserted among the N_k known peers and its characteristics will be evaluated in order to select actual neighbors for new query. The stored profiles are continually updated according to the peer interactions during the normal system functioning (i.e., matches between queries and responses). Moreover, a peer can directly request a newer profile when necessary. In Fig. 2, we summarize the main steps of the proposed adaptive algorithm.

4 Experimental Evaluation

4.1 The Simulator

Since in the studies on deployed P2P networks [6, 17, 18], the dynamics in peer lifetimes and the complexity of these networks make it difficult to obtain a precise comprehensive snapshot, we use simulation to perform a preliminary evaluation of the proposed approach. Simulation of P2P networks can provide a thorough evaluation and analysis of their performance. To study the behavior of peer interactions in our system, we have implemented a simple simulator that allows

Table 1. Parameters used in Scenario 1

Parameter	Value
Number of peers N	100
Number of actual neighbors N_a	5
Number of initial neighbors N_s	5
Maximum number of known peers N_m	99
Number of groups N_g	10
Time To Live of queries TTL	3
Number of time steps N_ts	1000

us to model synthetic peer networks and run queries according to the routing protocol adopted. The goal of the simulator in the preliminary experiments reported below is to analyze the topology properties of emergent peer networks. Our simulator takes a snapshot of the network for every time step. In a time step of the simulator, all of the peers process all of their buffered incoming messages and send all of their buffered outgoing messages. This may include the generation of a local query as well as forwarding and responding to the queries received by other peers.

4.2 Simulation Scenarios

In the current evaluation, we perform four different kinds of experiments, considering four scenarios, each of them different from the others for a single simulation parameter. For each scenario, the aim is to study how the network statistics change when the parameter value changes. Since the initial random topology can affect the final results, for each scenario, we perform several independent simulations and we average the results. In order to study whether the proposed algorithm can generate network topologies that capture user interests, thus reducing query flooding problems, we model synthetic peers belonging to different groups of interest (let N_g denote the number of groups in the network). Each group is associated with a general topic. Within each topic, the resources are further classified into categories and sub-categories. In this preliminary evaluation, we consider four scenarios: (*i*) Scenario 1, used as a baseline for all the others experiments (its simulation parameters are reported in Table 1), (*ii*) Scenario 2, characterized by the N_a value variation, (*iii*) Scenario 3, characterized by the N_s value variation, (*iv*) Scenario 4: characterized by the TTL value variation.

4.3 Evaluation Metrics

For the analysis of experimental results, we consider the two network metrics previously introduced: the clustering coefficient and the characteristic path length. The clustering coefficient is computed in the directed graph based on each peer's N_a neighbors, with a total of $N_a(N_a - 1) = 20$ possible directed links between neighbors. The overall clustering coefficient $C(G)$ is computed by averaging across all peer nodes. The characteristic path length $L(G)$ is defined as the

average shortest path length across all pairs of nodes. Since in our simulations the network is not always strongly connected we used an alternative way to average shortest paths:

$$\tilde{L} = \left(\frac{1}{P} \sum_{p=1}^{\infty} l_p^{-1} \right)^{-1}, \tag{1}$$

where p is a pair of nodes and $P = N(N-1)$ is the number of all possible pairs. The characteristic path length $\tilde{L}(G)$ thus defined can be computed from all pairs of nodes irrespective of whether the network is connected. Another interesting network metric taken in account is the number of connected components. A connected component, which in the case of a directed graph is called strongly connected component (SCC), is a strongly connected subgraph, S, of the directed graph G, such that no vertex of G can be added to S and it still be strongly connected (informally, S is a maximal subgraph in which every vertex is reachable from every other vertex). C, $\tilde{L}$, and SCC are measured at each time step and averaged across simulation runs.

4.4 Simulation Results

Scenario 1 (a baseline). In the first scenario we investigate the system behavior when the simulation parameters are set to some representative, base values. The results obtained are used as a baseline to compare the algorithm performances in the others scenarios, when significant parameters are varied. Fig. 3

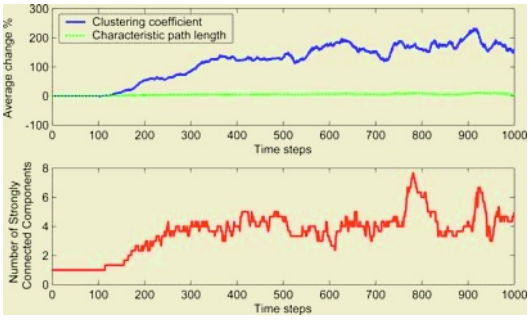
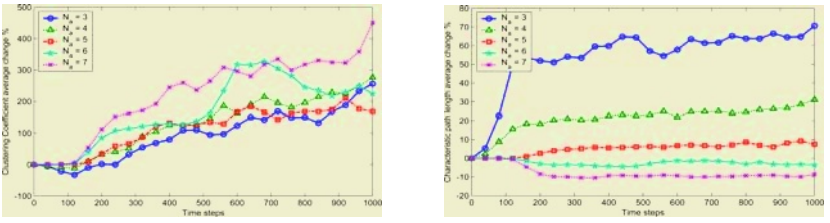


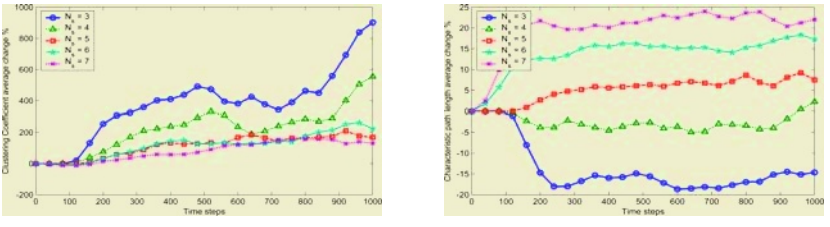
Fig. 3. Scenario 1: clustering coefficient, characteristic path length, and SCCs

plots the evaluation metrics taken into account. In particular, the top chart shows that the characteristic path length remains roughly equal to the initial random graph characteristic path length while the clustering coefficient increases rapidly and significantly, reaching a value that is, in average, 150% larger than that of the initial random graph. These conditions define the emergence of a *small world* topology in our peer network [19]. This is a very interesting finding, indicating that the peer interactions cause the peers to route queries in such a way that communities of users with similar interests cluster together to find quality results



(a) Clustering coefficient (b) Characteristic path length

Fig. 4. Scenario 2



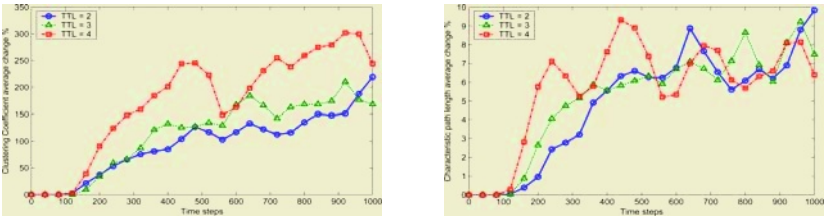
(a) Clustering coefficient (b) Characteristic path length

Fig. 5. Scenario 3

quickly, while it is still possible to reach any peer in a small number of hops. In the bottom of Fig. 3, the number of strongly connected components (SCC) is reported. It is easy to observe that the network splits into a few SCCs (it remains connected in the weak sense, based on undirected links). The number of SCCs is smaller than the number of groups: while the network is becoming more clustered and localized, peers continue to have access to most peers in other groups.

Scenario 2 (varying N_a). In this experiment, we choose to vary the N_a value, i.e. the maximum number of peers to which a query can be sent. The base N_a value is set to 5. We perform additional simulations investigating also systems with $N_a = 3, 4, 6, 7$. In Fig. 4(a) the clustering coefficient is plotted for the different values of N_a. The x-axis reports the time steps and the y-axis the percentage average variation of clustering coefficient. It is expected that the coefficient grows when N_a becomes larger, denoting the constitution of bigger peer groups with shared interests. The characteristic path length for different N_a values is plotted in Fig. 4(b). We observe that the characteristic path length stabilizes around a value that is lower for larger values of N_a. That is because there are more links and the network has a greater degree of compactness.

Scenario 3 (varying N_s). In this experiment, we study how the N_s value (i.e., the number of peers known at the simulation start-up) affects the system



(a) Clustering coefficient (b) Characteristic path length

Fig. 6. Scenario 4

performance. We perform simulations with $N_s = 3, 4, 5, 6, 7$. From the results (see Fig. 5) we note that the emerging networks present similar topological structures. When the N_s values grow the statistical differences between the initial network structure and the final one are less evident.

Scenario 4 (varying the TTL). In this experiment, we consider how to choose the TTL value, i.e. the maximum number of links traversed by a query. The base TTL value is set to 3. We perform additional simulations investigating also systems with $TTL = 2, 4$. It can be seen from Fig. 6 that the emerging of a *small world* topological structure is faster for higher TTL values. This is mainly caused by the fact that, during query propagation, an higher TTL value is associated with a deeper exploration of the network graph, allowing a peer to learn in a faster way the characteristics of the other peers.

5 Conclusion

In this paper, we presented a novel mechanisms for improving search efficiency in unstructured P2P networks, and we evaluated its efficiency. To address major limitations of unstructured P2P networks, we propose an adaptive routing algorithm in order to dynamically change the topology of the peer network, based on commonality of interests among users. Preliminary results confirm the idea that adaptive routing can properly work and that *small world* network topological structure can emerge spontaneously from the local interactions between peers. This is a very interesting finding, indicating that the peer interactions cause the peers to structure the overlay in such a way that communities of peers with similar interests cluster together in order to find quality results quickly, while it is still possible to access any network peer in a small number of hops.

References

1. Lv, Q., Caoa, P., Cohen, E., Li, K., Shenker, S.: Search and replication in unstructured Peer-to-Peer networks. In: Proc. ACM ICS. (2002)
2. Stoica, I., Morris, R., Karger, D., Kaashoek, M.F.: Chord: A scalable Peer-to-Peer lookup service for internet applications. In: Proc. ACM SIGCOMM 01. (2001) 149–160

3. Rowstron, A., Druschel, P.: Pastry: Scalable, distributed object location and routing for large-scale Peer-to-Peer systems. In: Proc. of the IFIP/ACM International Conference on Distributed Systems Platforms. (2001)
4. Ratnasamy, S., Francis, P., Handley, M., Karp, R., Shenker, S.: A scalable content-addressable network. In: Proc. ACM SIGCOMM 01. (2001) 161–172
5. Limewire: The Gnutella protocol specification (ver. 0.4). (http://www.limewire.com/developer/gnutella_protocol_0.4.pdf)
6. Ripeanu, M., Foster, I., Iamnitchi, A.: Mapping the Gnutella network: Properties of large scale Peer-to-Peer systems and implications for system design. IEEE Journal on Internet Computing, Special Issue on Peer-to-peer Networking (2002)
7. Adamic, L.A., Lukose, R.M., Puniyani, A.R., Huberman, B.A.: Search in power law networks. Phys. Rev. E64 (2001) 46135–46143
8. Liu, Y., Liu, X., Xiao, L., Ni, L.M., Zhang, X.: Location-aware topology matching in P2P systems. In: Proc. IEEE INFOCOM. (2004)
9. Yang, B., Garcia-Molina, H.: Effcient search in Peer-to-Peer networks. In: Proc. IEEE ICDCS. (2002)
10. Crespo, A., Garcia-Molina, H.: Routing indices for Peer-to-Peer systems. In: Proc. IEEE ICDCS. (2002)
11. Milgram, S.: The small world problem. Psychology Today **2** (1967) 60–67
12. Zhang, H., Goel, A., Govindan, R.: Using the small-world model to improve Freenet performance. In: Proc. IEEE INFOCOM. (2002)
13. Iamnitchi, A., Ripeanu, M., Foster, I.: Small-world file-sharing communities. In: Proc. IEEE INFOCOM. (2004)
14. Manku, G.S., Bawa, M., Raghavan, P.: Symphony: Distributed hashing in a small world. In: Proc. of USENIX Symposium on Internet Technologies and Systems. (2003)
15. Comellas, F., Mitjana, M.: Broadcasting in small-world communication networks. In: Proc. 9th Int. Coll. on Structural Information and Communication Complexity. (2002) 73–85
16. Bloom, B.H.: Space/time trade-offs in hash coding with allowable errors. Communications of the ACM **13** (1970) 422–426
17. Saroiu, S., Gummadi, K., Gribble, S.: A measurement study of Peer-to-Peer file sharing systems. In: Proc. ACM Multimedia Conferencing and Networking. (2002)
18. Sen, S., Wang, J.: Analyzing Peer-to-Peer traffic across large networks. IEEE/ACM Transactions on Networking **12** (2004) 212–232
19. Watts, D., Strogatz, S.: Collective dynamics of small-world networks. Nature **393** (1998) 440–442

Enhancing UDDI for Grid Service Discovery by Using Dynamic Parameters

Brett Sinclair, Andrzej Goscinski, and Robert Dew

School of Information Technology,
Deakin University,
Geelong, Vic 3217, Australia
{brettwil, ang, rad}@deakin.edu.au

Abstract. A major problem for a grid user is the discovery of currently available services. With large number of services, it is beneficial for a user to be able to discover the services that most closely match their requirements. This report shows how to extend some concepts of UDDI such that they are suitable for dynamic parameter based discovery of grid services.

1 Introduction

Naming and discovery are critical components of grids. Without the ability to discover the names of services that are available on a grid these services cannot be invoked. A grid service can be named by not just a name that is globally unique, but also by the parameters it portrays [6]. Service's parameters can be dynamic or static and can be used to expose important details of a service, such as its state. By using dynamic parameter based discovery a user is able to refine their search for services. This reduces the number of results received and increases the number of relevant results.

UDDI (Universal Description Discovery and Integration) [7] is a generic system for discovering Web services. It presents high level business information such as contact details and service lists in a format similar to a telephone directory. The indexing service of Globus, which is a toolkit for creating only one kind of grid, offers a registry for grid services that can be used for discovery.

This report shows the outcome of an initial study into dynamic parameter based discovery in grids by extending UDDI and a proof-of-concept in the form of a dynamic parameter based discovery broker suitable for service discovery in a grid.

2 Related Work

Grids are an emerging technology; the number of discovery systems which are being employed on grids today is small: UDDI is the current standard for discovery of Web services; and the index service is part of the Globus Toolkit.

UDDI [8] is a system to enable businesses to publish and locate Web services. Businesses can publish information about themselves and the Web services they provide. Interaction with a UDDI registry is done using a set of XML interfaces, protocols and programming APIs. Information about Web services is stored similar to

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 49 – 59, 2005.

that of a phone book. There are white pages which contain information about individual Web services, yellow pages which classify Web services, and the green pages which store information about individual businesses [2]. However, UDDI does not have the ability to aggregate information about the state of a grid service.

Globus contains, as a base service, a component called the index service. The index service is designed to aid in the discovery of resources and assist in the management of service data. Service data are additional parameters related to a service. The index service provides three systems [5]: a registry; service data providers; and a data aggregator. The registry provides a list of services currently published to the index service. This can be used for the discovery of resources. A service data provider provides additional information about a service which has been generated by an external program. The data aggregator can collect all information generated about a service and provide a simple access point to it. Thus, a name of a service can be discovered from the index service; however to obtain the most consistent service data the service must be invoked.

3 Dynamic Parameter Based Discovery Broker – Logical Design

This section presents a logical design of a centralized resource discovery broker that allows a published grid service to be discovered based on dynamic parameter values.

3.1 UDDI Data Model and UDDI Deficiencies for Grid Service Discovery

A fundamental concept of the UDDI standard is to represent data and meta-data about Web services. This information is represented in a standard way, catalogued and classified to allow for Web service discovery [1]. UDDI can be used to discover grid services; a grid service is a special case of a Web service.

The information about businesses and services are stored in four structures:

- businessEntity, businessService and bindingTemplate are in a parent child relationship, a parent can contain many children;
- businessEntity is the parent to businessService;
- the businessService is the parent to bindingTemplate structures; and
- the tModel structure is outside this parent child relationship.

One extension to Web services that a grid service introduces is service data. Service data provide information about a service, which the interface alone cannot provide. Every grid service has a basic set of service data elements, which contain information about the service, its interface, and its location. UDDI's data model is adept at storing information about Web services, however lacking the ability to record the service data of a grid service.

UDDI's inability to index service data increases the processing complexity of a client. A client wishing to discover a grid service and obtain it service data using UDDI is forced to complete two steps. First a client must query the UDDI registry for a set of grid services. Second the client must query the discovered grid services to obtain their current service data. The retrieved service data is then used to determine if the service meets the requirements of the client.

3.2 A Logical Design of a Centralised Discovery Broker by Enhancing UDDI

Using the concepts of notifications described by the OGSA [3] it is possible to create a discovery broker that has the ability to dynamically index and keep a consistent view of service data. The published service is effectively given a dynamic parameter name. This allows the discovery of a service to be based on static and dynamic information about the service. To store the additional information, the UDDI data model, is enhanced to incorporate the storage of service data.

A new data type called serviceData, which is the child of a single bindingTemplate, can be used to store service data. The serviceData type stores the current service data of the grid service described in the parenting bindingTemplate. This is shown in Figure 1. Furthermore, this figure shows the relationship between the UDDI data structures, which provide information about Web service and the businesses that are responsible for them.

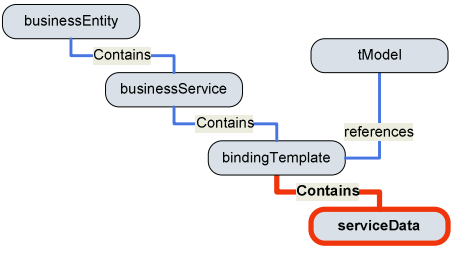


Fig. 1. Proposed data model incorporating service data

Service data are dynamic and change with respect to the state changes of a grid service. The OGSA describes the concept of notifications to monitor the changes of service data. Notifications can be used by the discovery broker to keep the data that is stored by the broker, consistent with the service data of a service. When a service is published to a discovery broker, the broker subscribes to that service to be notified of service data changes. When a grid service state changes a notification of the change, as well as the changed service data is sent to the discovery broker. This enables a registry to contain up to date information about a grid service. In this manner a client has the ability to query a centralized discovery broker on the high level business data, as well as the current service data of a grid service. The proposed model for a centralized discovery broker is shown in Figure 2.

A grid service is published to a discovery broker (message 1). In the publication process, data about the grid service and the service data are presented to the broker. The discovery broker then subscribes to the grid service, such that it is notified when the service data of the grid service changes (message 2). A grid service notifies the broker of any changes made to its service data (message 3). A client wishing to discover a service, based on its dynamic parameter name, only has to query the discovery broker (message 4).

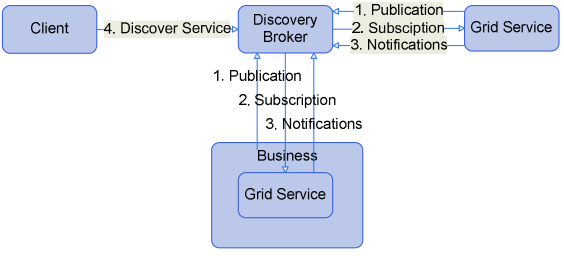


Fig. 2. Dynamic parameter based grid service discovery using proposed broker

The proposed discovery broker is composed of four major components: the interface; the query manager; the notification manager; and the published service information database, as shown in Figure 3.

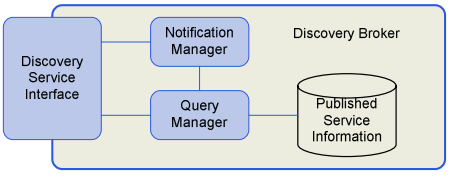


Fig. 3. Proposed discovery broker

This new model reduces the complexity of a client wishing to discover a grid service by removing the need for interaction between the client and the grid service; and reducing the need for the aggregation and querying of service data from multiple grid services. The model also decreases the demands on the grid service: a grid service is not continually queried by clients for its service data; and when the service data of a grid service change, only the centralized registry needs to be notified.

4 Centralized Dynamic Parameter Based Discovery Broker

This section demonstrates the construction of a centralized dynamic parameter based discovery broker following the design presented in Section 3. The discovery broker has been implemented as a grid service using the tools provided by the Globus toolkit. A grid service has two desirable properties. Firstly, it preserves its state between invocations, allowing the discovery broker to maintain information about publicized services. Secondly, it interoperates with all other grid services. The discovery broker developed is centralized. A UML diagram of the discovery broker showing the four major components is illustrated in Figure 4.

The DiscoveryBrokerImpl class implements all the functionality of the discovery broker; its interface is defined by the DiscoveryBroker interface. It inherits all functionality which the OGSA requires a grid service to implement from the

GridServiceImpl class. The notification manager, the query manager, the database and the interface are all contained by the DiscoveryBrokerImpl class.

The Discovery Broker Interface. This interface describes the operations that a client can invoke on the discovery broker. All operations which can be invoked on a grid service are specified by its portType. Section 3 identified three categories of operations that can be invoked on the discovery broker: publications, notifications and queries. The implementation of the discovery broker specifies the operations in each of these categories.

The operations for the publication of a service are:

- publishService(), which is used to publish a service; and
- unpublishService(), which is used to un-publish a service.

Queries are used to discover services; a number of different operations can be invoked to discover services in different ways. The names of the operations were derived from data structures used to store service information. The operations are:

- searchBusinessEntity(), search for a business based on a business's name;
- searchBusinessService(), search for a business by the type (category) of services it provides;
- searchBindingTemplate(), search for a particular service by name; and
- searchServiceData(), searches for a grid service that has a particular service data element with a specified value.

Only one operation is in the notifications category, deliverNotification(). It accepts messages from published grid services when their service data changes. This operation is a common operation to all grid services and is described by the OGSI [9].

These operations are not the only operations that the discovery broker must implement – all grid services must implement the GridService interface described by the OGSI and since the discovery broker accepts notifications the NotificationSink interface must also be presented. Figure 5 shows a UML diagram of the discovery broker interface. The interface, DiscoveryBroker, inherits the operations from the GridService and NotificationSink interfaces, which are provided by Globus.

Globus provides an implementation of all OGSI interfaces. It is therefore not necessary to define the messages for operations inherited from the GridService or NotificationSink interfaces. The only messages which must be specified are the messages unique to the DiscoveryBroker; they are presented in [8].

Query Manager. The query manager provides a simple means of querying the data stored about published services. Messages are received from either the discovery broker or the notification manager requesting that information is either retrieved from, modified in or stored in the database. A major functionality of the query manager is its ability to understand and query data stored in an XML document. Service data is encoded in XML; the messages received via notification come in this form. The query manager has the ability to parse this document and extract specific service data elements and their values for appropriate action.

Notification Manager. The notification manager implements one function, deliverNotification(), which receives push notification messages from grid services

whose service data has been modified. The message received is an XML document; the notification manager parses it, extracts service data, and then passes the service data to the database.

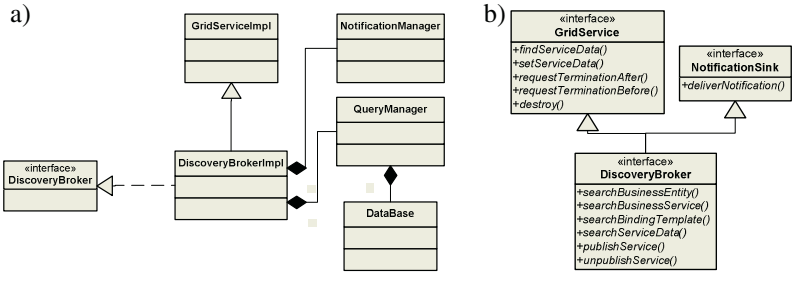


Fig. 4. UML representations of a) the discovery broker and b) the discovery broker interface

Database. There is a need for a data model for storing the published information about a service. The data model is hierarchical: each BusinessEntity contains a list of BusinessServices, each BusinessService contains a list of BindingTemplates, and each BindingTemplate contains a list of ServiceData. This hierarchical structure makes it relatively easy for the query manager to traverse when searching for records that need to be accessed, modified, deleted, or inserted.

5 Testing

This section shows that the proposed enhancement to UDDI and the application of the concepts of naming, notification and service data of OGSA are sound using evaluation of the dynamic parameter based discovery broker within a simple grid.

5.1 Test Environment

The dynamic parameter based discovery broker was implemented and tested on a small grid, consisted of two Pentium based machines connected by a hub. On each machine the Globus toolkit was installed and a set of simple services were deployed.

Two services were developed to test the discovery mechanism: a simple mathematics service, MathService; and a printing service, PrintService. Each service contained a single service data element that is dynamically updated when the service is invoked. This service data element is used to test discovery based on the dynamic parameters of a service.

The mathematics service allows two operations, add and subtract, to be conducted on a single value stored in the service. This value is retained between invocations of the service. A service data element, MathData, is exposed by the service. The printing service allows three operations to be conducted; turn on printer, turn off printer, and print a page. This service contains one service data element, PrintData, that contains two variables; the value in one indicates if the printer is on or off, and the other variable holds the total number of pages printed.

Two hypothetical businesses were created to own the services: business1 and business2. Each business owns two services, a MathsService and a PrintService, which are hosted on two separate machines.

The machine Globus01 is running three services a MathService owned by business1, a PrintService owned by business2, and the discovery broker. The machine Globus02 is running two services: a MathService owned by business2, and a PrintService owned by business1. These machines and hypothetical businesses where used to complete a series of tests on the capabilities of the discovery broker.

Clients interact with the discovery broker that is hosted on Globus01. Operations are invoked on the discovery broker by sending SOAP messages to the broker using HTTP. A response from the broker is returned to the client as SOAP message.

5.2 Test Scenarios

To test the discovery broker, two scenarios were designed. Each scenario is comprised of a logical sequence of steps. The output of each step was predicted and compared with the actual output. Scenario 1 tested the discovery of services using static information that emulates the discovery methods provided by UDDI. Scenario 2 tested discovery using the dynamic parameters of a service.

Scenario 1. Scenario 1 is used to test publication, un-publication and discovery of services. The following steps were conducted:

1. Publish the MathService owned by business1, using the publishService operation.
2. Publish the MathService owned by business2, using unpublishService.
3. Test whether services are published correctly by using searchBindingTemplate to search for services that have names that match the name “MathService”.
4. Un-publish the MathService owned by business1, using unpublishService.
5. Test, using the searchBindingTemplate operation, whether the service has been unpublished by searching for services that have the name “MathService”.

The results of publishService invoked in steps 1 and 2 are shown in Listing 1. These results have been gathered from the discovery broker and show the received information about the services which are being published.

Listing 1. Publication of MathSerices

```
Publish service:
Business Name: business1
Service Category: Math
Service Name: MathService
Location: http://globus01:8080/Math

Publish service:
Business Name: business2
Service Category: Math
Service Name: MathService
Location: http://globus02:8080/Math
```

The results from searchBindingTemplate invoked in step 3 are shown in Listing 2.

Listing 2. Search for MathService

```
Search for service: MathService
==========
business1
==========
Service Category: Math
Service Name: MathService
Location: http://globus01:8080/Math
==========
business2
==========
Service Category: Math
Service Name: MathService
Location: http://globus02:8080/Math
```

These results were gathered from the client application and were produced by a search for all services called MathService. They show that both published MathServices were discovered (which was the expected results).

Step 4 attempts to un-publish the MathService owned by business1 using the unpublishService operation. The output generated from this operation being applied to the discovery broker is shown in Listing 3.

Listing 3. Un-publication of MathService

```
Unpublish Service:
Business Name: business1
Service Category: Math
Service Name: MathService
Location: http://globus01:8080/Math
```

Step 5 searches (again) for the services named MathService, Listing 4. This time the results show that the only published service with the name MathService is the service owned by business2.

Listing 4. Results of search

```
Search for service: MathService
==========
business2
==========
Service Category: Math
Service Name: MathService
Location: http://globus02:8080/Math
```

This step has shown the un-publication of a service succeeded, as expected.

The testing from Scenario 1 has shown that the discovery broker allows for discovery based on a service's name.

Scenario 2. This scenario is designed to assess discovery based on dynamic parameters. This discovery is based on the current values of dynamic parameters of a service. The steps are:

1. All services are published to the discovery broker, using publishService.
2. Both printer services are turned off.
3. Using searchServiceData, a search is conducted for printer services which are on.
4. The printer service from business1 is turned on.
5. A search is conducted for printer services which are turned on.
6. The searchBusinessService operation is used to search for businesses which have services which provide mathematics (maths) services.

Listing 5 shows the results returned by the discovery broker when the publishService operation is invoked in step 1.

Listing 5. Publication of all services

```
Publish service:
Business Name: business1
Service Category: Math
Service Name: MathService
Location: http://globus01:8080/Math

Publish service:
Business Name: business2
Service Category: Math
Service Name: MathService
Location: http://globus02:8080/Math

Publish service:
Business Name: business1
Service Category: Printing
Service Name: PrintService
Location: http://globus02:8080/Print

Publish service:
Business Name: business2
Service Category: Printing
Service Name: PrintService
Location: http://globus01:8080/Print
```

Step 2 required both printers to be turned off. This was completed by calling the operation turnOff for each printing service; this operation modifies the service data element called PrintData. The invocation of the turnOff operation on a printing service generates notifications messages that are sent to the broker. The output generated by the discovery broker, when the notifications occur, show the source of the notification and the service data element which has changed. This can be seen in Listing 6. These results show that the notifications of service data are received.

Listing 6. Turning printers off

```
Notification received from: http://globus01:8080/Print
sde: PrintData
      status: OFF
      pages: 0
Notification received from: http://globus02:8080/Print
sde: PrintData
      status: OFF
      pages: 0
```

Step 3 searches for printers which have been turned on. As there were no printers turned on, no results were received. The output from the client is shown in Listing 7.

Listing 7. Search for printers which are on

```
Search for service: PrintService
Service data name: PrintData
Service data element name: status
Value: ON
null
```

Step 4 turns on the printer owned by business1. Listing 8 shows that the notification was received by the discovery broker.

Listing 8. Turn Printer On

```
Notification received from: http://globus02:8080/Print
sde: PrintData
      status: ON
      pages: 0
```

Step 5 searches (again) for printers which were on. One printer was found and that was the printer which was turned on in step 4 (which is what was expected). Listing 9 shows the output generated by the discovery broker.

Listing 9. Search for printers which are on

```
Search for service: PrintService
Service data name: PrintData
Service data element name: status
Value: ON
==========
business1
==========
Service Category: Printing
Service Name: PrintService
Location: http://globus02:8080/Print
```

Step 6 searches for all businesses that provided maths services. The results of this operation are shown in Listing 10. It is shown that both businesses are returned, as the both host math services, which is what was expected.

Listing 10. Search for Math services

```
Search for service category: Math
business1
business2
```

These six steps demonstrate the ability to complete discovery based on the dynamic parameters of a service. The scenario of discovering a printing service emphasizes the benefits of dynamic parameter based discovery. A user is able to specify a more detailed query using dynamic parameter values which returns higher quality results.

6 Conclusions

The design and development of a dynamic parameter based discovery technique for service discovery in a grid environment has been presented in this paper. This technique has taken advantage of an extension to UDDI such that it is suitable for dynamic parameter based discovery of grid services, and the concepts of naming, notifications and service data which were introduced by the OGSA.

The current implementation of the discovery broker is limited by the types of queries which can be conducted because currently only a small set of queries are defined. Future research could be conducted into the implementation of a flexible XML based query language to conduct queries, which is programming language and platform independent. To provide scalability research is currently being carried out where several brokers cooperate to produce a distributed dynamic parameter based discovery broker.

References

1. T. Bellwood, L. Clément, and C. v. Riegen, UDDI Spec Technical Committee Specification, http://uddi.org/pubs/uddi_v3.htm, accessed September 2004
2. M. Deitel, P. Deitel, B. Waldt, and L. Trees, Web Services a Technical Introduction: Prentice Hall, 2003.
3. I. Foster, C. Kesselman, J. Nick, and S. Tuecke, "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration.," presented at Open Grid Service Infrastructure WG, 2002.
4. A. Iamnitchi and I. Foster, "On Fully Decentralized Resource Discovery in Grid Environments," International Workshop on Grid Computing, Denvor, CO, 2001.
5. J. Joseph and C. Fellenstein, Grid Computing: Prentice Hall, 2004.
6. Y. Ni and A. Goscinski, "Evaluation of attributed names," Distributed Systems Engineering, vol. 2, pp. 87-101, 1995.
7. OASIS, Introduction to UDDI: Important Features and Functional Concepts, UDDI Technical White Paper, www.uddi.org, October 2004.
8. B. Sinclair, A. Goscinski, R. Dew, "An Attributed Naming Based Enhancement of UDDI to Offer Service Discovery in Globus 3", Technical Report TRC04/15, Deakin University, December 2004.
9. S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maquire, T. Sandholm, D. Snelling, and P. Vanderbilt, "Open Grid Services Infrastructure (OGSI) Version 1.0," 2003.

A New Approach for Efficiently Achieving High Availability in Mobile Computing

M. Mat Deris[1], J.H. Abawajy[2], and M. Omar[3]

[1] University College of Tun Hussein Onn,
Faculty of Information technology and Multimedia,
P.O.Box 101, 86400 Parit Raja, Batu Pahat, Johor
mustafa@kustem.edu.my
[2] Deakin University, School of Information Technology, Geelong, VIC, Australia
jemal@deakin.edu.au
[3] University of Malaya,Institute of Science Mathematics,
50603 Kuala Lumpur
mohd@um.edu.my

Abstract. Recent advances in hardware technologies such as portable computers and wireless communication networks have led to the emergence of mobile computing systems. Thus, availability and accessibility of the data and services become important issues of mobile computing systems. In this paper, we present a data replication and management scheme tailored for such environments. In the proposed scheme data is replicated synchronously over stationary sites while for the mobile network, data is replicated asynchronously based on commonly visited sites for each user. The proposed scheme is compared with other techniques and is shown to require less communication cost for an operation as well as provide higher degree of data availability.

1 Introduction

Mobile computing is born as a result of remarkable advances in the development of computer hardware and wireless communication technologies. One of the main objectives of mobile database computing is to provide users with the opportunity to access information and services regardless of their physical location or movement behavior. By enabling data and resources to be shared anywhere and anytime, mobile technology enhances distributed applications and allows higher degrees of flexibility in distributed databases [1, 2]. As a result, mobile database systems have been evolving rapidly [10] and it is expected that in the near future millions of users will have access to on-line distributed databases through mobile computers [13]. However, the restrictions imposed by the nature of the wireless medium and the resulting mobility of data consumers and data producers make the capability of the mobile computing to provide users with access to data where and when they need it a very difficult proposition.

Replication techniques can be used to increase data availability and accessibility to users despite site failure or communication disconnections [6]. In this paper, a new data replication and management scheme tailored for mobile computing environments

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 60 – 68, 2005.

is discussed. Mobile computing is basically an ad hoc network commonly built on a mix of stationary and mobile networks [3] where the stationary network consists of fixed servers and storage subsystems while mobile network consists of mobile hosts. In the proposed replication scheme, data will be replicated synchronously in a manner of logical three dimensional grid structure on the fixed network. For the mobile network, data will be replicated asynchronously based on commonly visited sites for each user. The proposed scheme is then compared with the Tree Quorum (TQ) [8] and Grid Configuration (GC) [5] techniques and shown that the proposed technique provides higher data availability.

The rest of the paper is organized as follows: In Section 2, the system architecture, and the problem statement are discussed. In Section 3, the proposed replica management technique is presented. In Section 4, the performance comparison with TQ and GC techniques is given. Finally, the conclusion and future directions are given in Section 5.

2 Background and Problem Statement

In this section, we present the main architecture of mobile computing on top of which we define our replica control and management model. The problem statement is discussed.

2.1 System Architecture

As in [4], we view a mobile DBMS computing environment as an extension of a distributed system. As in [7, 11–13], the system consists of two basic components: fixed network component and mobile network component.

2.1.1 Fixed Component

The fixed component consists of Wired Network (FN) and Fixed Host(FH) units. A FH is a computer in the fixed network which is not capable of connecting to a mobile unit.

The model consists of three layers: the source system, the data access agent, and the mobile transaction. The Source System represents a collection of registered systems that offer information services to mobile users. Examples of future services include white and yellow pages services, public information systems including weather, and company-private database/information systems. In our model, a system in this layer could be any existing stationary system that follows a client-server or a peer-to-peer architecture. Data in the source system(s) is accessed by the mobile transaction through the Data Access Agent (DAA). Each base station hosts a DAA. When it receives a transaction request from a mobile user, the DAA forwards it to the specific base stations or fixed hosts which contain the needed data and source system component. Each DAA contains location tables which indicates, by transaction or sub-transaction, the correct MDBS (Multidatabase System) or DBMS to process the request. When the mobile user is handed over to another base station, the DAA at the new station receives transaction information from the old base station.

2.1.2 Mobile Component

The mobile component consists of Wireless Network (WN), Mobile Unit (MU), Base Stations (BS) and Base Station Control (BSC) units. Base stations are capable of

connecting with a mobile unit and are equipped with a wireless interface. Each BS covers a particular area, called a cell and acts as an interface between mobile computers and fixed hosts. Its responsibilities include keeping track of the execution status of all mobile transactions concurrently executing, logging recovery information, and performing needed checkpointing.

The wireless interface in the base stations typically uses wireless cellular networks. Ericson GE Mobidem is an example of a cellular network that uses packet radio modems. The wireless interface can also be a local area network, of which NCR WaveLan is an example. Operations of multiple BSs are coordinated by Base Station Controller (BSC).

A MU is a mobile computer which is capable of connecting to the fixed network via a wireless link. Mobile units can move from one cell to another. This process of moving is referred to as a handoff. The mobile user accesses data and information by issuing read and write transactions. We define the Mobile Transaction as the basic unit of computation in the mobile environment. It is identified by the collection of sites (base stations) it hops through. These sites are not known until the transaction completes its execution.

2.2 Problem Statements

The use of wireless communication makes the availability of data and the reliability of the systems one of the most important problems in mobile computing environments. The problem is, given a system composed of F fixed sites and M mobile sites both of which are prone to failure, how to provide the users the ability to access information and services regardless of their physical location or movement behavior or system component state (i.e., operational or failed). A site is either operational (i.e., connected) or failed (i.e., disconnected or completely failed). When a site is operational, the copy of the data at that site is available; otherwise it is unavailable. The system should be able to provide as much functionality of network computing as possible within the limits of the mobile computer's capabilities. In the context of database applications, mobile users should have the ability to both query and update the databases whenever they need and from anywhere.

However, mobile computing introduces a new form of distributed computation in which communication is most often intermittent, low-bandwidth, or expensive, thus providing only weak connectivity. One way to hide this variability from users and to provide a responsive and highly available data management service is by replicating service state and user data at multiple locations. Mobile computing environment poses several unique challenges to replication algorithms, such as diverse network characteristics (latencies and bandwidths), node churn (nodes continually joining and leaving the network). Moreover, it is expected that in the near future millions of users will have access to on-line distributed databases through mobile computers [3,4]. Hence, a scalable replication technique that ensures data availability and accessibility, despite site failure or communication failure, is an important infrastructure in mobile computing environments. Moreover, to efficiently manage replicas that are not uniformly well accessible to each other, replication algorithms must take their non-uniform communication costs into account, and allow applications to make different

tradeoffs between consistency, availability and performance. Their complexity motivates the need for a reusable solution to manage data replication and consistency in new distributed services and applications.

3 Proposed Replication Technique

3.1 Hybrid Replication

Data replication is generally a complex task resulting from potential. In the fixed network, the data file is replicated to all sites, while in the mobile network, the data file is replicated asynchronously at only one site based on the most frequently visited site. We assume that the mobile network consists of M sites labeled $S_1,S_2,\ldots,S_M$. For the mobile network, a site will replicate the data asynchronously analogous to the concepts of *Check-out* proposed in [3]. In check-out mode, the site S_i wants to disconnect and be able to update a set of data items $I(S_i)$. The disconnected site has complete and unlimited access to $I(S_i)$ while the remaining system has complete access to the database other than $I(S_i)$. The 'commonly visited site' is defined as the most frequent site that request the same data at the fixed network (the commonly visited sites can be given either by a user or selected automatically from a log file/ database at each center). This site will replicate the data asynchronously, therefore it will not be considered for the read and write quorums.

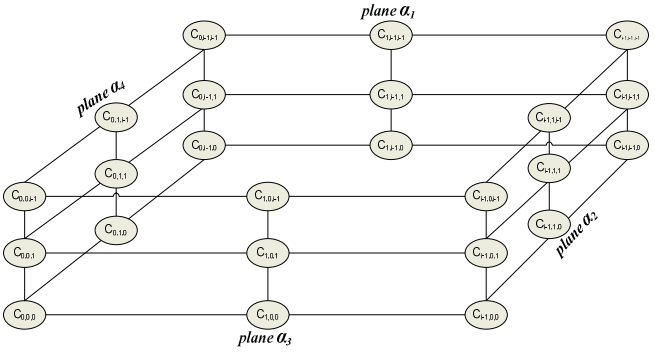


Fig. 1. The organization of the fixed networks with four planes (i.e., α_i) denotes planes for fixed network. The circles in the grid represent the sites

The proposed replication strategy is called hybrid replication as it has different ways of replicating and managing data on fixed and mobile networks. As in [5, 6, 8, 10, 12], our approach imposes a certain structure on part of the system. However, our approach differs from [5, 8, 10, 12] in that such systems are highly inefficient, and are only useful in systems with a high read/write operation ratio. Our environment consists of two types of networks, i.e., the fixed network and the mobile network. For the fixed network, all sites are logically organized in the form of three-dimensional

grid structure (TDGS). For example, if the network consists of N sites, TDGS is logically organized in the form of box-shape structure with four planes (i.e., α_1, α_2, α_3, and α_4) as shown in Fig. 1.

The environment has two types of networks, i.e., the fixed network and the mobile network. For the fixed network, all sites are logically organized in the form of three dimensional grid structure (TDGS). For example, if the network consists of N sites, TDGS will logically organized in the form of box-shape structure with four planes (i.e., α_1, α_2, α_3, and α_4), while in the mobile sites, each copy of the object is located at each site. A site is either operational or failed and the state (operational or failed) of each site is statistically independent to the others. When a site is operational, the copy at the site is available; otherwise it is unavailable.

3.2 Consistency Maintenance

3.2.1 Read Operations on Fixed Network

For a TDGS quorum, read operations on an object are executed by acquiring a read quorum that consists of any hypotenuse copies. In Fig. 1, copies $\{C_{0,0,0},C_{l-1,l-1,l-1}\}$, $\{C_{0,0,l-1},C_{l-1,l-1,0}\}$, $\{C_{0,l-1,l-1},C_{l-1,0,0}\}$, or $\{C_{l-1,0,l-1},C_{0,l-1,0}\}$ are hypotenuse copies any one pair of which is sufficient to execute a read operation. Since each pair of them is hypotenuse copies, it is clear that, read operation can be executed if one of them is accessible, thus increasing the fault-tolerance of this technique. We assume for the read quorum, if two transactions attempt to read a common data object, read operations do not change the values of the data object.

3.2.2 Write Operations on Fixed Network

For the fixed network, a site $C_{(i,j,k)}$ initiates a TDGS transaction to write its data object. For all accessible data objects, a TDGS transaction attempts to access a TDGS quorum. If a TDGS transaction gets a TDGS write quorum without non-empty intersection, it is accepted for execution and completion, otherwise it is rejected. Since read and write quorums must intersect and any two TDGS quorums must also intersect, then all transaction executions are one-copy serializable

Write operations are executed by acquiring a write quorum from any plane that consists of: hypotenuse copies and all vertices copies. For example, if the hypotenuse copies, say $\{C_{0,0,0}, C_{l-1,l-1,l-1}\}$ are required to execute a read operation, then copies $\{C_{0,0,0}, C_{l-1,l-1,l-1},C_{l-1,l-1,0},C_{0,l-1,l-1},C_{0,l-1,0}\}$ are sufficient to execute a write operation, since one possible set of copies of vertices that correspond to $\{C_{0,0,0},C_{l-1,l-1,l-1}\}$ is $\{C_{l-1,l-1,l-1},C_{l-1,l-1,0},C_{0,l-1,l-1},C_{0,l-1,0}\}$. Other possible write quorums are $\{C_{0,0,0},C_{l-1,l-1,l-1},C_{l-1,l-1,0},C_{l-1,0,l-1},C_{l-1,0,0}\}$,$\{C_{l-1,l-1,l-1},C_{0,0,0},C_{0,0,l-1},C_{l-1,0,l-1},C_{l-1,0,0}\}$,$\{C_{l-1,l-1,l-1},C_{0,0,0}, C_{0,0,l-1},C_{0,l-1,l-1},C_{0,l-1,0}\}$, etc. It can be easily shown that a write quorum intersects with both read and write quorums in this technique.

3.2.3 The Correctness of TDGS

In this section, we will show that the TDGS technique is one-copy serializable. We start by defining sets of groups called coterie [9] and to avoid confusion we refer the sets of copies as groups. Thus, sets of groups are sets of sets of copies.

Definition 3.1. Coterie. Let U be a set of groups that compose the system. A set of groups T is a coterie under U iff,

i. $G \in T$ implies that $G \neq \emptyset$ and $G \subseteq U$.
ii. If $G, H \in T$ then $G \cap H \neq \emptyset$ (intersection property)
iii. There are no $G, H \in T$ such that $G \subset H$ (minimality).

Definition 3.2. Let η be a group of hypotenuse copies and ω be a group of copies from any plane that consists of hypotenuse copies and all copies which are vertices as shown in Fig. 2. A set of read quorum, R, can be defined as

$R = \{\eta_i \mid \eta_i \cap \eta_j = \emptyset, i \neq j\}$, and
a set of write quorum, W, can be defined as
$W = \{\omega_i \mid \omega_i \cap \omega_j \neq \emptyset, i \neq j$, and $\omega_i \cap \eta_j \neq \emptyset$ for $\eta_j \in R\}$

By definition of coterie, then W is a coterie, because it satisfies all coterie's proper ties.

3.2.4 Correctness Criterion

The correct criterion for replicated database is one-copy serializable. The following theorem shows that TDGS meets the one-copy serializable criterion.

Theorem 3.1. The TDGS technique is one-copy serializable.

Proof: The theorem holds on condition that the TDGS technique satisfies the quorum intersection properties, i.e., write-write and read-write intersections. Since W is a coterie and by Definition 3.2, then it satisfies read-write and write-write intersection properties. □

3.2 Write Operation on Mobile Network

For the Check-out in mobile network, if site S_i wants to disconnect and be able to write a particular data object, it declares its intention to do so before disconnection and "check-out" or "takes" the object for writing. This can be accomplished by obtaining a lock on the item before disconnection. An object can only be checked out to one site at a time. In order to maintain serializability in check-out mode, some of the sites are prevented from accessing the objects that do not 'belong' to it. By using two-phase locking mechanism, if a site that wishes to disconnect, say, S_i, it acquires a write lock on the object it wants to update while disconnected. The disconnect procedure is as follows:

i. S_i tells the nearest site "proxy" from the fixed network to check-out.
ii. At the same time, S_i initiates a pseudo-transaction to obtain write locks on the items in $I_L(S_i)$
iii. If the pseudo-transaction is successful, S_i disconnects with update privileges on the items in $I_L(S_i)$.

Otherwise, these data items are treated as read-only by S_i
During the reconnection, S_i must go through a procedure as follows:

i. When S_i reconnects, it contact the proxy from the fixed network.
ii. The updated objects from S_i will be replicated to the proxy
iii. S_i release the corresponding lock

The site wishing to disconnect, S_i, check-out the desired data items with pseudo-transactions and sign-off. S_i has read/write access of the check-out items, $I_L(S_i)$ only. The remaining connected sites in the system have read/write access other than $I_L(S_i)$. The proxy applies the TDGS technique to replicate updated data items from S_i to diagonal sites.

In order to preserve correctness, it must be possible to serialize all of the transactions executed by S_i during disconnection at the point in time of disconnection. This can be done if:

i. Only the data items in S_i write locked by transaction at disconnect time can be modified during disconnect.
ii. The data items which are write locked at disconnect time can neither be read nor written by other site.
iii. Data items not write locked by transaction at disconnect time are treated read-only by S_i during disconnect.

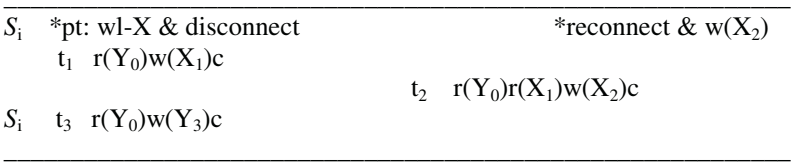


Fig. 2. Check-out mode with mobile read

Fig. 2 shows an example of serializability executions during disconnection. Time proceeds from left to right and the (*) indicates the disconnection and reconnection points in time. X_i indicates version of data item X written by transaction *i*. In Fig. 2, S_i first acquire a write lock on X with pseudo-transaction (pt) and disconnects. t_1 and t2 are examples of transactions that may be executed at the disconnected site S_i. S_j remains connected and executes transaction t_3. Notice that S_i can execute transaction during its disconnection that read all other database items without getting read locks before disconnection on those items. This is due to the fact that all of S_i's transactions will be reading versions that existed at disconnect time.

This will guarantee serializability because each transaction at a disconnected site respects two phase locking. Thus, in Fig 2, t_3 executes under the condition that X is write locked by a pseudo-transaction that were ongoing when S_i disconnected. The equivalent and correct serial order of these transactions is t_1,t_2,t_3.

4 Performance Comparisons

Let p be the probability that each site is operational. The write availabilities of TDGS, GC and TQ techniques are compared in Fig. 3 when the number of copies is set to 40 (i.e., N = 40) . We assume that all copies have the same availability. Fig. 3 shows the TDGS technique has better performance of availabilities when compared to the GC and TQ techniques.

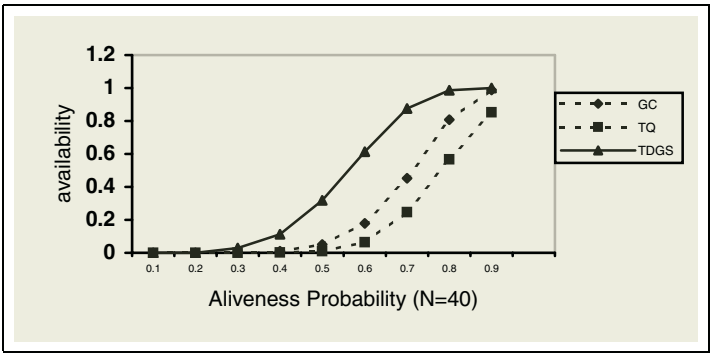


Fig. 3. Comparison of the write availability between TDGS, GC and TQ for N=40

For example, when an individual copy has availability of 70%, the write availability in the TDGS is more than 87%, whereas the write availability in the GC is approximately 45% and the write availability in the TQ is approximately 25%.

5 Conclusions and Future Directions

One of the main objectives of mobile database computing is to provide users with the opportunity to access information and services regardless of their physical location or movement behavior. However, this new infrastructure presents tremendous challenges for data management technology, including huge-scale, variable, and intermittent connectivity; location- and context-aware applications; bandwidth, power, and device size limitations; and multimedia data delivery across hybrid networks and systems. In the presence of frequent disconnection failures, data availability and accessibility from anywhere at any time is not easy to provide. One way to cope with this problem is through data and service replications. To this end, we proposed a new replication technique to manage the data replication for mobile computing environments. The proposed approach imposes a logical three-dimensional grid structure on data objects. We showed that the proposed approach presents better average quorum size, high data availability, low bandwidth consumption, increased fault-tolerance and improved scalability of the overall system as compared to standard replica control techniques.

References

1. M. Nicola and M. Jarke,"Performance Modeling of Distributed and Replicated Databases",*IEEE Trans. And Knowledge Engineering*, vol. 12, No.4, pp. 645-670, 2000.
2. O. Wolfson, S. Jajodia, and Y. Huang,"An Adaptive Data Replication Algorithm,"*ACM Transactions on Database Systems*, vol. 22, no 2 (1997), pp. 255-314.
3. J.Holliday, D.Agrawal, and A.El. Abbadi,"Disconnection Modes for Mobile Databases",*Journal of Mobile Network*, *Kluwer*, Vol.8, pp. 391-402,2002.
4. Budiarto, S. Noshio, M. Tsukamoto,"Data Management Issues in Mobile and Peer-to-Peer Environment", *Data and Knowledge Engineering, Elsevier*, 41 (2002), pp. 183-204.

5. D Agrawal and A.El Abbadi, "Using Reconfiguration For Efficient Management of Replicated Data,"*IEEE Trans. On Know. and Data Eng*.,vol.8,no. 5 (1996), pp. 786-801.
6. J.F. Paris and D.E. Long, "Efficient Dynamic Voting Algorithms," Proc. Fourth *IEEE Int'l Conf. Data Eng*, Feb.1988,pp. 268-275.
7. M.Mat Deris,J.H. Abawajy,H.M. Suzuri," "An Efficient Replicated Data Access Approach for Large-Scale Distributed Systems", *Proc.IEEE International Symposium on Cluster Computing and the Grid*, vol 2660, Chicago, April 2004,pp 223-229.
8. D Agrawal and A.El Abbadi, "The Tree Quorum technique: An Efficient Approach for Managing Replicated Data,"*Proc.16th Int'l Conf. On Very Large Data Bases* ,Aug. 1990, pp. 243-254.
9. D Agrawal and A.El Abbadi, "The generalized Tree Quorum Technique: An Efficient Approach for Managing Replicated Data,"*ACM Trans. Database Systems*,vol.17,no.4,(1992),pp.689-717
10. S.Y. Cheung, M.H. Ammar, and M. Ahmad,"The Grid Protocol: A High Performance Schema for Maintaining Replicated Data," *IEEE Trans. Knowledge and Data Engineering*,vol. 4, no. 6,pp. 582-592, 1992.
11. S.Jajodia and D. Mutchles,"Dynamic Voting Algorithms for Maintaining the Consistency of a Replicated Database,"*ACM Trans. Database Systems*,vol 15,no. 2(1990), pp. 230-280.
12. M. Maekawa," A √n Algorithm for Mutual Exclusion in Decentralized Systems,"*ACM Trans. Computer Systems*,vol. 3,no. 2,pp. 145-159, 1992.
13. Budiarto, K. Harumoto, M. Tsukamoto, S. Nishio, T. Takine, Replica allocation strategies for mobile databases, IEICE Transactions on Information and Systems E81-D (1) (1998) 112-121.

A Flexible Communication Scheme to Support Grid Service Emergence

Lei Gao and Yongsheng Ding

College of Information Sciences and Technology,
Donghua University, Shanghai 200051, P. R. China
ysding@dhu.edu.cn

Abstract. The next generation grid systems exhibit a strong sense of automation. To address such a challenge, Our previous work has viewed a grid as a number of interacting agents and applied some key mechanisms of natural ecosystems to build a novel grid middleware system, where a collection of distributed agents are searched, assembled, organized and coordinated to emerge desirable grid services. All these actions of agents depend on an effective communication scheme.

In this paper, we design a flexible communication scheme to implement the complicated coordination strategies among agents, including a RMI-IIOP-based transport mechanism, an ecological network communication language, and an ecological network interaction protocol from low to high implementation strategy. To test our hypothesis that grid services with desired properties can emerge from individual agents via our communication scheme, simulations of resource discovery service are carried out. The results prove that the scheme can well support this kind of bottom-up approach to build desirable services in grid environments.

1 Introduction

Grid systems have evolved over nearly a decade to enable large-scale flexible resource sharing among dynamic virtual organizations (VOs) [1]. Next generation grid systems are heading for globally collaborative, service-oriented and live information systems that exhibit a strong sense of automation [2], such as autonomic configuration and management. These "automatic" features, resembling the self-organizing and the self-healing properties in natural ecosystems, give us inspiration for designing next generation grid systems by associating them with some key concepts and principles in ecosystems.

Our previous work have viewed a grid as a number of interacting agents and applied some key mechanisms of natural ecosystems to build a grid middleware system named Ecological Network-based Grid Middleware (ENGM) [3]. Grid services and applications in ENGM can emerge from a collection of distributed and autonomous agents as the creatures living in a large ecosystem. Searching, assembling, organizing and coordinating of agents to emerge a desirable grid service depend on effective communication and cooperation scheme. However, we

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 69–78, 2005.

focus the discussion solely on issues of the ENGM architecture and its dynamic service design, and do not refer the communication scheme of ENGM in [3].

The implementation of communication scheme in most of current grid middleware systems relies on traditional communication standards, such as Message Passing Interface (MPI) and Parallel Virtual Machine (PVM). For example, MPICH-G2 [4] is a typical grid-enabled MPI implementation built on top of services provided by the Globus Toolkit. Agent technologies have recently become popular in the grid community, thus the research on grid computing with agent communication is becoming a hotspot. The start-point/endpoint paradigm, a key concept of the Nexus communication, has been successfully used to the world of agents, which resulted in the Southampton Framework for Agent Research [5]. The Control of Agent Based System Grid [6] is based on Java Remote Method Invocation for inter-agent communications, to integrate heterogeneous agent-based systems, mobile agent systems, object-based applications, and legacy systems.

In this paper, we describe how to combine agent communication with grid computing for emerging desired services in ENGM, and explain the rationale behind our design. We expect, via our communication scheme, emergent grid services have not only corresponding common functionalities, but also some life-like properties such as survivability and adaptability. Agent communication is generally broken down into three sub-sections [7]: interaction protocol, communication language, and transport protocol. Interaction protocol is the high-level strategy pursued by the agents that govern their interactions with other agents. The communication language is the medium through which the attitudes regarding the exchange content are communicated. The transport protocol is the actual transport mechanism using the communication languages.

Hence, next section first overviews ENGM architecture. Based on it, our communication scheme is presented including agent inner-communication module, Ecological Network Communication Language (ENCL), transport mechanism based on Remote Method Invocation over Internet Inter-Orb Protocol (RMI-IIOP [8]), and Ecological Network Interaction Protocol (ENIP) in Section 3. Considering interaction protocol is service dependent, we take the example of resource discovery service to depict the design of ENIP. To test our hypothesis about desired grid services can emerge from agents via the schema, a simulation on resource discovery is carried out in Section 4. Section 5 concludes our efforts.

2 The Architecture of Ecological Network-Based Grid Middleware

From the implementation point of view, from the bottom up, the architecture of ENGM system is presented as three-layers: Heterogeneous and Distributed Resources layer, ENGM layer, and Grid Applications for VOs layer.

(1) *Heterogeneous and Distributed Resources* consist of different types of resources available from multiple service providers distributed in grids. Via a java virtual machine, an ENGM platform can run on a heterogeneous distributed sys-

tem that established in a network node. (2) *ENGM* provides the services support a common set of applications in distributed network environments. It is made up by ENGM functional modules, ENGM core services, grid agent survivable environment, and emergent grid common service. (a) In ENGM functional modules layer, low-level functions related to management of networks and systems are dealt with. (b) ENGM core services layer provides a set of general-purpose runtime services that are frequently used by agents such as naming service and niche sensing service. (c) Grid agent survivable environment is runtime environment for deploying and executing agents that are characterized by high demand for computing, storage and network bandwidth requirements. (d) Emergent grid common services part is kernel of middleware and responsible for resource allocation, authentication, information service, task assignment, and so on. These common services are emerged from those autonomous agents. (3) *Grid Applications for VOs* use developing kits and organize certain agents and common services automatically for special purpose applications.

ENGM is fully implemented by Java language. Technically, an agent in it is a combination of a Java object, an execution thread, a remote interface for network communication, and a self-description. These four simple pieces together create an agent, a full-fledged autonomous process in ENGM.

3 Agent Communication Scheme

3.1 Agent Inner-Communication Module

Each agent follows a set of simple behavior rules (such as migration, replication and death) and implements a functional component related to its service. An application is created out of the interactions of multiple agents. A collection of agents can form a community as to a certain law. A community is able to emerge a higher-level service. These formed communities can be viewed as emerging agents that organize a higher-level community to provide complex services.

These agents fall into two categories: grid user agents (GUAs) and grid service agents (GSAs). A GUA represents a kind of user tasks. While GUAs are used to comprise the main components of ENGM, such as Grid Information Ser-

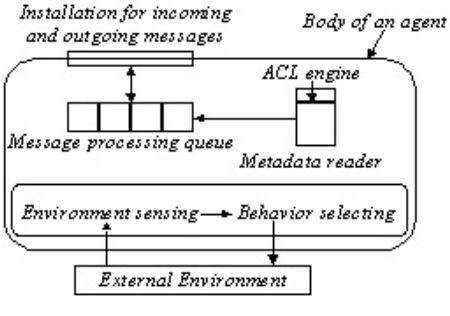


Fig. 1. Structure of agent inner-communication module

vice Agent (GISA). An agent is represented on its unit function that is defined as a metadata structure. Fig. 1 demonstrates the structure of a generic agent inner-communication module. The *metadata reader* is used for an agent to obtain another agent's metadata. The metadata of an agent consists of *agentID* (a global unique identifier of an agent), *agentAddress* (location of this agent), *serviceType* (service type of this agent), *serviceDescription* (description of service information) and *relationshipDescription* (information about the its acquaintances, agents know each other are called *acquaintances*). When an agent receives a message via the *installation for incoming and outgoing messages* from other platform or other agent, it inserts the message in its *message processing queue*. Each agent uses an independent thread to continuously sense the nearby environment, invoke most suitable behavior for the current environment condition.

3.2 Ecological Network Communication Language

An agent's communications are structured and constrained according to predefined speech act-based messages, usually referred to as *performatives*, which together make up an agent communication language (ACL). We reuse FIPA ACL [9] to design a high-level language called Ecological Network Communication Language (ENCL), as oppose to traditional specified interface definitions to achieve flexible interactions.To communicate using this language, each agent has to have an interpretation engine, as shown in Fig. 1.

An ENCL message contains a set of one or more parameters that can be arranged randomly in the message. Precisely needed parameters for an effective communication will vary according to the different situations. These message parameters are selected from those in FIPA ACL, including *performative-name*, *sender*, *receiver*, *content*, *language*, *ontology*, *conversation-id*, *reply-with*, *in-reply-to*, and *interaction-protocol*. We reuse part of FIPA ACL performatives and specify some new performatives. Performatives available in the ENCL are listed in the following: *request*, *agree*, *refuse*, *not-understood*, *inform*, *failure*, *defray*, *query-ref*, *query-if*, *advertise*, and *convene*. Among them, *advertise*, *convene*, and *defray* are specific to ENCL. An agent may send an advertise message to notify nearby agents of its existence (publishing its information and its capabilities such as the service it provides). An agent may send convene message to ask for interaction partner(s) to complete a certain task. Whenever receiving a service, the GUA returns a defray with a credit (a reward or a penalty) to an agent for its service. Also, we use XML (eXtensive Markable Language) as primary coding language of ENCL because of its strong capability of data description and metadata aspect. By combining the advantages of the ACL and XML, ENCL can flexibly implement the complicated coordination strategies.

3.3 Transport Mechanism

ENGM message transport is required to offer a reliable, orderly and accurate messaging stream, to detect and report errors to the sender, and to facilitate agent mobility. ENGM uses RMI-IIOP as transport protocol for ENCL messages. RMI-IIOP, the Sun Microsystems-IBM vision for distributed communi-

cation, provides the robustness of CORBA (Common Object Request Broker Architecture) and the simplicity of Java RMI. It is an application-level protocol on top of TCP. What's more, RMI-IIOP makes it possible that serialized Java objects can transmit among different components. Adopting this technology is also based on the following considerations: 1) It takes Java technology as the core, which offers simple and effective way for Java-based agent distributed computing. 2) It can find, activate and collect the useless agents. It allows the migration of data and codes, which facilitates the software implementation of autonomous and mobile services. 3) Using it, ENGM can flexibly implement agent-to-agent, community-to-community and agent-to-community communications.

Although a community is formed by the interactions of multiple assembled agents, it stresses on the global pattern. That is, it can act as a whole to communicate with other agents or communities. For example, when interacting, an agent needs to send a message to a community that provides a specific service. To address such a challenge, we introduce the concept of "channel" that represents a specific queue. If two communities use different channels, they would not affect each other. Sending or receiving messages in a certain channel is similar to tuning to a certain TV channel or radio frequency.

3.4 Ecological Network Interaction Protocol

To ensure interoperability among agents in wide-area grid environments, a sequence of interactions between two agents (a community can also be viewed as an emerging agent) is defined by an Ecological Network Interaction Protocol (ENIP). ENIP provide a reusable solution that can be applied to various kinds of message sequencing we encounter between agents. Only the agents that have implemented the same an ENIP are allowed to interact. Besides some conventional ENIPs defined by the ENGM platform, service developers may also define an ENIP relevant to their applications.

In the paper [10], we proposed a social network-inspired resource discovery approach, including some key models such as resource matching model, request forwarding model, and service defraying model. In this section, we further develop the approach as a resource discovery service of the ENGM platform and take the service as an application case to introduce ENIPs that are depicted as the Agent Unified Modeling Language (AUML) [11].

The next generation grid systems will exist in an unstable environment where a large number of nodes join and leave the grid frequently. Resource discovery service is a critical activity in such an environment: given a description of resources, it will return a set of contact addresses of resources that match the description. We consider a collection of agents that form a relationship network to discover desired resources. The presented approach is proposed as three-phases: GISA relationship construction, request processing strategy and trust-based reconstruction of relationship. GISA relationship construction is responsible for collecting and updating information on the currently participating peers and forming the relationship network; request processing strategy performs the search itself;

trust-based reconstruction of relationship makes the necessary preparations for a more efficient search. We describe the latter two phases using AUML.

Fig. 2 depicts a protocol expressed as an AUML sequence diagram for the ENGM request forwarding protocol. When invoked, a GISA acting as a request sender sends a request message with a given TTL (Time to live) to an acquaintance agent that has biggest probability to respond the request. The request receiver can then choose to respond to the sender by: refusing to process the request, responding the request and then forwarding the request (or only forwarding the request, which depends on the conditions), notifying the attempt to process is failed, or indicating that it did not understand. Depending on the conditions it contains, the symbol of decision diamond indicates it can result in zero or more communications. To implement an ENIP, the protocol specification

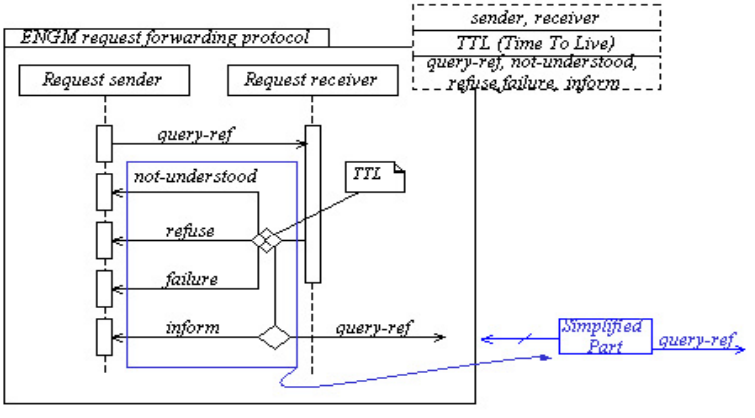


Fig. 2. A generic ENGM request forwarding process expressed as a template package

requires demonstrating the detailed processing that takes place within an agent. The AUML statechart can express the process that higher-level agents consist of aggregations of lower-level agents in ENGM interactions. In addition, it can also specify the internal processing of a single agent. Fig. 3 depicts the detailed request-respond process of proposed discovery approach using a combination of diagrams. On the bottom of it, the statechart that specifies request processing that takes place within a request originator. Here, the sequence diagram indicates that the agent's process is triggered by a query-ref message and ends with a defray message which can update and strengthen the relationship network to perform discovery better. To show explicitly, the sequence diagram has been simplified. We use a box with "Simplified Part" to replace some communicate operations, as shown in Fig. 2. A GISA relies in some GISAs and mistrusts in others to achieve its purpose. The reliability is expressed through a *trustCredit* value with which each GISA labels its acquaintances. *trustCredit* stands for how a relationship partner performed in discoveries in the past. In addition, information on similar requests user evaluated previously in *collaborationRecord*

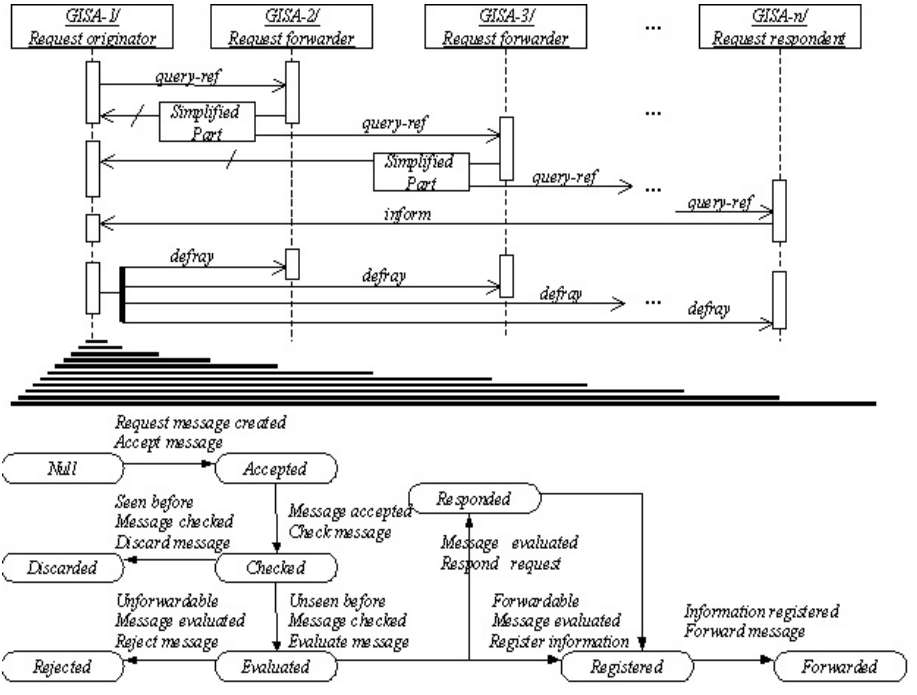


Fig. 3. The interaction protocols of discovery request-respond process using a combination of diagrams, on the bottom of which is the statechart that specifies request-processing behavior for a GISA (request originator)

attribute may help users get desired resources. GISA can use *collaborationRecord* to distinguish which of the two GISAs are more likely to be useful for satisfying user demand. Thus, a decision can be easily made by a GISA regarding which acquaintance is more likely to lead towards desired grid resources. If the condition is satisfied according to *request forwarding model*, the GISA will forward the request message. The GISA will discard the message if it neither responses the request (according to *resource matching model*) nor forwards the request.

As shown in the sequence diagram of Fig. 3, on receiving a request hit, the request originator returns a defray message including a collaboration record and a credit that stands for user evaluation of request hit, according to *service defraying model*. If the relationship of a GISA does not contain the corresponding collaboration record, the collaboration record including initial user evaluation is added to *relationshipDescription* attribute. A credit could be a reward or a penalty, which indicates the degree of its preference of the received request hit. This message is propagated through the same path where the discovery request was originally forwarded. When an intermediate GISA on the path receives message, it adjusts the *trustCredit* value of the relationship that was used to forward the original request. The *trustCredit* value is increased for a reward (i.e., high degree of the request originator's preference), and is decreased for a penalty (i.e.,

low degree of the request originator's preference). Some relevant designs and models such as *resource matching model*, *request forwarding model* and *service defraying model* are described in our paper [10]. Due to focusing on communication schema here, we will not carry on discussing the detailed contents.

4 Simulation Experiment

To examine: (a) whether the design of ENIP is expressive enough to describe complex grid services and (b) whether desired services can emerge from our service-building approach and communication solution, we develop a simulator of discovery service on ENGM platform to check above two aspects.In the simulator, minimum capabilities of ENGM such as *agent communication service*, *agent evolution state management* and *community niche sensing* are implemented.

4.1 Simulator Implementation

(1) Initial Network Topology and Simulation Time We adopt the generation method of network topology proposed in [12]. Suppose that each time there is only one request message sent by a GISA and the same message can be sent only once by the request originator during the entire simulation. Define that a cycle starts from a request message sent by a GISA and ends till all the relevant messages disappear in the system, and 100 cycles is a generation.

(2) Resource distribution. We make a set of common resources (contains 20000 different resource vectors) and a set of new-type resources (contains 2000 different resource vectors that are completely different from common resources). We experiment on two strategies. (a) Balanced distribution strategy: initially each GISA provides 3-5 resource vectors, which are randomly picked out from the common resource set. (b) Unbalanced distribution strategy: a few number of GISAs provide most of the resource vectors from common resource set, while the large number of GISAs share the small part of the resources.

(3) User requests and user evaluation. Requests are initiated at a fixed percentage of randomly selected GISAs and contain resource vectors. The resource attributes of each request have the same weight. Two user request patterns are studied: (a) unbiased user request scenario (requesting all the vectors randomly) and (b) biased user request scenario (using a great probability to request a small part of and special vectors available in the simulation network).

4.2 Simulation Results

We study the effect of different resource distributions and user request patterns on discovery service. A parameter η is introduced to act as the weight given to *trustCredit* and *collaborationRecord*. As shown in Fig. 4, the random forwarding has the lowest efficiency, though it is low-cost (no need to store any discovery information in GISAs). For four experimental environments and different η values, in general, our discovery service can improve its performance adaptively. At the beginning of simulation, relationships of agents are random, and discovery

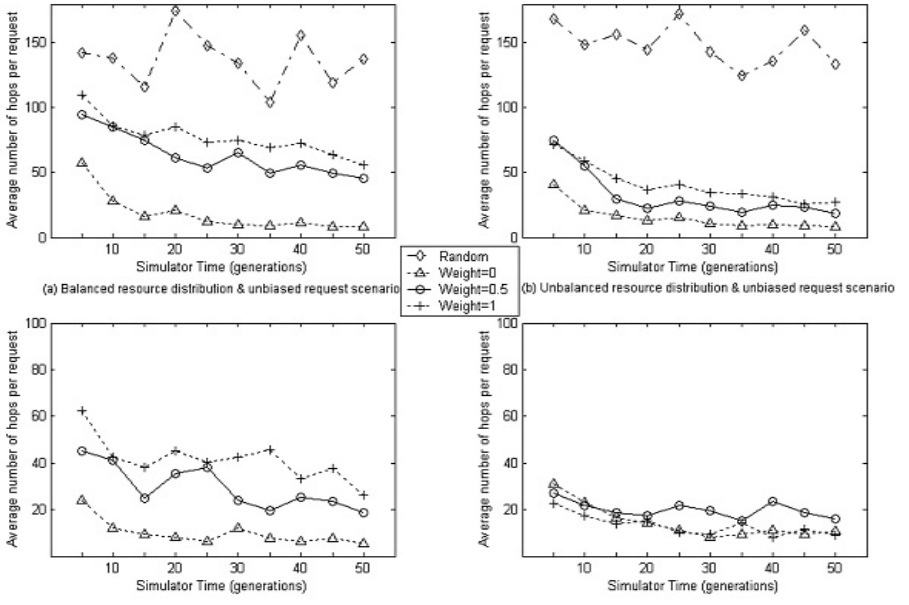


Fig. 4. Average number of hops per request as a function of simulation time in four environments. (a): balanced resource distribution and unbiased user request scenario; (b): unbalanced and unbiased; (c): unbalanced and biased; (d): unbalanced and biased. The number of GISA is 5000.The TTL values are set to 200

performs poorly. Many hops need to be visited to hit the target GISA (hops is a measurement of node amount). As more simulation cycles elapse, GISA gradually obtains many relationships similar to them, leading to improved performance in discovery process. We find that such improvement results from the clustering of request-matched GISAs. With enough hops, GISAs are likely to meet some relative resource clusters during the discovery process and enter the clusters to find the required resources. The clusters have not formed in the process of random forwarding simulation, and the discovery will go aimlessly till it meets the matched GISA. It can be seen that the discovery service can form clusters and improve the discovery performance.

The result has proved that, through exploiting the ENGM platform and its communication implementation, the above process can be well-carried out. The design of ENIP is expressive enough to describe complex discovery service and support a group of agents to emerge discovery service that can adapt to dynamically changing environments.

5 Conclusions

We design and implement a communication schema that couples grid techniques to enable flexibly searching, assembling, organizing and coordinating of agents for

emerging desired grid services. This schema is built on the ENGM platform that would not only be more resistant to failure, but also would provide many useful functions and services for the agents to invoke. To evaluate the communication design and its supported ability of service emergence, we take the service of resource discovery in grid environments as an application case and conduct a series of simulation experiments. The experimental results demonstrate via our communication solution, a service with desired properties can emerge.

Acknowledgments. This work was supported in part by the National Nature Science Foundation of China (No. 60474037 and 60004006) and Specialized Research Fund for the Doctoral Program of Higher Education from Educational Committee of China (No. 20030255009).

References

1. Foster, I., Kesselman, C., and Tuecke, S.: The anatomy of the grid: enabling scalable virtual organizations. Int. J. Supercomputers Applications.**15** (2001) 205–220
2. De Roure, D., Jennings, N. R., and Shadbolt, N. R.: The Evolution of the Grid. In Berman, F., Fox, G. and Hey, A. J. G.(eds.): Grid Computing: Making the Global Infrastructure a Reality , NJ: John Wiley and Sons Ltd. Publishing (2003) 65–100
3. Gao, L., Ding, Y.-S., and Ren, L.-H.: A Novel Ecological Network-based Computation Platform as Grid Middleware System. Int. J. Intell. Sys. **19** (2004) 859–884
4. Karonis, N. T., Toonen, B., and Foster, I., MPICH-G2: A Grid-enabled Implementation of the Message Passing Interface, J. Para. and Dist. Comp.**63** (2003) 551–563
5. Moreau, L., Zaini, N. M., Cruickshank, D., and De Roure, D.: SoFAR: An Agent Framework for Distributed Information Management. In Plekhanova, V. (eds.): Intelligent Agent Software Engineering, PA: Idea Group Publishing, (2003) 49–67
6. Kahn, M. L. and Cicalese, C. D. T.: CoABS Grid Scalability Experiments. Autonomous Agents and Multi-Agent Systems. **7** (2003) 171–178
7. Finin, T., Labrou, Y., and Mayfield, J.: KQML as An Agent Communication Language. in Bradshaw, J.(Eds.): Software Agents,AAAI/MIT Press (1997)
8. Java RMI-IIOP Documentation. http://java.sun.com/j2se/1.3/docs/guide/rmi-iiop
9. Labrou, Y., Finin, T., and Peng, Y.: Agent Communication Languages: The Current Landscape. IEEE Intelligent Systems. **14** (1999) 45–52
10. Parunak, H. V. D. and Odell, J.: Representing Social Structures in UML. in Wooldridge, M., Ciancarini, P., and Weiss, G.(eds.): Agent-Oriented Software Engineering Workshop II. Springer publishing, Berlin (2002) 1–16
11. Gao, L., Ding, Y.-S., and Ren, L.-H.: A social Network-Inspired Resource Discovery Approach in Grid Environments. IEEE Trans. Sys. Man and Cybe. submitted
12. Barabsi, A.-L. and Albert, R.: Emergence of scaling in random networks. Science. **286** (1999) 509–512

A Kernel-Level RTP for Efficient Support of Multimedia Service on Embedded Systems

Dong Guk Sun and Sung Jo Kim

Chung-Ang University, 221 HukSuk-Dong,
DongJak-Ku, Seoul, Korea, 156-756
{dgsun, sjkim}@konan.cse.cau.ac.kr

Abstract. Since the RTP is suitable for real-time data transmission in multimedia services like VoD, AoD, and VoIP, it has been adopted as a real-time transport protocol by RTSP, H.323, and SIP. Even though the RTP protocol stack for embedded systems has been in great need for efficient support of multimedia services, such a stack has not been developed yet. In this paper, we explain *embeddedRTP* which supports the RTP protocol stack at the kernel level so that it is suitable for embedded systems. Since *embeddedRTP* is designed to reside in the UDP module, existing applications which rely on TCP/IP services can be processed the same as before, while applications which rely on the RTP protocol stack can request RTP services through *embeddedRTP*'s API. Our performance test shows that packet-processing speed of *embeddedRTP* is about 7.8 times faster than that of UCL RTP for multimedia streaming services on PDA in spite that its object code size is reduced about by 58% with respect to UCL RTP's.

1 Introduction

Multimedia services on embedded systems can be classified into on-demand multimedia services like VoD and AoD, and Internet telephone service like video conference system and VoIP. H.323[1], SIP[2] and RTSP[3] are representative protocols that support these services. In fact, these protocols utilize RTP[4] for multimedia data transmission. Therefore, embedded systems must support RTP in order to provide multimedia services.

Most recent researches related to RTP have focused on implementation of RTP library for their own application. Typical implementations include RADVISION company's RTP/RTCP protocol stack toolkit[5], Bell Lab's RTPlib[6], common multimedia library[7] of UCL(University College London) and vovida.org's VOCAL(the Vovida Open Communication Application Library)[8]. These libraries can be used along with traditional operating systems such as LINUX or UNIX. Among these implementations, RADVISION Company's RTP/RTCP toolkit[5] is a general library that can be used in an embedded system as well as a large-scale server system. This RTP/RTCP toolkit, however, is not suitable for embedded system; this is because it did not consider

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 79–88, 2005.

typical characteristics of embedded system such as memory shortage. No RTP for embedded systems has been developed so far.

In this paper, we design and implement RTP at the kernel level(called as **embeddedRTP**) to support smooth multimedia service in embedded systems. The primary goals of *embeddedRTP* are to guarantee the small code size, fast packet-processing and high resource utilization. Furthermore, *embeddedRTP* can resolve resource-wasting problem caused by RTP that was implemented redundantly by applications.

This paper is organized as follow. In Section 2, we discuss current RTP's problems and propose a solution to them. Then, Section 3 explains design and implementation of *embeddedRTP*. Section 4 presents performance evaluation of *embeddedRTP* and finally Section 5 concludes our work and discusses future work.

2 Current RTP's Problem

Because RTP is not a mandatory protocol to use Internet, network modules of embedded system do not necessarily include RTP. Therefore, each application requiring RTP in an embedded system should contain its own RTP as shown in Fig.1. In this case, if embedded system must support various kinds of multimedia services such as H.323[1], SIP[2] and RTSP[3], these applications must implement RTP redundantly. Therefore, this method is inappropriate to embedded systems, which do not have sufficient amount of memory.

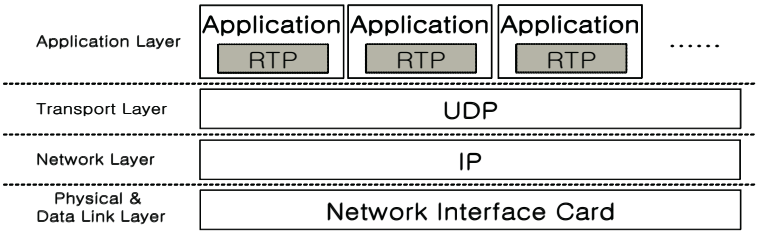


Fig. 1. Protocol stack in case that RTP is included in each application

RTP can be also offered as a library for application layer as shown in Fig.2. This method can resolve RTP's redundancy problem. However, when data is transmitted from UDP to RTP and subsequently from RTP to an application, overhead due to memory copy and context switching may occur. Since various applications must share one library, library compatibility problem may also occur.

If we implement RTP at the kernel level as shown in Fig.3, the protocol redundancy and context switching problems can be resolved. Moreover, we can reduce the code size of applications. To check if a received packet is in fact a RTP packet, port numbers of all UDP packets should be examined. This checking

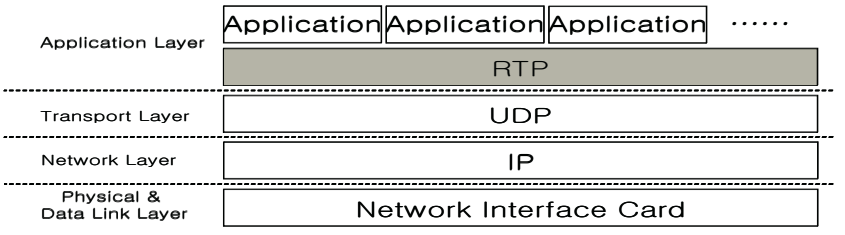


Fig. 2. Protocol stack in case that RTP is supported by library

may cause overhead to other application layer protocols that use UDP. However, such application layer protocols as TFTP, DHCP and ECHO are not used in multimedia communication and do not cause much traffics. Moreover, these do not require high performance. Consequently, we can expect that overhead caused by RTP checking does not have serious effect on performance of these application layer protocols.

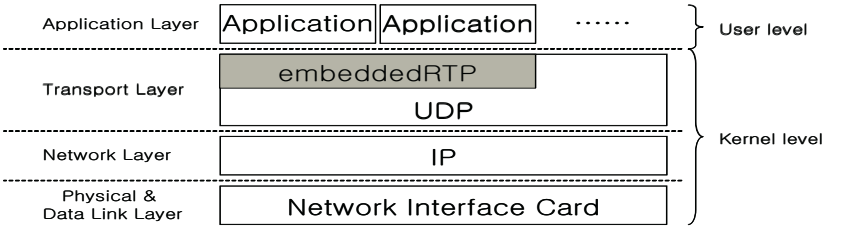


Fig. 3. Protocol stack in case that *embeddedRTP* is implemented at the kernel level

3 Design and Implementation of *embeddedRTP*

3.1 Overall Structure

For efficient memory usage, fast packet processing and removal of redundancy problem, *embeddedRTP* is designed to reside in the UDP module as shown in Fig.3. Fig.4 shows overall structure of *embeddedRTP*.

EmbeddedRTP is composed of API, session checking module, RTP packet-reception module, RTP packet-processing module, RTCP packet-reception module, RTCP packet-transmission module and RTCP packet-processing module. When an application uses TCP or UDP, BSD socket layer can be utilized as before. On the other hand, applications requiring RTP can utilize *embeddedRTP*'s API for RTP services.

3.2 Communication Mechanism

Since RTP is implemented at the kernel level in *embeddedRTP*, communication channel between RTP module of application layer and UDP module of the

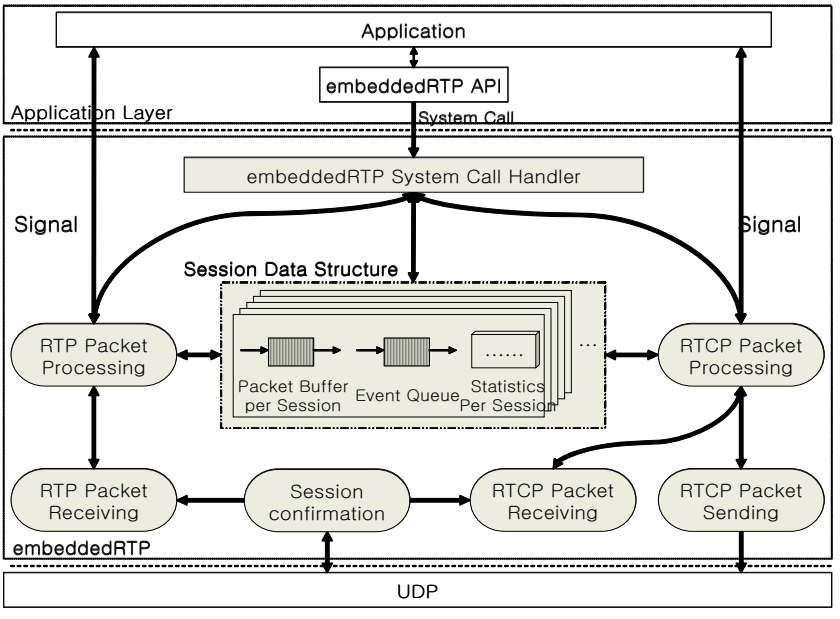


Fig. 4. Overall Structure of *embeddedRTP*

kernel should be changed to communication channel between *embeddedRTP* and UDP module of the kernel. For communication between application and *embeddedRTP*, special mechanisms supported by OS should be provided.

Application Calls *EmbeddedRTP*: In this case, we use system call as a communication channel. System call is the one and only method by which application can use kernel modules in LINUX system. However, the current LINUX kernel does not include RTP-related system calls. Therefore, we register new system calls at the kernel so that application can call *embeddedRTP*.

***EmbeddedRTP* Calls Applications:** In this case, we use signal and event queue as a communication channel. Signal is a message transmission mechanism that is used generally in LINUX, but has following problems. First, since we can only use predefined signals, the signal mechanism cannot be used in data transmission. Second, since the types of signals that can be transmitted are not diverse, this mechanism's usage should be restricted. Finally, there is no mechanism to store signals. Therefore, if another signal arrives before the current signal is handled, the current one will be overridden. Nonetheless, it is enough for notifying the occurrences of events. The signal overriding problem can be resolved by event queue. When an event occurs, it is stored in the event queue and a signal is sent to an application to notify occurrence of an event. Event queue used in *embeddedRTP* manages FIFO(First-In First-Out) queue which is the simplest form of queue. And each queue entry stores the multimedia session ID and event arisen.

3.3 Session Management

Session management is important in *embeddedRTP* because it should be used by all applications which need RTP services. When RTP is implemented at application layer, it is enough for applications themselves to manage their session. However, in *embeddedRTP*, all multimedia sessions must be managed by the kernel.

As shown in Fig.5, multimedia session is managed by one-dimensional array combined with circular queue. Each element of the array contains session information, which includes transmission related data such as the server address and port number, event queue and packet buffer related data structures, and statistical information. In this mechanism, each session can be identified by the index of array; so session information can be accessed quickly. Moreover, *front* and *rear* pointers can prevent the queue from overflowing. However, since there is trade-off between the number of multimedia session and memory usage, it is difficult to decide the proper queue size.

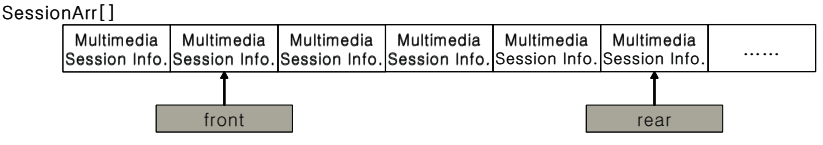


Fig. 5. Data Structure for Session Management

3.4 Session Checking Module

When an RTP packet is received, UDP module does all processing related to UDP packets such as checksum. After that, UDP module stores the packet in socket buffer. At this time, this module checks whether there is a multimedia session associated with the received packet. If a multimedia session exists, RTP/RTCP packet-reception module is called according to the packet type.

3.5 RTP Packet-Reception Module

RTP packet-reception module reads in a packet from socket buffer and sets packet header structure. After that, it checks the version field and length of packets to examine that packets are valid. If valid packets are received, it calls RTP packet-processing module. Received RTP packet is managed by *rtp_packet* structure. This structure consists of *rtp_packet_data* structure that composes packet buffer and *rtp_packet_header* that stores header information.

3.6 RTP Packet-Processing Module

RTP packet-processing module inserts a packet into packet buffer, and updates statistical information. After that, this module informs application that the RTP packet be received.

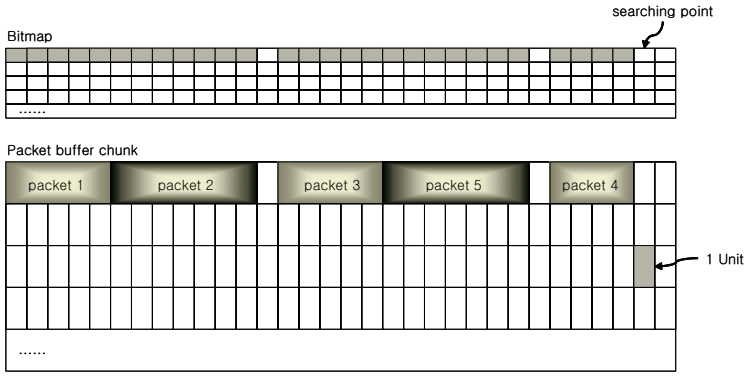


Fig. 6. Packet Buffer

Packet Buffer: Packet buffer stores received RTP packets. As shown in Fig.6, packet buffer is consisted of packet buffer chunk and bitmap. The former consists of units with fixed size, while the latter is a bit-stream that shows whether each unit of packet buffer chunk is available or not.

When a RTP packet is received, RTP packet-processing module checks if it can be stored contiguously from a unit in packet buffer chunk corresponding to searching point[1] of bitmap. If so, it is stored in packet buffer chunk; otherwise, searching starts from the next available searching point repeatedly until enough units can be found. Once it has been stored in packet buffer chunk, bitmap corresponding to the allocated units are set.

This mechanism utilizes memory more efficiently as the size of unit becomes smaller. On the other hand, as it becomes smaller, the size of bitmap becomes larger, resulting in longer searching time for bitmap. Since there is trade-off between memory efficiency and searching time, the unit size should be determined appropriately according to the resources of target systems and the characteristics of applications.

Packet Ordering: Since RTP uses UDP as a transmission protocol, RTP cannot guarantee that packets are received sequentially(See Fig.6). Therefore, in order to guarantee timeliness of multimedia data, *embeddedRTP* must maintain the order of packets that a server transmits. We use circular queue using doubly linked list to maintain the order of packets. Each node of circular queue stores information on a packet stored in packet buffer chunk. *Head* and *tail* pointers of circular queue indicate the first and last packet in the ordered list, respectively.

Each received packet will not be inserted into packet buffer from the unit which is adjacent to the unit pointed by *tail* pointer, unconditionally. Instead, it is inserted in a suitable place of the queue according to its sequence number. If the sequence number is the same as a stored packet, it is dropped because it

[1] The location of bitmap corresponding to a unit of packet buffer chunk from which to search for enough unit(s) to store a packet.

is duplicated one. Consequently, the application can access packets sequentially from the head pointer.

3.7 RTCP

RTCP packet-reception module acts similar to RTP packet-reception module. It reads in a packet from socket buffer and checks the validity of packet. After that, it calls RTCP packet-processing module.

RTCP packet-processing module checks if received packets are in a multimedia session using the SSRC(Synchronization Source) field of RTCP header. If so, it calls relevant processing routine according to the RTCP message type. After RTCP messages are processed, it stores reception event into event queue and sends a signal to the application to inform that RTCP message be received.

When an RTCP packet needs to be transmitted, RTCP packet-transmission module creates and transmits the packet with calculated statistical information. This packet includes various RTCP messages such as RR(Receiver Report), SDES(Source Description) and BYE. Upon transmitting the packet, it resets statistical information and terminates its execution.

4 Performance Evaluations

To evaluate performance of *embeddedRTP*, we measure the packet-processing time, memory requirement, the code size, and session checking overhead. We then compare them with those of UCL RTP library, which has been used in MPEG4IP project[9].

4.1 Packet-Processing Time

Packet-processing time is one of the most important factors to evaluate the performance of protocol stack. Packet-processing time is defined as time between the moment when RTP packets are confirmed in multimedia session and the moment when the application consumes all the packets, after RTP packet-processing has been done in UDP module. While tens of RTP packets are transmitted per second, RTCP packets are transmitted about every 5 seconds. Therefore, we excluded RTCP packet-processing time in our measurement. Table 1 shows the measured packet-processing time.

Table 1. RTP Packet-Processing Time(ms/packet)

	EmbeddedRTP	UCL RTP
Packet-Processing Time	0.289	2.253

As shown in Table 1, *embeddedRTP* is about 7.8 times faster than UCL RTP library. The biggest cause of this difference is that *embeddedRTP* uses statically allocated memory, while UCL RTP library allocates packet storage data

structure dynamically. Another cause is the time when packet is stored. While *embeddedRTP* stores packets at the kernel level, UCL RTP library stores packets at the application layer. Therefore, UCL RTP library can be preempted to other process by kernel scheduler while received RTP packets are being processed. On the other hand, *embeddedRTP* is not preempted because RTP packets are processed in the LINUX kernel.

4.2 Memory Requirement

To compare the memory requirement, four sample animations which have different bit and frame rates are used in measurement. We measure the average payload size of sample animation, the average number of packets being buffered and the number of units that are used in *embeddedRTP*. Since the average payload size of animations used in the measurement tends to be multiple of a roughly 200 bytes, we adopt 200 bytes as the size of unit. The media buffering time when to start animation playback greatly affect the amount of memory to be used. While buffering time is about 5 - 10 seconds for playing sample animations, in general, we allow only 2 seconds as buffering time to reduce memory requirement without affecting animation playback.

Table 2. Memory requirement of UCL RTP and *embeddedRTP*

Sample	Payload (bytes)	Buffered Packets	Units	Memory Requirement(bytes)		Bit Rates	Frame Rates
				UCL RTP	*EmbeddedRTP*		
1	798	27	4.683	40,087	25,031	125	15
2	849	57	4.935	85,580	56,315	277	30
3	389	87	2.778	131,180	47,469	101	30
4	280	58	2.099	86,594	24,231	68	30

As shown in Table 2, *embeddedRTP*'s memory requirement is about 28% - 65% than UCL RTP's. Also, the amount of memory required by sample animations varies according to their characteristics. Sample 2 and 4 have the similar number of buffered packets, but their payload sizes differ as much as 570 bytes. Since UCL RTP library uses the equal sized data structure regardless of the payload size, memory requirement are similar. On the other hand, in *embeddedRTP*, sample 2 and 4 require about 65% and about 28% of memory required by UCL RTP library, respectively.

The payload size of sample 1 is greater than that of sample 4, while the average number of sample 1's buffered packets is much less than sample 4's. Considering these two samples, *embeddedRTP*'s memory requirements are similar in both samples. However, in UCL RTP library, memory requirements increase proportionally by the average number of buffered packets since the library allocates pre-determined amount of memory per packet. Since sample 3 has smaller payload size comparing with other samples, but has more average number of buffered packets, sample 3 shows the largest difference in memory requirement between *embeddedRTP* and UCL RTP.

4.3 Code Size

Because embedded systems have much smaller physical memory than desktop PC or server system, the code sizes of software is very important in embedded systems. Table 3 shows the code size of *embeddedRTP* and that of UCL RTP library. The former is determined by the total size of *embeddedRTP* modules that are implemented at the kernel level and *embeddedRTP*'s API. The latter is determined only by modules relevant to packet reception. As shown in Table 3, *embeddedRTP*'s code size is reduced to 42% of UCL RTP's.

Table 3. Code Size(bytes)

EmbeddedRTP		UCL RTP Library
EmbeddedRTP Modules	*EmbeddedRTP*'s API	
25,080	804	61,132
25,884		

4.4 Session Checking Overhead

In *embeddedRTP*, all packets received by network system should be checked whether they are in fact RTP packets. Therefore, we measure session checking overhead for RTP packets which are not in a multimedia session and other UDP packets which are not a RTP or RTCP packet. In the worst case, it takes 0.014 ms per packet. This time is relatively short comparing with packet-processing time of *embeddedRTP*(0.289 ms). Note that such application protocols that use UDP but are not be used in multimedia services as TFTP or DHCP, neither generate much traffics nor require high performance. Consequently, the overhead caused by session checking does not have serious effects on performance of these application layer protocols.

5 Conclusions and Future Work

In this paper, we explained *embeddedRTP* which supports the RTP protocol stack at the kernel level so that it is suitable for embedded systems. Since *embeddedRTP* is designed to reside in the UDP module, existing applications which rely on TCP/IP services can be processed the same as before, while applications which rely on the RTP protocol stack can request RTP services through *embeddedRTP*'s API. *EmbeddedRTP* stores received RTP packets into per session packet buffer, using the packet's port number and multimedia session information. Communications between applications and *embeddedRTP* is performed through system calls and signal mechanisms. Additionally, *embeddedRTP*'s API makes it possible to develop applications more conveniently. Our performance test shows that packet-processing speed of *embeddedRTP* is about 7.8 times faster than that of UCL RTP for multimedia streaming services on PDA in spite that its object code size is reduced about by 58% with respect to UCL RTP's.

To improve the result of this paper, the followings should be investigated further. Static packet buffer used in *emdbeddedRTP* can reduce packet-processing time but restricts the extensibility. In order to overcome this shortcoming, extensible packet buffer using overflow buffer has to be implemented. In addition, research on how to determine the size of unit dynamically is needed to obtain optimal performance. Furthermore, we need to investigate a mechanism to reduce the number of memory copies between protocols stack and the applications.

References

1. ITU-T Recommendation H.323: Packet based multimedia communications systems. Feb. (1998).
2. J.Rosenberg, et al. : SIP: Session Initiation Protocol , RFC 3261, Jun. (2002).
3. H.Schulzrinne, et al.: Real Time Streaming Protocol (RTSP)., RFC 2326, Apr. (1998).
4. H.Schulzrinne, et al.: RTP: A Transport Protocol for Real-Time Applications. , RFC 1889, Jan. (1996).
5. RADVISION : RTP/RTCP Toolkit., `http://www.radvision.com/TBU/Products/RTP-RTCP+Toolkit/default.htm`.
6. Lucent Labs : Lucent Technologies Software distribution. , `http://www.bell-labs.com/topic/swdist`.
7. University College London, "UCL Common Multimedia Library", `http://www-mice.cs.ucl.ac.uk/multimedia/software/common/index.html`.
8. VOVIDA.org : Vovida.org., `http: // www.vovida.org/protocols/downloads/rtp` .
9. MPEG4IP : MPEG4IP - Open Streaming Video and Audio., `http://www.mpeg4ip.net`

Group-Based Scheduling Scheme for Result Checking in Global Computing Systems*

HongSoo Kim[1], SungJin Choi[1], MaengSoon Baik[1], KwonWoo Yang[2], HeonChang Yu[3], and Chong-Sun Hwang[1]

[1] Dept. of Computer Science & Engineering, Korea University,
1, 5-Ga, Anam-Dong, SungBuk-Gu, Seoul, 136-701, Rep. Korea
{hera, lotieye, msbak, hwang}@disys.korea.ac.kr
[2] Gongju National University of Education,
376, BongHwang-Dong, GongJu city, 314-711, Rep. Korea
kwyang@gongju-e.ac.kr
[3] Dept. of Computer Education, Korea University,
1, 5-Ga, Anam-Dong, SungBuk-Gu, Seoul, 136-701, Rep. Korea
yuhc@comedu.korea.ac.kr

Abstract. This paper considers the problem of correctness to fault-tolerance in global computing systems. Global computing system has been shown to be exposed to intentional attacks, where authentication is not relevant, network security techniques are insufficient. To guarantee correctness for computation, fault-tolerance schemes have been used to majority voting and spot-checking but these schemes intend to high computation delay because are not applied scheduling scheme for result checking. In this paper, we propose a new technique called GBSS(Group-Based Scheduling Scheme) which guarantee correctness and reduce computation delay by using fault-tolerant scheduling scheme to the result checking. Additionally, simulation results show increased usable rate of the CPU and increased performance of the system.

1 Introduction

Global computing systems are computing paradigm to run high-throughput computing applications by using idle time of internet connected computers[12]. One of main characteristics of the systems is that computation nodes are freely leaved or joined according to their volatile property. Moreover, there exists different administrator for each nodes. These projects have ventured to study and develop global computing systems such as SETI@home [2], Korea@home [14], Distributed.net [9], Entropia [8], Bayanihan [3], XtremWeb [4][12], Charlotte [6], Cilk [5], GUCHA [13].

Global computing systems assume relatively unreliable computing resources because of the computation interference or submitting bad result of malicious

* This work was supported by grant No. G- 04-GS-01-03R-7 from the Korea Institute of Science and Technology Information.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 89–98, 2005.

workers. If malicious workers are submitting bad result, then all results could be canceled. Practically, in the SETI@home, malicious workers returned bad result as change original code[2]. Therefore, global computing systems have to consider a result checking scheme for the computed result.

In previous work, a result checking scheme for the computed result presented majority voting and spot-checking. A majority voting[7] scheme adapt result if results are same value compare to least three result. This scheme has an expected redundancy of 2k+1, so this scheme is becoming insufficient because of a waste of the resource. In spot-checking scheme[3], the master node does not redo all the work objects two or more times, but instead randomly gives a worker a spotter work object whose correct result is already known or will be known by checking it in some manner afterwards. These previous works are not applied a scheduling scheme, so they have been the high computation delay and the high error rate.

In this paper, we propose fault-tolerant scheduling scheme through organized group as credibility of worker. First, we calculate credibility of each workers by spot-checking scheme. Second, we organize group through credibility-based group organization scheme(CBGOS). Finally, we propose group-based scheduling scheme. The GBSS guarantee correctness and reduce computation delay by using fault-tolerant scheduling scheme.

The rest of paper is structured as follows. In section 2, we introduce executing mechanism and assumption for global computing system model. In section 3, this paper describe group-based scheduling scheme using the credibility of each workers. In section 4, we describe implementation and performance evaluation. Finally, in section 5, we discuss conclusions and future work.

2 Global Computing System Model and Assumptions

In this paper, we assume a work pool-based master-worker model. This model is used in practically all internet-based distributed computing systems. As show in **fig. 1**, a computation is divided into sequence of batches, each of which consists of many mutually independent work objects. This work objects are located into work pool by task allocation server(TAS). Work objects are assigned to each workers by group-based scheduling scheme, and then they are executed computation in worker, after then scheduler return results to the TAS. Each workers request new work object in work-pool and then they complete a batch as execute a work objects by work stealing[5]. Also, we assume single-instruction multiple-data(SIMD) model which execute single code through each different data.

For some applications, the acceptable error rate can be relatively high, about 1% or more. These include applications such as image or video rendering. But also applications as climate forecast, acceptable error rate can be relatively low, near 0%. Therefore, global computing systems need the result checking mechanism to correctness for result. In this paper, we assume malicious failure model as follows. Malicious failure means worker return malicious bad results. So, we focus this problem that worker return bad results by malicious failure. Therefore, we propose schemes to solve this problem.

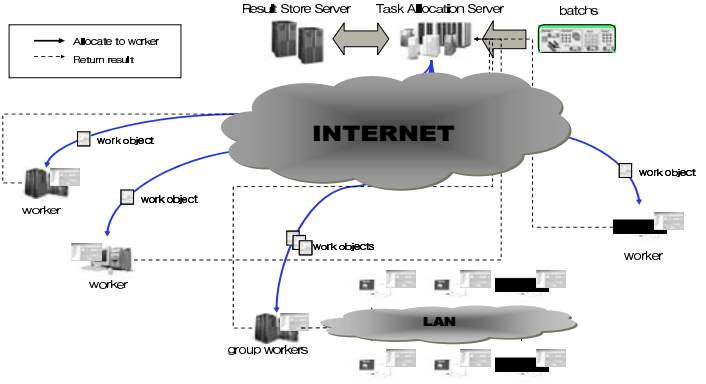


Fig. 1. Global Computing System Model

3 Group-Based Scheduling Scheme for Result Checking

We describe the CBGOS and the GBSS proposed in this paper. Although work has been made on the credibility-based fault tolerance scheme using spot checking and majority voting proposed in [3], this scheme is not applied fault-tolerant scheduling scheme. In our work, we propose scheduling scheme as credibility-based grouping of each workers. Therefore, we can be guaranteed correctness for computed result and reduce the computation delay by result checking scheme.

3.1 Overview

The GBSS is executed by grouping using the credibility threshold as **fig. 2**. We apply GBSS as credibility-based grouping to guarantee correctness of executed result by workers. First of all, this scheme have to calculated credibility C_v of a worker v_i by majority voting and spot-checking before assign work object to the workers. If we only accept a result for a work object when the probability of that

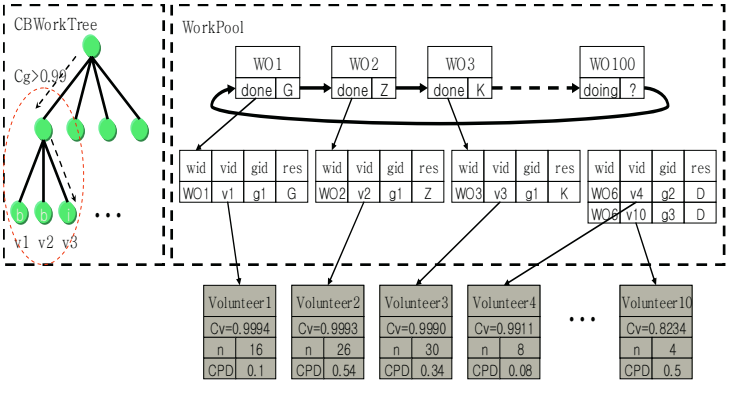


Fig. 2. Work Pool and Work Tree to Group-based Scheduling

result being correct is at least threshold θ, then the probability of accepting a correct result would be at least θ. θ has different value as each applications and calculated as $\theta = 1$ - ε by acceptable error rate ε.

In the global computing system, the executing applications have each different error rate *f*. An executed result by worker could be accepted if results have verify the reliability by the majority voting. First of all, when scheduler assigns work object, it execute concurrently assignment for point of group by credibility of group. As error rate, it determine a number of redundancy $k/(1-f)$ and assign to itself or to upper group when execute k+1 majority voting scheme. k+1 majority voting adapt more than half in k+1, this is to reduce redundancy. If it lead to bad result, executed result by majority voting scheme as a redundancy, then we have one more redundancy. In the next section, we guarantee more higher performance than previous schemes by group-based scheduling scheme.

3.2 Credibility

We calculate credibility of workers to guarantee correctness for executed results by workers. This paper defines credibility of a worker as follows.

Definition 1. *Credibility(*C*). The credibility is a determining foundation of correctness as executed computation result by worker. This use error rate* f*, participated degree in computation and the number of spot-checks passed by a worker,* n*, to estimate likely a worker is to give a good result.*

$$C_{v_i} = 1 - \frac{f}{n} \cdot CPD(v_i)(n > 0) \tag{1}$$

$$C_{v_i} = 1 - f \cdot CPD(v_i)(n = 0) \tag{2}$$

Definition 2. *Computation Participation Degree(*CPD*). The CPD define that worker is the rate of how long does it participate at the computation, so we calculate really participated time at the computation. The participation degree at computation $CPD_k(v_i)$ for a worker v_i calculate as follows.*

$$CPD_k(v_i) = 1 - \frac{CJT_k(v_i)}{CJT_k(v_i) + CLT_k(v_i)} \tag{3}$$

In **equation 1**, C_{v_i} is the credibility of the worker v_i and n is the number of spot-checks returned by a worker and f is the probability that a worker chosen at random would be bad. If n is 0, then it would be calculated by **equation 2**. And participation degree of computation $CPD_k(v_i)$ is calculated by the rate of really participated term for really participated term of worker and leaved term of worker. In **equation 3**, $CJT_k(v_i)$ is participated term of v_i in computation and $CLT_k(v_i)$ is leaved term of v_i in computation. Therefore, if leaved term in computation is more than really participated term in computation, then credibility of v_i is relatively the less. Also, the higher participation degree of computation have the higher credibility.

```
var
    n := theNumberOfWorkers; //the number of workers
    θ := credibilityThreshold; //credibility threshold
    v_i := worker; //ith worker
    g_id := groupID; //ID of group
    C_{v_i} := credibilityOfWorker; //credibility of worker v_i

OrganizeGroup()
    for i := 0 to n do
        C_{v_i} := ClassifyWorkersByCredibility();
        g_id := organizeCBWT(C_{v_i});
    endfor
```

Fig. 3. Group Organization Algorithm

3.3 Credibility-Based Group Organization Scheme

The CBGOS organize group by credibility of a worker. Before assign a work object, it complete work tree as group organization algorithm with information of the workers as calculated credibility value C_{v_i} by spot-checking. If the credibility of group is more higher than the least threshold when it organize group, then it is place at the highest level. And, workers update work tree by CBGOS whenever they return result. As **fig. 2**, the credibility-based work tree determine that it assign to which worker selected next work object in cycle linked list of work pool. A state of worker in the work tree have idle(i) which is computation available state, busy(b) which is computation unavailable state and die(d) which is stop failure state.

After organized credibility-based work group when assign computation to each group, scheduler assign identifier to know which work object assigned to which worker and organized in which group. The elements of work object is as follows.

$$WO(v_{id}, w_{id}, g_{id}) \tag{4}$$

A v_{id} is the identifier of a worker and a w_{id} is the identifier of a work object and a g_{id} is the identifier of a group. Also, these identifiers are divided into CBWT.

3.4 Group-Based Scheduling Scheme

A work scheduling of this paper apply basically eager scheduling scheme with added GBSS. We firstly execute spot-checking in order to calculate credibility of a worker before allocate a work object. After then, we organize credibility-based group. As **fig. 4**, an allocation server of task allocates a task W_i to a worker through GBSS. If allocated worker's credibility is more less than credibility threshold, then it execute majority voting. Executed result R_i by two workers confirm result and return to task allocation server if scheduler give good result through CBGOS. In this time, GBSS allocate a same work object to a worker of group to more upper level through CBWT. When it execute scheduling, we have to consider state of work object as follows. A work object of work pool have three

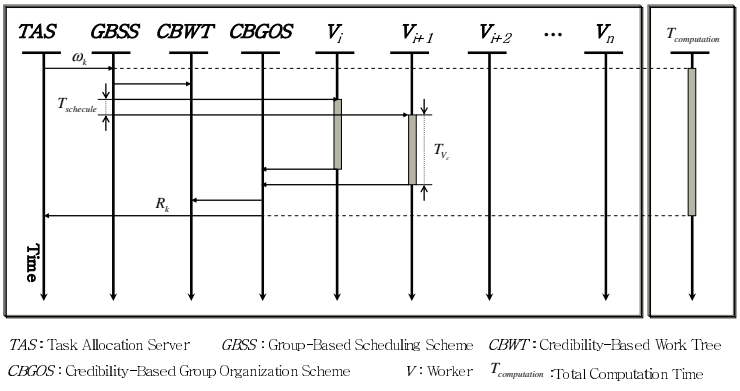


Fig. 4. Scheduling Sequence Diagram through GBSS

```
var
   wo := WorkObject;
   wp := WorkPool;
   wt := CBWT;
   w_id := IDofWorkObject;
   v_id := IDofWorker;
   g_id := IDofGroup;

while wp.nextWorkPoolObject() != null &
     wp.nextWorkObjectState == "undone" do
    wo := wp.nextWorkObject();
    if initial computation then
       allocateWorkObject(wo);
    else
       for wt.everyCredibilityGroup do
          if( wt.g_id > wo.g_id )
             v_id := wt.searchWorker("idle");
          endif
       endfor
       wo.g_id := wt.getGroupID(v_id);
       allocateWorkObject(wo);
    endif
endwhile
```

Fig. 5. GBSS Algorithm

states such as "done", "undone" and "doing". The "done" state means return correct result after executed by worker. The "undone" state means worked state or work state before execution. Also, the "doing" state means working state by a worker. In **fig. 5**, we present a GBSS algorithm.

4 Implementation and Performance Evaluation

We measure performance of added component to global computing system, "Korea@home"[14]. This system is able to data centric distributed computing system because of consisted of huge amount of data in set of single instruction.

Type	The Number of Workers	Rate
Xeon	50	1.20%
P4 2.0G over	1,145	27.37%
P4 2.0G under	602	14.39%
P3	1,067	25.48%
P2 under	211	5.05%
Intel Celeron	148	3.54%
Intel Mobile	73	1.75%
AMD Athlon XP	296	7.07%
AMD Athlon	138	3.30%
AMD Duron	38	0.91%
AMD	32	0.76%
AMD Mobile	4	0.10%
Etc	2	0.05%
Not Aware	378	9.03%
Total	4,184	100%

Fig. 6. Organizing Distribution as CPU type

Execution between each data have to independent relation to mutual exclusive execution.

In Korea@home, applications such as "fundamental research to protein folding" and "fundamental research to discover a new medicine" are executed during from November 2003 to February 2004 by participation of 1,845 users and 4,184 workers. In **fig. 6**, workers of Korea@home have various CPU types. We present statistical information of workers's CPU type as follows.

– **Organized Distribution as CPU type**. The workers of 27% for all have Pentium4 2.0Ghz and workers of about 25% have performance of Pentium3. The Pentium4 and Xeon type have relatively high performance as about 43%. Also, AMD CPU has about 12%.

4.1 Performance Evaluation of Group-Based Scheduling Scheme

Group-based scheduling scheme for result checking proposed in this paper presents more low computation delay than previous scheme. First, this scheme to result checking guarantee the reliability for a result as consist of group by the credibility of a worker and then assigns a work object to the organized group. Therefore, this scheme shows more low computation delay than previous scheme. Second, We measured computation delay. As shows in **fig. 4**, we can be reduced computation delay as execute fault-tolerant scheduling in the same time.

– **Guarantee of correctness for result**. We could verify to guarantee correctness for result by our proposed GBSS. First of all, we measure credibility of worker and then apply scheduling scheme by measured credibility of workers. Therefore, we describe equation of proposed scheme in this paper as follows.

$$R_T = \prod_{i=1}^{n} R_{g_i} \tag{5}$$

$$Also, R_g = 1 - \prod_{j=1}^{m}(1 - R_i) \tag{6}$$

In **equation 5**, R_{g_i} denoted reliability for group, R_T denoted reliability for all system. Also, in **equation 6**, R_i denoted reliability for a worker. In previous scheme, it is dependent of reliability of each workers to guarantee correctness, but as **equation 5**, we scheme can shows reliability for all computation by reliability of group R_g. Therefore, we does show reliability for all computation of group by reliability of group verify higher than previous scheme. So, probability to voting normal workers is presented by the "***Bernoulli Trials***". Therefore, we verify group-based scheduling scheme have voting probability worker of more higher credibility than previous scheme.

$$P_x = \binom{n}{s} | \frac{F}{N - \sum_{j=k}^{n} N_G} |^s | 1 - \frac{F}{N - \sum_{j=k}^{n} N_G} |^{N-S} \tag{7}$$

In **equation 7**, P_x is probability to normal worker and n denoted n trials, sequence of independent repetitions, and s denoted s successes and F denoted the number of normal worker and N denoted the number of all workers. In **equation 7**, N_{G_j} denoted the number of workers of jth group and F is probability of voted normal worker in rest workers of N minus other groups. We are compared with two equation. In previous scheme, it present the number of assigned worker to prior to executed worker among total workers, but in proposed scheme, it has probability the higher correctness because of voted worker in more upper group than itself by GBSS.

– **Measurement of CPU usable rate applied group-based scheduling scheme.** We measure CPU usable rate applied group-based scheduling scheme. **fig. 7** shows experimental results we plot ranking of CPU usable rate of participated worker's PCs during a month period, from SEP. 2004 to OCT. 2004 at Korea@home.

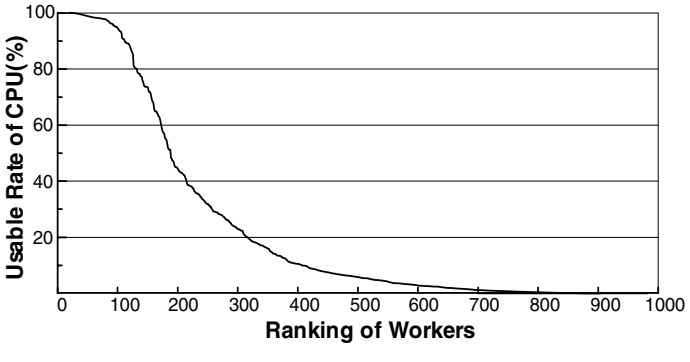


Fig. 7. CPU Usable Rate of Workers

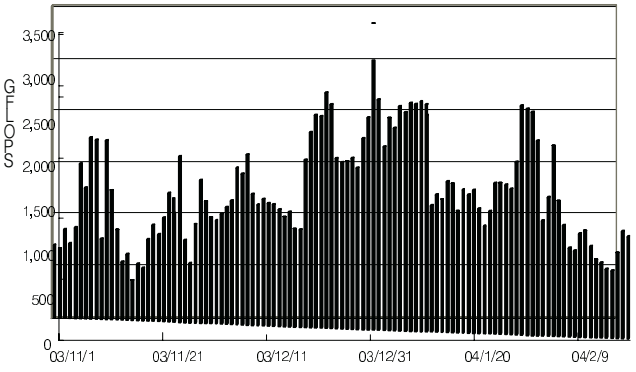


Fig. 8. Performance of Korea@Home measured during about three month

The usable rate of CPU means rate of turnaround time, computation time of all at worker's PC. In here, turnaround time of worker's PC presents normally executing time of worker's PC. First of all, we are shown as follows when we are not apply GBSS. Generally, while measured term of CPU usable rate, worker's PCs of 958 executed task from server and shown CPU usable rate of average 9.68%. But, when we are apply group-based scheduling scheme, worker's PCs of 1598 executed task from server and shown CPU usable rate of average 38.19%. We can know discarded results in previous work, and also we can know high usable rate by reduction of discarded results by GBSS. Therefore, we can getting high performance by GBSS.

– **Performance Evaluation by applied GBSS**. We measured performance by apply GBSS in Korea@home. **fig. 8** shows experimental results it shows measured result during about three month period.

The workers participated average 412(max 942) while a day. We can get 3,185 GFlops more 13% than before and performance of average 1,457 GFlops. In this way, GBSS can be guaranteed the higher correctness.

5 Conclusion and Future Work

In this paper, we proposed GBSS by credibility-based group organization in global computing system. This scheme apply scheduling scheme through organization of group as participated degree in computation and credibility of worker. We can be guaranteed correctness for executed result by workers and also reduced computation delay as result checking.

In near future, we should research for advanced scheduling scheme and dependency model between work objects to implement in Korea@home. Also, we should research for the least turnaround time to the fast response time by scheduling scheme using availability of workers.

References

1. M. O. Neary, P. Cappello. "Advanced Eager Scheduling for Java Based Adaptively Parallel Computing," JGI'02, November 3-5, 2002.
2. D. Molnar. "The SETI@home Problem".
3. L. Sarmenta. "Sabotage-Tolerance Mechanism for Volunteer Computing Systems," FGCS, 18(4). 2002.
4. C. Germain, N. Playez. "Result Checking in Global Computing Systems," ICS'03, June 23-26, 2003.
5. R. D. Blumofe, C. F. Joerg, B. C. Kuszmaul, C. E. Leiserson, K. H. Randall, and Y. Zhou. "Cilk: An Efficient Multithreaded Runtime System," In 5th ACM SIAPLAN Symposium on Principles and Practice of Parallel programming (PPOPP '95), pages 207-216, Santa Barbara, CA, July 1995.
6. A. Baratloo, M. Karaul, Z. Kedem, and P. Wyckoff. "Charlotte: Metacomputing on the Web," In Proceeding 9th International Conference on Parallel and Distributed Computing Systems, 1996.
7. L. LAMPORT, R. SHOSTAK, M. PEASE. "The Byzantine Generals Problem," ACM Transactions on Programming Languages and Systems, Vol. 4, No. 3, July 1982.
8. Entropia home page. http://www.entropia.com.
9. A. L. Beberg, J. Lawson, D. McNett, distributed.net home page. http://www.distributed.net.
10. M. O. Neary, A. Phipps, S. Richman. "Javelin 2.0: Java-Based Parallel Computing on the Internet". In Proceedings of Euro-Par 2000. Munich, GERMANY, August 28 - September 1, 2000.
11. N. Camiel, S. London, N. Nisan, and O. Regev. "The POPCORN Project: Distributed Computation over the Internet in Java," In 6th International World Wide Web Conference, Apr. 1997.
12. C. Germain et al, "Global Computing Systems," In Proceeding of SciCom01. LNCS 2179. Springer, 2001.
13. L. F. Lau, A. L. Ananda, G. Tan, W. F. Wong, "GUCHA: Internet-based Parallel Computing using Java," ICA3PP, pp. 397-408. December 2000.
14. Korea@home home page. http://www.koreaathome.com.

Service Discovery Supporting Open Scalability Using FIPA-Compliant Agent Platform for Ubiquitous Networks*

Kee-Hyun Choi, Ho-Jin Shin, and Dong-Ryeol Shin

School of Information and Communication Engineering,
Sungkyunkwan University, Korea
300 Cheoncheon-dong, Jangan-gu, Suwon, 440-746, Korea
{gyunee, hjshin, drshin}@ece.skku.ac.kr

Abstract. Service discovery protocol is the main element that determines the efficiency in middleware platform of ubiquitous networks. Recently, a large number of middlewares supporting scalability which focus on the distributed computing and data sharing among tremendous peers has been proposed. However, due to the distributed nature of ad-hoc networks, peers may not be able to find other peers and corresponding resources persistently. Furthermore, current service discovery engines do not provide *open scalability* that can make them to interoperate with each other. In this paper, we propose a simple mechanism, which provides cross-platform interoperability through directory federation combined with DHT mechanism to support *open scalability* and lightweight service discovery in ad-hoc network based on FIPA compatibility agent framework.

1 Introduction

These days, the computing environment is becoming more and more pervasive, ubiquitous and mobile. As the number of Internet services grows, it becomes increasingly important for network users to be able to locate and utilize those services which are of interest to them on the Internet. As a result, service discovery plays a key role in highly dynamic networks, for example, in ad-hoc networks. Especially, with the increasingly wide utilization of mobile devices such as PDAs, the dynamic discovery of services in a visited foreign network and automatic system configuration is becoming very important.

Much of the previous work that has been done on service discovery is based on centralized registries such as UDDI [2]. However, this approach suffers from the traditional problems of centralized systems, namely performance bottlenecks and single points of failure. In addition, centralized registries may be more vulnerable to denial of service attacks. Moreover, storing vast numbers of advertisements on centralized registries prevents their timely update as changes are sometimes made in the availability and capabilities of the providers involved.

* This work was supported by a grant from the CUCN, Korea.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 99 – 108, 2005.

In contrast to the traditional 'client/server' architecture which is server-centric, the "peer-to-peer" (P2P) architecture allows every peer in the network to act as both a client and a server, i.e. every peer has equivalent capabilities and responsibilities. In such an environment, a peer can be any network-aware device such as a cell phone, PDA, PC or anything else you can imagine that passes information in and out. Alternatively, a peer might be an application distributed over several machines. The entire premise of the Internet is centered on the sharing of information and services. The focus of P2P networks is on how a peer provides services to other peers and how it discovers the services available from other peers, i.e. the methods of service discovery that are based on different P2P applications.

A great deal of papers dealing with service discovery has recently been published. Although many efficient and fast searching schemes for service discovery are proposed in these papers, they do not consider the problem of inter-operability among discovery protocols. In this paper, we focus on service discovery in a large scale network. In such a pervasive environment, heterogeneous middlewares exist, which should be able to inter-operate with each other. Because most of the proposed schemes provide scalability only within their own domain, they cannot communicate with peers in different domains and, thus, cannot find each other. For this reason, in this study, we developed a new agent platform in order to provide open scalability. Open scalability refers to expanded scalability, namely the capacity to communicate with peers in different domains without any modifications being required to the existing systems (protocols/middlewares). We define scalability according to the system's interoperating capability.

Closed Scalability: Existing systems (e.g. Jini, UPnP[5], JXTA) may support scalability, but only in their own domains. In a ubiquitous environment, however, users with wearable/portable devices (e.g., notebooks, PDAs, cellular phones) can move among domains. Such systems, which only support closed scalability, cannot provide users with inter-communication capability.

Open Scalability: As mentioned above, open scalability allows users to communicate with each other without any modifications being required to the existing systems. For example, if user A in Jini[3] and user B in JXTA[4] want to communicate with each other, these two users would need to modify their systems. In contrast, a system supporting open scalability would have the inherent capacity to provide inter-operability automatically.

Some approaches (e.g. NaradaBrokering[6], BASE[7], and ReMMoC[8]) which support open scalability have already been published. The ReMMoC project, currently being carried out at Lancaster University in collaboration with Lucent Technologies, is examining the use of reflection and component technology, in order to overcome the problems of heterogeneous middleware technology in the mobile environment. Narada is an event brokering system designed to run on a large network consisting of cooperating broker nodes. Communication within Narada is asynchronous and the system can be used to support different interactions, by encapsulating them in specialized events. The BASE is the uniform abstraction of services as well as device capabilities via proxies as the application programming interface. Consequently, the middleware delivers requests to either device services in the middleware or transport protocols. In our proposed architecture, we use a FIPA-compatible Agent Platform to provide open scalability in a ubiquitous environment. A more detailed explanation will be given in Section 3.

In order to satisfy interoperability and open scalability between heterogeneous service platforms in an ad-hoc environment, we designed a modified algorithm based on a distributed hash table and implemented an extended agent platform based on a FIPA compatible agent platform and a lightweight service discovery protocol suitable for ad-hoc networks. The main contribution of this paper is the association of the DHT algorithm with the DF of the agent platform, in order to guarantee open scalability.

The remainder of the paper is organized as follows. Section 2 describes the basic service discovery protocols and FIAP agent platform. Sections 3 and 4 provide the details of the proposed agent platform and lightweight service discovery protocol for use in ad-hoc networks. Finally, Section 5 gives our conclusions and the scope of future works.

2 Related Works

2.1 Agent Platform

Many service discovery protocols have been proposed for wired networks (e.g. Jini, SLP[1], UPNP), all of which are based on the assumption of there being a common data model for the services involved. Agent platforms have also been designed to facilitate flexible service/agent discovery within an agent community. In all of these platforms, an agent uses a Directory Facilitator (DF) and an Agent Management System (AMS) to register/manage the services and agents [9]. AMS manages the life cycle of the agents, the local resources and the communication channels. It also acts as a "white pages", which allows agents to locate each other by their unique names. The DF provides a "yellow pages" service that identifies which agent provides what services. The reference FIPA model described in the FIPA Agent Management Specification also introduces Agent Communication Channels (ACC); ACC manages the interchange of messages between agents on the same or different platforms.

Recently, FIPA published the FIPA Agent Discovery Service Specification (FIPA ADS), which describes the interoperability of agents in Ubiquitous Computing environments with ad-hoc communication. The DF federation mechanism is not suitable for flexible service discovery with mobile ad-hoc communication, because it takes a long time to find a remote DF and to search for the services registered there. For these reasons, FIPA ADS uses the existing Discovery Middleware (DM) (e.g. Jini, JXTA and Bluetooth) for service discovery in the ad-hoc environment. The FIPA JXTA Discovery Middleware Specification deals with the use of JXTA as a discovery middleware (DM) for the Agent Discovery Service (ADS). Although these specifications provide solutions to the problem of the agents' interoperability, it is difficult to implement all of the different discovery middlewares used in the various agent platforms in order to provide sufficient scalability. This means that FIPA AP with DM provides scalability only in a few domains, even if we attach a new DM whenever necessary. In this paper, we propose a new mechanism for DF federation using a distributed hash table (DHT). DF federation with DHT provides open scalability to FIPA AP without the assistance of other DMs. This implies that we can find services/agents in different domains without the need for any other discovery middlewares. In addition, we introduce a lightweight discovery middleware for mobile devices in ad-hoc networks,

since the currently available discovery middlewares are too heavy to be used to find services in an ad-hoc environment. We describe the design method in more detail in Sections 3 and 4.

2.2 Service Discovery Overview

In this section, we discuss several of the different service discovery protocols and middlewares, which have been proposed in order to facilitate dynamic cooperation among devices/services with minimal administration and human intervention. In order to be able to support the impromptu community, these protocols and middlewares should have the means to announce their presence to the network, to discover the services that are available in their neighborhood, and to gain access to these services. Basically, SLP, Jini, UPnP and Salutation[10] all address these concerns, but in different ways. A direct comparison between these different schemes is attempted here, since it is helpful to understand each of them, although they each put different weights on the above mentioned functionalities.

SLP (Service Location Protocol): The SLP is an IETF standard which allows IP network-based applications to automatically discover the location of a required service. The SLP defines three agents, viz. the User Agent, which performs service discovery on behalf of client software, the Service Agent, which advertises the location and attributes on behalf of services, and the Directory Agent, which stores information about the services announced in the network.

Jini: Jini is a technology developed by Sun Microsystems. Jini's goal is to enable truly distributed computing by representing hardware/software as Java objects. The service discovery mechanism in Jini is similar to the Directory Agent in SLP, and is called the Jini Lookup Service.

UPnP: UPnP uses the Simple Service Discovery Protocol (SSDP) as its discovery protocol. SSDP can work with or without the central directory service, called the Service Directory. When a service wants to join the network, it first sends an announcement message by multicasting, in order to notify the other devices of its presence. Thus, all of the other devices can see this message, and the Service Directory registers the announcement.

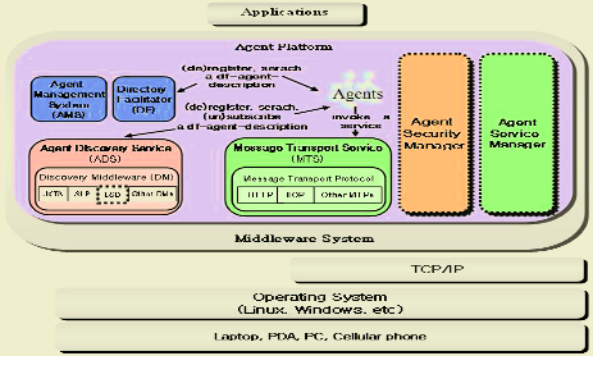


Fig. 1. Agent Platform

Salutation: Salutation is an architecture used for looking up, discovering and accessing services and information. Its goal is to solve the problems associated with service discovery and utilization among a broad set of applications and devices in an environment of widespread connectivity and mobility.

All of the above mentioned service discovery infrastructures and architectures, such as the Service Location Protocol (SLP), Jini, Universal Plug and Play (UPnP) and Salutation, have been developed in order to explore the service discovery issues in the context of distributed systems. While many of these architectures provide good foundations for developing systems with distributed components in networks, they do not adequately solve all of the problems that may arise in a dynamic domain and a large scale network. Thus, in this study, we propose a new scheme to provide an efficient and more scalable architecture, by using a FIPA compatible agent platform. To accomplish this, we modify the FIPA standard specification to provide open scalability.

3 Proposed Agent Platforms

In this section, we describe the proposed Agent Platform (AP) architecture. The overall design of the architecture is shown in Figure 1. We adopt the FIPA reference model and add several new components. The Agent Security Manager is a managing component which monitors agents that register and de-register with the DF. This component provides these agents with secure communication by using authentication, authorization, etc. The Agent Service Manager is a routing solution for ad-hoc networks. In the proposed AP, we create a virtual network overlay on top of the existing physical network infrastructure, as in the case of JXTA. Even though this virtual network allows a peer to exchange messages with any other peer, independent of its location, it does not take into consideration physical network status information such as congestion, bandwidth and other network parameters. With this in mind, we introduce a new component, the Agent Service Manager, to provide our system with physical network context information.

In fact, the FIPA DF federation mechanism is not suitable for flexible service discovery in mobile ad-hoc networks, as described Section 2. FIPA ADS uses the existing discovery protocols, as shown in Figure1. The functionality of ADS is similar to that of DF. The ADS maintains one or more DM modules, each of which provides access to a certain ad-hoc network. Although this architecture follows the FIPA specification to discover agents/services in other domains by using the existing discovery protocols, there are still some problems which need to be resolved. Firstly, in an ad-hoc network, there are many devices and services available and the users are susceptible to move from domain to domain. We cannot know in advance which domain a given user wants to go to or how many DMs he or she needs. Because mobile devices (e.g. PDAs) have limitations such as power, memory, etc, we need to add all DMs present in the AP, in order to cover all domains. Secondly, since the AP can support inter-communication capability by using DM modules, each DM maps its description to a DAD (DF-agent-description). However, due to the DM's peculiar description, it is difficult for these DMs to map each other completely. For these reasons, FIPA ADS cannot support open scalability. To overcome the problems described above and with

the aim of supporting open scalability, we introduce an algorithm based on distributed hash tables (DHTs) in the proposed agent platform. Before describing the extended AP architecture, we summarize some of the currently available DHT algorithms.

CAN: A Content Addressable Network[11] is a mesh of N nodes in a virtual d-dimensional dynamically partitioned coordinate space. The CAN mechanism uses two core operations, namely a local hash-based look-up of a pointer to a resource, and the subsequent routing of this look-up request to the resource using the acquired pointer.

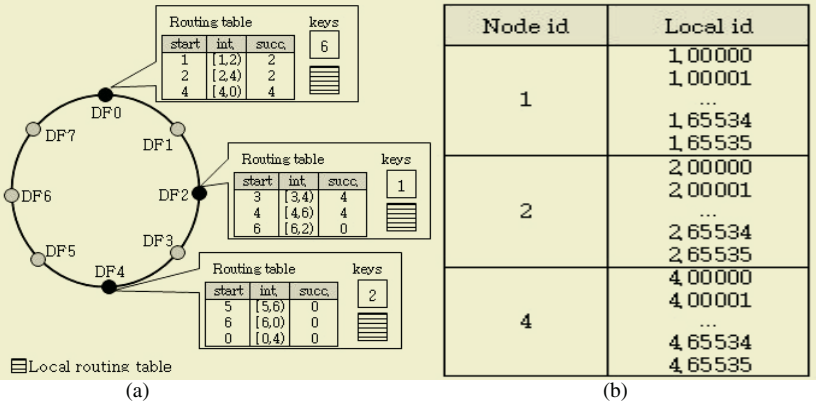
Chord: Chord[12] addresses key location and routing in an overlay network forming a ring topology. Keys are assigned to nodes using a consistent hashing algorithm, enabling a node to locate a key.

Pastry: Pastry[13] is a prefix-based routing protocol. Each node in the Pastry network has a unique identifier in the form of a 128-bit circular index space. The Pastry node routes a message to the destination node, using the identifier that is numerically closest to the key contained in the message.

Note that DHTs were designed to be used in peer-to-peer networks in order to guarantee scalability, and they are unsuitable for pervasive computing, because they rely on the presence of fairly stable underlying networks. The approach proposed in this study is taken partially from one of the DHT algorithms, which operates on top of the agent platform, in order to maintain open scalability. We use Chord-like DHT algorithm for DF federation to support open scalability.

3.1 Directory Federation with DHT

P2P systems are distributed systems without any centralized control or hierarchical structure, in which the application running at each node is equivalent in functionality. Unfortunately, most of the current P2P designs are not scalable or, even if they do provide scalability, it is effective only within its own domain. This means that the current P2P systems cannot provide open scalability which is independent of their supporting scope. The DHT enables the proposed AP architecture to be extended. In such a system, users can communicate with each other independently of the domains in which they are located and without the need for any additional DMs, other than the



(a) (b)

Fig. 2. (a) Routing tables in DF (b) Local routing table of node DF0

ones they are currently using. In the extended AP, we modify ADS so as to enable DHT functionality.

Unlike Chord, the DF using DHT (D^2HT) is a hierarchical mapping algorithm. The FIPA-compatible AP uses DF as its yellow pages in each domain. Thus, DF has access to the local domain information pertaining to agents/services. In the local domain, we can locate agents/services simply by using the DF functionality. However, in a large scale network, the searching mechanism based on DF federation is insufficient, as mentioned in Section 2. For this reason, we adopt a Chord-like algorithm, in order to resolve these scalability problems. It is difficult to apply the Chord algorithm to an ad-hoc network, because peers join and leave the network dynamically in such an environment. Whenever the network configuration is changed, the Chord algorithm must redetermine which peer is in charge of the network, thereby complicating the situation, i.e. the overhead required to keep this information up-to-date is quite substantial. In D^2HT, however, this information does not need to be reconfigured, because the DF in D^2HT not only manages the agents/services in the local domain, but also acts as a superpeer (rendezvous-peer), which manages the local information of peers in JXTA. Also, we assume that nodes with the DF are reliable once they have been initiated. Thus, we adopt a Chord-like algorithm to federate the DFs. The architecture of D^2HT is depicted in Figure 2.

When we create an agent in an AP, the agent registers its name and property with the DF in the local domain. If an agent wants to find another agent, it sends a request message to the DF. If it cannot find all of the services in the local domain, the DF forwards the request message to the DFs in other domains (Figure 2(a) shows the architecture). Domains in different networks use different discovery protocols, as shown in Figure 2(b). The applications on the JXTA can find the services in domain 2 or 3 by using D^2HT. Note that we implement a trivial version of DM for LSD (lightweight service discovery), as explained in Section 4. By using these different versions of the DM, an application in a given domain is able to communicate with an application in another domain. We implement three DMs to validate our D^2HT mechanism.

3.2 Routing in D^2HT

In a dynamic network, nodes can join (and leave) at any time. The main challenge in implementing these operations in the Chord algorithm is preserving the ability to locate every key in the network. But in our proposed mechanism, D^2HT, there is no need to reconfigure the routing table, which is called the finger table in the Chord algorithm. As discussed above, we assume that the DF does not leave and join once it has been initiated. Thus, there is no need to reconfigure the routing table.

We allocate 8 bits per node to make the local routing table. Each node has 65536 entries in each table (Figure 2(b) shows an example of a local routing table). Thus, each node has a range of 65535 distinct values. Because there are nodes continually moving in and out of the network, we should reconfigure the local routing table each time a node configuration is changed. While the routing tables of the DF do not need to be reconfigured, the nodes in the vicinity of the DF may join and leave at any time. However, even so, we can reconfigure the tables only once in a while, because the DF is the mandatory server in each domain, and the network administrator who manages the DF can reconfigure the table whenever necessary. The remaining parts of the

D^2HT algorithm follow the Chord algorithm, with the exception of the node interval. Thus, we modify the Chord algorithm in order to generate the D^2HT algorithm.

As an example, consider the Chord ring in Figure 2(a). Suppose node DF2 wants to find the 1.00001. Since 1.00001 belongs to the circular interval [6,2) (since the prefix of 1.00001 is node id 1), it belongs to the third entry in DF0's routing table. Therefore, DF0 searches for 1.00001 in its local routing table.

As mentioned above, we assign 8 bits to the local routing table. In an ad-hoc network, the number of nodes may exceed the maximum number allowed. If such a situation arises, we need to install a new DF in the domain. Alternatively, all of the services may exist in the routing table for a certain interval (referred to as the time-to-live value). When this value is expired, we delete the entry from the table. This is similar to the concept of leasing, which already is used for service advertisements.

4 Lightweight Service Discovery Protocol

The LSD (Lightweight Service Discovery) protocol, denoted by a dotted line in Figure 1, is one of the discovery protocols which is mapping into DF-agent-description and vice versa by using DM. We designed a new discovery protocol for mobile ad-hoc networks. Such an environment, in which small handheld mobile devices with wireless connectivity communicate with each other, is vastly different from the traditional infrastructure-based environment. For this reason, a lightweight model is needed for the ad-hoc environment.

The service discovery has focused on the service management layer dealing with services provided by devices (e.g. printers, multimedia-systems, sensors, etc.). In an ad-hoc environment, small handheld devices join and leave the network dynamically and the number of devices and services tends to increase. In the proposed service discovery protocol, we use a cache manager which registers information about the services offered by peers in the network. All devices in the network should listen to all advertisement messages and save the local cache for a given period of time. When a user wants to find a particular service, he or she first searches for it in the local

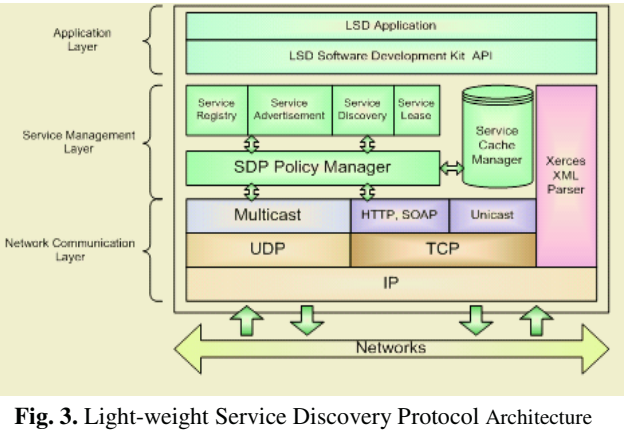


Fig. 3. Light-weight Service Discovery Protocol Architecture

cache. If there is no information in the local cache, the user sends a request message using local-multicasting. Figure 3 shows the LSD architecture.

We use LSD to provide the user with a light version of the service discovery mechanism, whose full functionality can be obtained by using other discovery protocols available in the AP architecture. We assume that the user wants to obtain light services such as printer services, audio-video services, fax services, etc. Thus, a complicated architecture is not suitable for use as a discovery protocol in an ad-hoc network. The LSD architecture shown in Figure 3 comprises three layers. The application layer provides the user with applications such as an audio-video player and a printing application. The service management layer provides the services associated with discovery. The network communication layer is divided into two protocols. The first is UDP, which is used for multicasting; the second is TCP, which is used for unicasting. Advertise/request messages are sent using UDP, and the data is retrieved using TCP. We implemented LSD using Personal Java 1.2 and J2ME CLDC/MIDP. For this implementation, we used WinCE based iPAQ's from Compaq and Jeode VM from Insignia Inc[14]. However, this architecture is a preliminary version, and we are now working on the development of a complete implementation. More details are to be found in [15]. The whole framework discussed so far may be summarized as follows. In an environment with little route stability, such as an ad-hoc network, broadcast service discovery/advertisement is likely to be more effective, as shown in LSD, On the other hand, more stable networks, which can exploit the greater capabilities of their powerful nodes, may be able to take advantage of a structured P2P algorithm such as DHT. The agent platform also enhances the functionality of the service discovery mechanism. The proposed scheme combines these methods in order to solve the open scalability, dynamicity and interoperability problems which are susceptible to arise in highly dynamic, heterogeneous environments, such as ubiquitous networks. Based on these ideas, we present an initial implementation of the combined algorithm and related protocols, as shown in Figures 1 through 3.

5 Conclusions

As more and more people start using mobile devices such as PDAs, cell-phones, notebooks, etc. the need to connect them to the Internet for ubiquitous sharing and for accessing data will increase. An efficient service discovery infrastructure will play an important role in such a dynamic environment. In this study, we reviewed a number of existing service discovery frameworks and found that they all have one problem in common, namely, a lack of inter-operability within the network. In this paper, we introduced an agent platform designed to support open scalability by using D^2HT and the lightweight service discovery protocol for ad-hoc environments. D^2HT, which adopts the Chord algorithm for DF federation, is one of the main contributions of the proposed system. Unlike in the case of the Chord algorithm, in D^2HT it is not necessary to reconfigure the routing table when nodes join or leave the network. We are currently evaluating the performance of D^2HT and working on the full implementation of LSD. In the second phase of our framework, which combines D^2HT with FIPA-compatibility, we will add security and service management components to the framework.

References

1. Service Location Protocol Version 2, Internet Engineering Task Force (IETF), RFC 2608, June 1999.
2. UDDI: The UDDI Technical White Paper.: http://www.uddi.org/ 2000.
3. Ken Arnold et al, The Jini Specification, V1.0 Addison-Wesley 1999. Latest version is 1.1 available from Sun.
4. Projects::JXTA Book, http://www.brendonwilson.com/projects/jxta.
5. UPnP Device Architecture Reference Specification. Microsoft Corporation.
6. G.C.Fox, and Pallickara, S., "The Narada Event Brokering system: Overview and Extensions", In Proc. of the 2002 (PDPTA'02).
7. Christian Becker, Gregor Schiele, Holger Gubbels, Kurt Rothermel., "BASE - A Micro-Broker-Based Middleware for Pervasive Computing", (PerCom'03).
8. L. Capra et al, "Exploiting reflection in mobile computing middleware". ACM SIGMOBILE Mobile Computing and Communications Review, 1(2), 2003.
9. FIPA: Agent Management Specification, Dec. 2002.
10. Salutation Architecture Specification, Version 2.0c, Salutation Consortium, June 1, 1999.
11. S. Ratnasamy, P. Francis, M. Handley, R. Karp, S. Shenker, "A Scalable Content-Addressable Network. SIGCOMM01", August 27-31, 2001, San Diego, California, USA.
12. I. Stoica et al, "Chord: A scalable peer-to-peer lookup service for Internet applications." Technical Report TR-819, MIT, March 2001.
13. A. Rowstron and P. Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems." IFIP/ACM International Conference on Distributed Systems Platforms (Middleware), November, 2001.
14. Insigna Inc. jeode PDA Edition, http://www.insignai.com/content/pda.shtml, 2002.
15. Kee-Hyun Choi, Dong-Ryeol Shin, "LSD: Lightweight Service Discovery in Ubiqitous Environment", Tech. Report, Aug., 2004.

A Mathematical Predictive Model for an Autonomic System to Grid Environments

Alberto Sánchez and María S. Pérez

Computer Science School,
Universidad Politécnica de Madrid,
Madrid, Spain
{ascampos, mperez}@fi.upm.es

Abstract. One of the most important aims of the Grid technology is using geographically distributed resources. Nevertheless, Grid environments have a great problem: the system management is very complex because of the large number of resources. Thus, improving the system performance is a hard task and it would be advisable to build an autonomic system in charge of the system management. The autonomic system must take decisions based on the analysis of the monitored data of the whole Grid trying to improve the system performance. These decisions should not only take into account the current conditions of the Grid but the predictions of the further future behaviour of the system too. In this sense, we propose a mathematical model to decide the optimal policy based on predictions made thanks to the known past behaviour of the system. This paper shows our model on the basis of the decision theory.

Keywords: Grid computing, Autonomic computing, performance models, predictive models.

1 Introduction

A Grid environment [Fos02] could be understood as the result of the geographically distributed computing and storage resources. As system, we should try to get the best performance in any kind of access to any grid element or the whole environment.

Although the Grid community has changed its directions towards a services model, as is described in [FKNT02] and [CFF+], the high number of resources, which constitute the grid, makes difficult the system management and therefore to obtain the maximum system performance.

In order to make easier the system management, administrators must not be in charge of solving the whole system complexity and they could be helped by an autonomic system. Autonomic computing [RAC] [Iac03] is used to describe the set of technologies that enable applications to become more self-managing. Self-management involves self-configuring, self-healing, self-optimising, and self-protecting capabilities. The word *autonomic* is borrowed from physiology; as a human body knows when it needs to breathe, software is being developed to enable a computer system to know when it needs to configure itself and in our case study optimise itself. Taking decisions based on the data analysis is one of the problems that must be solved if we want to

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 109–117, 2005.

obtain and use the strategy that optimises the system and maximises its expected performance.

This decision problem, seen in the field of system analysis and evaluation, can be found in another different areas. For example, in Health Sciences different strategies have been studied to help to take decisions about the illness of a patient, analysing the observed facts and symptoms found in him, like in [Dei] and [Abi99].

Sciences, like Phisics and Chemistry, are characterised by establishing deterministic laws. Thus, if an experiment is repeated in same conditions, same results are obtained because of they are defined by natural laws. Nevertheless, in our field, the usual situations are random and do not comply a natural law. Thus, it is necessary to make a probabilistic data analysis based on the statistic regularity principle. This principle states that if the number of repetitions of a experiment is large, its relative frequency tends to be stabilized in the value represented by its probability.

This probabilistic data analysis should analyse the current state of the system. However, only this feature is not enough. It is required to predict the future system behaviour to select the best strategy that maximises the expected system performance. We must increase the system performance in a concrete point of time and in future actions. Nowadays, there is a lack of mathematical predictive models to take decisions suitable for Grid environments. This paper shows a predictive model based on different techniques, which are used in decision theory and could be used in an autonomic system to make easier this hard task.

The outline of this paper is as follows. Section 2 defines the background of known mathematical models used to build our approach. Section 3 describes our proposal, which is based on the mathematical models shown below. Finally, Section 4 explains the main conclusions and outlines the ongoing and future work.

2 Statistic Models

Getting the best strategy adapted to the necessity of taking decisions is a hard problem. With the aim of simplifying it, first of all, we will use different known techniques.

2.1 Decision Theory

Keeney and Raiffa in [KR93] define the decision analysis as a procedure that helps to people to decide when they have to face difficult situations.

The are several methods to select the best decision. Each one shows a different perspective of tackling the problem.

Maximin. This method uses the criterion of selecting the action that has the best of the worst possible consequences.

In this sense, it is possible that the selected action does not maximise the profit because the aim is obtaining a result that was not harmful. This method could be useful in an autonomic system that wants to improve the response time of the set of applications running over it without getting the minimum respond time of a particular application.

Expected Value. The idea of this method is assigning occurrence probabilities to each event and select the decision whose expected value was the highest one. For calculat-

ing the expected value of each action is necessary to add the result of multiplying the assigned values to each action, if the event occurs, and its occurrence probability.

The problem is how to obtain the occurrence probabilities of the events. These probabilities can be predefined or can be extracted from the past behaviour of the system. In this last case, it would be necessary to accept that the system behaves in the future in a similar way than in the past.

2.2 Decision Trees

In [RS61] the origins of decision trees can be found. Decision trees are a very useful resource to represent decision problems composed of several decisions, making easier the understanding and the resolution of complex problems.

In short, a decision tree is a probability tree that has three types of nodes:

1. Decision nodes. They have tree branches that represent possible decisions.
2. Chance nodes. They represent the possible states.
3. Value nodes. They are terminal nodes that indicate the utility associated to the taken decisions and states visited.

A decision tree can be seen as a set of policies that indicate action plans. The evaluation of decision trees aims identifying an optimum strategy, searching the strategy that maximises the expected utility.

The use of decision trees is useful to represent system policies and how they affect to the system. The disadvantage of solving the system using decision trees is that the change state probabilities must be defined "a priori", and this makes difficult the prediction of future behaviours.

2.3 Bayesian Approach

One of the most important strategies to take decisions is the Bayesian approach. This approach uses the famous Bayes theorem presented in 1763 in [Bay63].

The Bayes theorem is enunciated as the following: being A_1, A_2, ..., A_n a complete system of events with probability different to 0, and being B another event whose conditional probabilities are known $p(B|A_i)$, then the probability $p(A_i|B)$ is:

$$p(A_i|B) = \frac{p(A_i) \times p(B|A_i)}{\sum_{i=1}^{k} p(B|A_i) \times p(A_i)}$$

By using this theorem, we can represent an objective decision procedure taking into account the appearance probability of a certain event based on real facts. This allows us to calculate the probability an event happens while different additional information about the occurrence or not of related event is incorporated.

In short, the steps of the Bayesian approach are the following:

1. Calculate the "a priori" probability that A event occurs.
2. Calculate the probability that different events occur which depend of A, knowing that A has happened.
3. Calculate the "a posteriori" probability the A event occurs knowing that the other events have occurred by means of Bayes theorem.

The Bayesian approach is the most used selection method. In our case, it is not useful to make predictions about the system behaviour. In conclusion, the Bayesian approach allows us to know the probabilities of being in a certain state of the system based on monitored data. But this does not allow us to infer which is the probability that in the future the policy applied by this method maximises the system performance.

Although the Bayesian approach does not solve our full problem, it is an interesting start point to calculate the initial probabilities.

2.4 Markovian Approach

The Markovian approach is based on Markov chains. A Markov chain is a stochastic process that represents a system whose state can change. These changes are not predetermined, although the probability of getting the next state depending on the previous state, is known [Dav93]. In addition, this probability must be constant along time.

Markov chains include in the probabilistic analysis remunerations, numeric value associated to the transition between two states, without assuming nothing about the indicated value nature. This value could include every aspect needed for the model, as losses as profits. After a large number of steps, the transition probabilities between states reach a limit value called *stationary probability*. This one can be used to take decisions.

It is possible that the stationaries probabilities do not depend on the initial state. In this sense, all probabilities tend to a limit value since we increment the number of transitions among states. All states can be reached in the future and the behaviour does not change along time.

A Markovian process is said to have decision if for each transition a variable can be fixed and it is possible to choose between different sets of transition probabilities and different values from associated remuneration.

A Markov chain with remuneration is created from the definition of policies, rules that fixes for each state the value of the decision. For each policy it is necessary to define the probabilities of transition between the different states and the existing remunerations between them.

Due to the decision capacity of the system, the maximum remuneration in a large period of time,is obtained. This maximum remuneration policy is called *optimum policy*.

In order to establish the optimal policy, two methods can be used:

1. Iterating in the state space until the system converge to a certain policy that maximises the expected remuneration.
2. Iterating in the policy space. It allows the problem to be solved in an infinite horizon.

Thus, it will be necessary to define:

1. A matrix of transition probabilities between states.
2. A matrix of benefits or losses obtained when a transition between two states is made. At first the predefined values of remuneration will indicate the obtained benefits if this decision is taken.
3. The establishment of policies in the system. For each policy, probabilities and remunerations must be defined.

Although the Markovian approach allow us to infer the probability to reach a certain state in a further future (infinite horizon), it is necessary not to forget that the future behaviour is random and we are only inferring its behaviour in a probabilistic way.

3 Proposal

The concept of grouping is fundamental in every aspect of the life. Edwin P.Hubble, which is considered the founder of the observational cosmology, said in the thirties that the best place for searching for a galaxy is next to another one, describing the concept of galaxy grouping. Like in real life, computer science has a significant number of groupings, such as process group or user group, which are used for representing sets of objects from the computing field.

In a grid environment, the concept of grouping is very important. The use of different clusters that belongs to the grid generates the abstraction of the server concept, entrusting to the cluster instead of the server the storage of the information. This causes the need of knowing the parameters in a cluster level.

For knowing this parameters, a mathematical formalism should be defined. These parameters could be monitored to improve the decisions about its performance. Two types of parameters can be defined that must be managed by the system:

1. Basic parameters. They lead the operation of the system. By its influence on the performance we might emphasise:
   - Capacity (C). Occupation percentage of the hard disk of each server.
   - Load of the Network (N). Busy rate of the network.
   - Workload (W).
2. Advanced parameters. Parameters that have an influence in the performance of the autonomic system. It is very important:
   - Time window (T). Period of time in which the system monitors its performance. It should be calculated depending of the environment.

But it is necessary to consider and monitor every node belonged to the cluster to know the value of the cluster parameter. The parameter model based on the grouping concept is shown in Figure 1.

There are different methods to figure out the cluster parameters based on the nodes parameters. In the case seen in Figure 1, the parameters values for the cluster are the worst values of all nodes that compose it. In this sense, the system makes decisions based on the worst possible case. Other policies that the system can use are based on other variants.

The autonomic management of the system must face the decisions taken in a certain time trying to improve the future behaviour of the system. The decisions taken in the present should consider the future previsions. For predicting, it should be very useful to consider the past behaviour of the system.

In order to know the past behaviour of the system the events must be monitored each certain period of time corresponding to the time window (T). Then, we can calculate the occurrence probabilities of each event that are updated when new occurrences occur into the system.

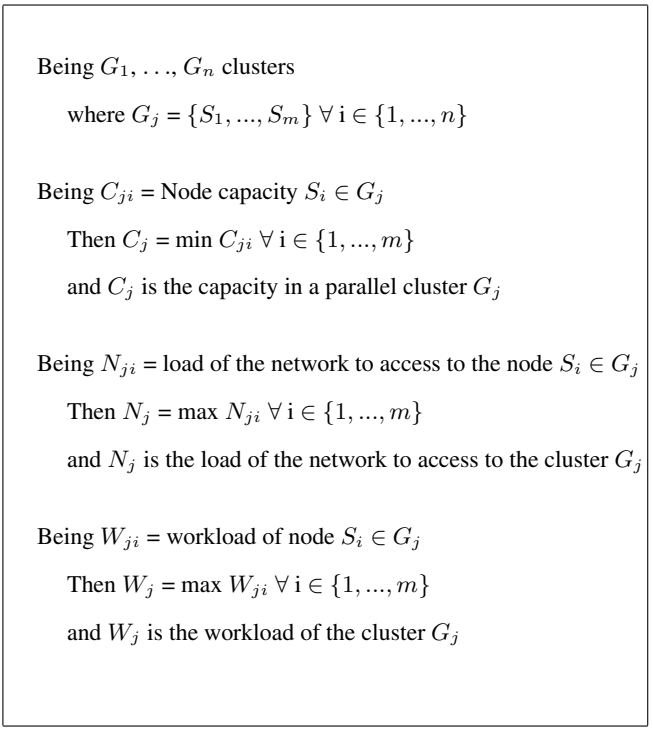


Fig. 1. Parameters model

The system evolves in a non deterministic way around a set of states. Due to this behaviour, the system operation can be formalised following the lines indicated by the statistical method of Markov chains. This statistical method provides the capability of taking decisions.

The initial probabilities for the Markovian approach can be calculated from the data acquired previously into the system. To obtain the values of probability considering the parameters that have an influence in the performance, the Bayesian approach with the formula of the inverse probabilities can be used, calculating the probability that the parameters of the system take certain values in a certain state. By means of Bayes theorem, the inverse probability can be obtained, that is, the probability of staying in a state when parameters have certain values.

This probability calculated for the state changes will not be determinate but it varies with the changes that occur into the system. The Markovian approach defines the probability of reaching a state as a constant, but our model must allows changes of probability along time.

Considering that the obtained probabilities of the system study will vary along time and the predictions can change, the system must be represented by different Markov chains. Each Markov chain will formalise the future operation of the system in a certain moment. A homogeneous operation of the system is assumed so that the probabilities

obtained by the data analysis in the past are taken as reference in the future to predict the system behaviour. Therefore, we build a new Markov chain each T time.

On the other hand, everything that helps to make better predictions should be taken into account. Therefore, the autonomic system should make possible the use of hints of the future system performance. To do this, the autonomic system must modify the remuneration predetermined to the transitions between states based on the extracted knowledge of the use of hints. The hints could change the predefined values of benefit assigned to the Markov model, fitting better the behaviour predicted of the system.

The number of states depends on the number of parameters that define the system. In a first approach, for each parameter, it should be advisable to define a state that represent its normal values, another with optimal values and a last state for values that make worse the system performance. The number of states by parameter could be increased to model in a better way the system, but the complexity of the problem grows exponentially and requires huge calculation time than it could not be assumed if a real time analysis is desired.

After a certain number of transitions sufficiently high, the information of the initial state will have been lost, and the probability that the system will be in a state will depend on this one. In this sense, predictions of the behaviour of the system can be made because the possibility of being in a state is known and it is possible to take decisions based on the collected data.

For the inclusion of the decision policies in the system and mainly with the aim of formalising its representation, decision trees can be used. In this sense, policies are represented by means of decision trees, but the Markovian approach will be applied to solve the optimal policy, being able to take decisions based on the values of transition probability between predicted states.

The different defined policies that manage the system operation can be transformed into a Markov chain with remuneration. Its evolution will be defined by an evolutionary sequence that follows a Markovian decision process that affects to the transition to the following stage. This kind of systems is denominated E/D (Evolution/Decision) systems.

4 Conclusions and Future Work

In this paper we have deeply analysed the way of obtaining a better performance in a Grid environment by using an autonomic system. We propose to improve the taken decisions selecting the optimum policy. This optimum policy should not only take into account the current conditions of the Grid but the future behaviour of the system. We have analysed different mathematical approaches to get the benefits that can contribute to our model.

In summary, our goal is to make a prediction about the future behaviour of the system according to the analysis of the current and the past system states. Thus, it will be necessary to do the following steps:

1. Finding the occurrence probabilities of being in a certain state, the values of the system parameters are in a certain rank. Thus, system logs and different monitoring tools could be used.

2. Finding the probability that the system is in one or another state knowing that the parameters of the system have a certain value by means of the Bayes theorem.
3. Creation of the initial matrix of probabilities of transition between states, from the previously collected data.
4. In case of the use of hints, it will be necessary to modify the matrix of remunerations obtained previously based on these ones, which indicate the future operation of the system. Thus, it is necessary to increase the remuneration in the transition towards those states that have a greater probability of occurrence according to indicated in the hints.
5. Calculating the matrix of remunerations, trying to lead the system towards the most beneficial states to reach an improvement in the system performance.
6. Establishment of policies in the system formalising them by means of decision trees. The Markovian problem will be affected so that the previous matrices can have different probabilities and remunerations for each policy.
7. Resolution of the problem of searching the optimal policy that maximises the expected remuneration following the proposed Markovian approach. This will be the decision that will be taken in real time on the system to improve the future performance of such system.

As future work, we aim to adapt this mathematical model to a real environment. Furthermore, we want to study the increase of the number of states that define the problem to adapt it in a better way to environment changes, and how this affects to the computing time. Finally, we will analyse different monitoring tools to find one which can be adapted to our needs and, therefore, which allows us to measure new important parameters.

Acknowledgements

This research has been partially supported by Universidad Politécnica de Madrid under Project titled “MAPFS-Grid, a new Multiagent and Autonomic I/O Infrastructure for Grid Environments”.

References

[Abi99] Syed Sibte Raza Abidi. Ai amidst the healthcare revolution: Towards an ’intelligent’ tele-healthcare environment. In Hamid R. Arabnia, editor, *IC-AI*, pages 172–176. CSREA Press, 1999.

[Bay63] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philosphical Transactions of the Royal Society*, 53:370–418, 1763.

[CFF+] Karl Czajkowski, Donald F. Ferguson, Ian Foster, Jeffrey Frey, Steve Graham, Igor Sedukhin, David Snelling, Steve Tuecke, and William Vambenepe. The WS-Resource Framework version 1.0.

[Dav93] M.H.A. Davis. Markov models and optimization. *Monographs on Statistics and Applied Probability*, 49, 1993.

[Dei] Stefano Rizzi Deis. Medical decision support in clinical record management systems.

[FKNT02] I. Foster, C. Kesselman, J. Nick, and S. Tuecke. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration, 2002.

[Fos02] Ian Foster. What is the Grid? A Three Point Checklist. *Grid Today*, 1(6), July 2002.

[Iac03] IBM and autonomic computing, editors. *An architectural blueprint for autonomic computing*, April 2003.

[KR93] R. L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press, 1993.

[RAC] IBM Research Autonomic Computing. *http://www.research.ibm.com/autonomic/*.

[RS61] H Raiffa and R Schlaifer. *Applied Statistical Decision Theory*. Harvard University, 1961.

Spatial Analysis: Science or Art?

Stefania Bertazzon

Department of Geography, University of Calgary,
2500 University Dr. NW, T2N 1N4Calgary, AB, Canada
bertazzs@ucalgary.ca

Abstract. Spatial Analysis is a relatively young discipline, descending from a modelling tradition where the analyst possesses all the knowledge and qualities that lead him/her to the definition of the optimal model. Numerous spatial analytical techniques are available these days in general-purpose GIS software, but their user interfaces are dry and do not offer structured choices in pull-down menus, as they do for more conventional operations in GIS. The average GIS user is often unprepared to identify the right solutions without guidance. Defining optimizing criteria and introducing them in structured software interfaces appears to be, at present, the best means to promote a widespread and appropriate use of spatial analysis. Defining such criteria constitutes at the same time an important line of research, potentially capable of furthering the theoretical underpinnings of the discipline, aiding its transition from infancy to maturity.

1 Introduction

As more and more sophisticated methods for spatial analysis become available, an increasing level of knowledge and understanding is expected on the user. The emergence of local as opposed to the more traditional global methods is the most recent example: which method is more appropriate for a given spatial process? A global estimator, which addresses the inefficiency induced by spatial autocorrelation, or a local one, that best deals with local non-stationarities? Is the average GIS user prepared to answer this question and to make the choices that stem from this decision? This paper outlines a discussion of the dangers that may arise when a non-expert user applies a powerful statistical technique, which requires delicate and informed choices, liable to alter the results of the analysis. The alternative solution is providing the inexpert user with a handful of guided choices, thus shifting the burden to the developer of the analytical software. The latter solution presents several advantages, and may be at present the most appropriate.

1.1 Background

When, in my lecture on multivariate statistics, I introduce principle components analysis, I start the lecture by writing on the board a bunch of variables and I ask my students to organize them into groups, in order to reduce their number. Every time, each class proposes several alternative groupings, all equally legitimate. In

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 118 – 125, 2005.

commenting the subjectivity of that exercise, I have occasionally used the term *alchemy*: the alchemist believes that he can extract gold from some less noble minerals, and he proceeds, guided by his religious and philosophical creed, to implement methods and make choices, some right, some wrong, but he cannot count on a rigorous methodology to determine specific procedures, temperatures and reagents, which would vary depending on the mineral at hand, the available tools, and environmental conditions. And he usually fails to achieve his goal.

Does an objective and rigorous method exist, to extract, if not gold, at least some meaningful groups from our set of variables? It certainly does: there exists a scientific method, based on the computation of correlations, eigenvectors and eigenvalues; rigorous and objective, applicable to any set of variables, and capable of guiding any researcher to identical results: it is what we call principal component analysis. Once the method has been applied, one must complete the next stage: the interpretation of those golden components. And here all the textbooks say more or less the same thing: the components, mathematically derived, are often difficult to interpret: it is up to the researcher, based on their sensibility and knowledge of the data, to understand and define the sense of those components [14; 11]. I suspect that my students fail to catch the subtlety, but I still feel that in this stage we are playing the alchemists again!

In the introductory class to my "advanced spatial analysis and modelling" class, I define spatial analysis[1] and place it in the context of GIS, I than encourage a discussion on the definition of GIS: the already classic debate between Geographic Information *Systems* and *Science* [4]. After the ritual oscillations, the class usually reaches a consensus on the idea that spatial analysis is the element which, more than any other, contributes to make GIS a Science. Once, however, a student surprised me by proposing the thesis that GIS is nor a System or a Science, but an Art. I believe I reacted to that comment with an intrigued smile but, having to pacify the rest of the class, I concluded that the definition of Art is not wrong, but certainly more appropriate to such sub-disciplines of GIS as cartography and visualization, than to spatial statistics.

Alas! While discussing the optimal number of variables in a multivariate model, the choice among different models, or the strategies to optimize their properties, I have caught myself stating that spatial analysis is an art! And the more I think about the techniques we use and the ways in which we use them (principal component analysis is only a trivial example – not necessarily applied to spatial data), the more I tend to think that that definition, apparently so extravagant, is in fact appropriate for spatial statistical analysis. Indeed, I tend to think that it is the only definition that can provide a key to resolve such contradictions as the one between the rigor of a method and the subjectivity of its interpretation.

2 Art and Science of Spatial Analysis

The lack of objective criteria in spatial statistical analysis certainly is not a new problem; the novelty is in the increasing availability of routines for the implementation of such analyses in standard commercial GIS software. Consequently,

[1] I am using the term *spatial analysis* to refer to models and quantitative techniques beyond simple operations such as overlay, spatial queries, buffering, etc., which yet are often termed *spatial analysis* in introductory GIS textbooks; see, for example, [6].

an ever growing number of users daily approaches such analyses, but often get trapped when questions arise to which they do not know the answer. What is even more disconcerting is the range of results that can be obtained from the same analysis by only varying some of its parameters[2].

What is important, in such situations, is not really which parameters are chosen in a specific application, but whether some criteria (and if so which criteria) should be set to guide such choices, or the choice of the parameters should be left to the user's discretion. The latter question may be tackled only by considering the skill and knowledge that may be expected from the user, and in turn it opens a wide range of solutions. On one end of the spectrum are the *ad hoc* solutions, subjective and specific to each application: what I take the liberty of calling Geographic Information Art, where the choice is left to the user's sensibility, from which the final results of the analysis depend. By narrowing down the possibilities, establishing borders and suggesting paths, one reaches the other end of the spectrum: the definition of objective and rigorous criteria. Defining such criteria corresponds to what is commonly known as Geographic Information Science, but such science[3] is usually exerted by those who are behind the commercial software, and the user can only choose within a narrow range of options, following the instructions provided.

How many variables should a regression model contain? What is the best cell size in a point pattern analysis? What is the ideal degree of the polynomial in a trend surface analysis? How many intervals should be defined for a variogram? Is a multivariate model better than an autoregressive one? These are the questions that my students ask me every day. But who are my students? The typical audience of my lectures consists of students of a Master in GIS: they are graduates, they often have a considerable background and experience, and in a few months they will go and work for the private industry or some government agency, as high level technicians or managers, with tasks involving the diverse GIS applications: from urban planning to national park management, from sales organization to the management of coastal traffic. They are the professionals who, within a few years, will use the science and the technology to direct the future of our planet. But there is more to this. A few years ago I was myself employed in the private sector, working in the GIS industry in Europe: the average specialization level was perhaps slightly lower, the applications equally ambitious, the problems to solve equally important. And the questions that my colleagues used to ask me every day were just the same. Why are the questions always the same? Is there a *fil rouge* that connects students and professionals, through diverse applications, in different countries and continents, and throughout the years?

2.1 Science and Users

Considering the series of questions that systematically puzzle the various users, it is clear that those questions do have at least one, apparently trivial, common trait: they are all questions to which no answer is provided in the pull-down menus of commercial GIS software. The average GIS user does not know the answer to those questions, gets hampered and, incapable of finding himself a solution, must revert to his instructor or his more expert colleague.

[2] The range of variation depends both on the routine and the data employed.
[3] I will not attempt a discussion on the definition of science in this context.

The scenery that is emerging is indeed disconcerting: these people, who should call themselves GIS professionals, behave instead as simple users. The analyses that these professionals conduct are the scientific ground for the "big decisions" that concern our future: safety of high speed trains, environmental impact assessments, etc.… Yet these professionals are experts in data management, they are masters in representing phenomena via vivid and effective maps, perhaps they know which analyses are the most appropriate, but when it comes to defining the details of such analyses – important details, liable to change their results- they do not have enough knowledge, they do not have enough critical judgment to be able to choose the appropriate parameters. And they must revert to someone else's advice.

I find the situation preoccupying, and I am personally convinced that this tendency is not limited to the GIS world, but pervades already the many disciplinary and professional realms where user-friendly interfaces have replaced the traditional systems, where the users were expected to have some knowledge and understanding of the tool they were using. Ten years ago Umberto Eco [8] referred to the former as Catholic, the latter as Calvinistic systems. I find Eco's (*ibidem*) essay, though ten years old, disconcertingly current. What should be noted of these ten years is the turn undoubtedly catholic imposed by Windows on the old DOS, and I believe that very few doubts are left about the triumph of such Catholicism and its champion[4]. I believe that if the community of professionals -of GIS and beyond- has slowly become a community of users, much of the responsibility should be attributed indeed to the unquestioned supremacy of that operating systems, which interfaces the software environment to which these professionals –to which indeed we all- are accustomed: the reassuring environment of the pull-down menus, where the user is never asked to make a real choice, where indeed the possible choices have been delimited by the software developer, and all the user is expected to do is choose one of a fistful of alternatives.

In this protecting and comfortable world, users are well aware that everyone is entitled to reach --if not the Kingdom of Heaven--the moment in which their beautiful map is printed; and it is only natural that they feel bewildered when they cannot obtain, with the same ease and confidence, the end results of their spatial statistical analyses. Addicted to the catechistic approach, they ask to be taken by the hand and guided in any choice: they cannot make the transition from users to professionals of GIS. They cannot exert the Art of spatial statistics.

2.2 Art and Professionals

Trying to define Geographic Information Art cannot proceed without a definition of art: an objective that exceeds by far the scope of this paper. Perhaps an acceptable idea is that art is associated with something that is esthetically pleasing. Is this the art that is required to perform good spatial analysis? I think that very few people would get aesthetic pleasure from a well calibrated spatial analytical model! On the other hand, very few –I believe- can appreciate that a good statistical model is not made

[4] This is not the appropriate place to discuss the destiny of the Macintosh and its influence on the dominant doctrines.

only of sound theories and rigorous applications. A *good* spatial analysis is certainly made of experience, theoretical knowledge, an in-depth understanding of the techniques, but also, and perhaps most of all, of intuition, logic, sensibility, intelligence, perspicacity: in short, of that set of personal qualities that make some individuals artists. Choosing the *optimal* number of parameters in a multivariate analysis; identifying the function that best approximates the *crucial* properties of a distribution; *interpreting* the components defined by correlations in a vector space; all this requires knowledge and understanding of the phenomenon under scrutiny – what is known as the *spatial process*. What is needed therefore is an abstract and theoretical knowledge of the analytical tools, a practical and applied knowledge of the spatial process, not less than those ineffable individual talents that make each individual take different choices: some better, some worse, some right, some wrong.

We have always thought that GIS fulfils its artistic vein in the production of the map: an exercise in color matching and relief rendering, whose result can indeed give aesthetic pleasure. But the age of hand drawn maps is long gone! When I think about the production of those contemporary beautiful maps, I can only picture a user who mechanically selects from a pull-down menu pre-packaged ranges of color and contrasts of light and shade; and when I think about the work behind that user, I picture a programmer who encodes light and colors in a dry and rigorous language. Paradoxically, it is indeed in the phases of definition and implementation of statistical analysis that the user is required to be a bit of an artist, he is asked to make difficult and subjective choices, it is expected that he can reach, by himself, the end results of the optimal model. And this is because spatial statistics descends from a modeling tradition that knows no users, but only professionals that have all the qualities that it takes: a deep knowledge of the theories on which the techniques are based, a profound comprehension of the spatial process under scrutiny, and finally those individual artistic talents that make only some persons reach the right model: the Art that complements the Science.

2.3 The Coming of Age of Spatial Analysis

I believe I am (and I am considered) a supporter of GIScience, and even more so a promoter of spatial analysis as a qualifying element of GIS. I also share the proclaimed needs for standardization and interoperability in GIS; I share them so much that I am myself suggesting that also spatial statistical analysis should implement uniform criteria, such that they can result in software solutions capable of making all users converge to an identical result. I am so convinced of it that a great deal of my recent work aims at the definition of criteria, flexible yet consistent, for the assessment of spatial autocorrelation in spatial regression models [5].

Am I thus siding with GIScience, much to the detriment of GI-Art? Faced with the widespread and criticized inaptitude in the professional practice of spatial statistical analysis, am I claiming to myself and a handful of colleagues (almost a sect) the right/duty of guiding users in the choices that they are not ready to make by themselves? Are we indeed (myself and my sect of spatial analysts) taking the role of catechists that we have criticized in those who have undertaken a lead role in software interfaces and operating systems? Does it make sense, these days, to propose uniform and optimizing criteria, as opposed to letting the users make their own choices? My answer is twofold.

On a superficial level, I observe a deep gap between the structured procedures for standard GIS operations served on a silver plate by GIS software on one hand, and on the other hand the almost complete lack of guidance in the interfaces of specialized spatial analytical software[5]. I am afraid that, to this day, GIS users are not ready to make independently the required choices, and the reaction of the user community to this *impasse* takes different forms, ranging from the incorrect use of spatial analytical techniques to –even worse- the disregard or even refusal of techniques which are difficult to understand and demanding to implement.

There is, however, a recent tendency to fill that gap, by means of software products packaged with more interactive and cheerful interfaces, more inclined to take the user's hand and guide him step by step – Anselin's GeoDa [1] is perhaps the best example. I am pleased to notice that such tendency is matched by the literature: traditionally spatial analysis textbooks were unfriendly and specialized [7; 9], and there was a deep gap between these and introductory texts -indeed only very introductory and general texts. Some recent textbooks [2, 12] approach the subject from the round, basic concepts, and push it up to the most recent and advanced developments of the discipline. Providing a more structured set of solutions, by means of guidelines and criteria, may prove to be the best way to encourage a more widespread and appropriate use of spatial analysis. The tendency that is appearing both in software and textbooks is not only reinforcing the validity of structured interfaces, it also suggests a further means to promote knowledge and an appropriate use of spatial analysis: teaching it since the early stages of university and college GIS programs. This permits not only to build the necessary foundations, but also to promote an appreciation for spatial analysis, by showing that it represents a necessary tool in the toolbox of the GIS professional.

On a deeper level, I am convinced that a common criterion, an "objective" criterion, if not a universal solution – in many cases does indeed exist. In the discipline of geography we are not accustomed to searching for such criteria, as we often "let our data speak", an attitude that may lead to a dangerously wide span of solutions. In my quest for a uniform criterion for the measurement of distance in a spatial autocorrelation model, the objective is measuring the variance associated with the regression model [5]. My work therefore represents an example of how the quest for a standardizing criterion represents by itself an important theoretical solution for the discipline.

Spatial analysis is a relatively young disciplinary field, where many of the current solutions are borrowed from other disciplines [3]: such solutions are consequently necessarily limited in scope and meaning, sometimes so crude that in attempting to solve one aspect of a problem they open up a new one. Such strategies might have been acceptable during the infancy of the discipline, but presently a call is heard from many parts, both within an outside the spatial analysis community [10; 13] for spatial analysis to develop a body of its own theories and techniques, such as to tackle the root of the problems, and to provide, in the long term, encompassing solutions, such as to open new research frontiers and to define a mature and independent discipline. Instrumental to the evolution of the discipline is the diffusion of a spatial analysis culture, awareness of the available techniques and of the benefits that the application

[5] By they stand-alone programs, or modules of lager packages.

of such techniques can provide: on one hand the superiority of specifically spatial techniques in contrast to traditional methods, on the other hand the complementarity of such techniques to other methods for the comprehension of spatial processes. In order to achieve a spatial analysis culture it is necessary that a growing number of users and professionals be willing to use the techniques, but it is also necessary that such techniques be used in a rigorous and correct manner.

Providing criteria and guidelines may be simply a means, necessary in a specific historic phases, but it might be the only possible and practicable means, because it is the best means to communicate in the only language (the software interface) spoken today by professionals and users all over the world.

3 Conclusion

Spatial analysis is a young discipline, descending from a quantitative tradition in which the analyst possesses all the knowledge and skills that make him achieve the best model. Numerous analytical techniques are available today in standard GIS packages, but their user interfaces are arid and do not offer structured options in pull-down menus, as it is the norm for most standard GIS operations. The average GIS user is often incapable of identifying the best solutions in the absence of guidelines.

The definition of optimizing criteria and their inclusion in structured interfaces of GIS software may constitute, in the present historic phase, the best means to promote a widespread and correct use of spatial analysis. The definition of such criteria represents, at the same time, an important line of research, potentially capable of developing the theoretical bases of the discipline, helping its transition beyond the threshold of infancy.

Acknowledgements

I wish to thank GEOIDE (NCE- Network of Centers of excellence) for funding our project “Multivariate Spatial Regression in the Social Sciences: Alternative Computational Approaches for Estimating Spatial Dependence”; my students and colleagues at the University of Calgary for the ideas shared in many discussions, that helped me mature the contents of this paper.

References

[1] Anselin L 2003, *GeoDa 0.9 User's Guide*, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics and CSISS, University of Illinois, Urbana, IL.

[2] Bailey and Gatrell (1995) Interactive Spatial Data Analysis. Prentice Hall.

[3] Berry B., 2004, Spatial Analysis in Retrospect and Prospect. Spatially Integrated Social Science. Edited by M. Goodchild and D. Janelle. 2004 Oxford University Press pp. 443-445.

[4] Bertazzon S (2001) The Name of GIS: Geographic Information *Systems* towards a *Science*. *Rivista Geografica Italiana*, Vol. 3, No. CVII, September 2001, pp. 409-440.

[5] Bertazzon S. (2004) Determining an optimal distance metric for the spatial autocorrelation model in spatial regression analysis. GIScience 2004. University of Maryland, October 20-23, 2004
[6] Burrough, P.A., McDonnell R.A. *Principles of Geographical Information Systems*, Oxford: Clarendon Press, 1998.
[7] Cressie, N.A.C. *Statistics for Spatial Data*, New York: Wiley, 1993.
[8] Eco, U, 1994, MAC *vs* DOS. In U. Eco, La bustina di Minerva. 20000 Bompiani. Pp. 165-167
[9] Fotheringham A S, Brundson C, Charlton M, 2000, *Quantitative geography. Perspectives on Spatial Data Analysis*. London: Sage.
[10] Goodchild M., Janelle D., 2004, Thinking Spatially in the Social Sciences. Spatially Integrated Social Science. Edited by M. Goodchild and D. Janelle. 2004 Oxford University Press pp. 3-17.
[11] Griffith, D. A., and Amrhein, C. G., 1991. Statistical Analysis for Geographers. Prentice Hall, Englewood Cliffs, New Jersey.
[12] Rogerson, P. 2001. *Statistical Methods for Geographers*. London: Sage.
[13] Sui, 2003, Terrae Incognitae and Limits of Computation: Whither GIScience? Computers, Environment and Urban Systems. Vol. 25, Issue 6, pp. 529-533.
[14] Tabachnick, B. G., Fidell, L. S., 2001, *Using Multivariate Statistics*. Fourth Edition.2001, Allyn & Bacon

Network Density Estimation: Analysis of Point Patterns over a Network

Giuseppe Borruso

Università degli Studi di Trieste, Dipartimento di Scienze Geografiche e Storiche,
Piazzale Europa 1, 34127 Trieste, Italia
giuseppe.borruso@econ.units.it

Abstract. This research focuses on examining point pattern distributions over a network, therefore abandoning the usual hypotheses of homogeneity and isotropy of space and considering network spaces as frameworks for the distribution of point patterns. Many human related point phenomena are distributed over a space that is usually not homogenous and that depend on a network-led configuration. Kernel Density Estimation (KDE) and K-functions are commonly used and allow analysis of first and second order properties of point phenomena. Here an extension of KDE, called Network Density Estimation (NDE) is proposed. The idea is to consider the kernel as a density function based on network distances rather than Euclidean ones. That should allow identification of 'linear' clusters along networks and the identification of a more precise surface pattern of network related phenomena.

1 Introduction

This research is focused on examining point pattern distributions over a network, therefore abandoning the usual hypotheses of homogeneity and isotropy of space and considering network spaces as frameworks for the distribution of point patterns. Many human related point phenomena are distributed over a space that is usually not homogenous and that depend on a network-led configuration. That is the case of population or commercial facilities, which locate in space depending on road network infrastructures. The study of networks themselves can also be biased by assumptions on homogeneity and isotropy. The problem of analyzing networks and their influences over space can be reduced to analyzing their point structure, as proposed by some authors ([1]; [2]), but the consequent analyses focus on clusters in a homogenous space.

The starting point of the present research is point pattern analysis, which both in spatial analysis and GI science is commonly adopted and used. Kernel Density Estimation and K-functions (Ripley's and Diggle's) are commonly used and allow analysis of first and (reduced) second order properties of point phenomena. Authors ([3]) have proposed methods for estimating K-functions over a network structure. Here an extension of KDE, called Network Density Estimation (NDE) is proposed. The idea is to consider the kernel function as a density function based on network distances rather than Euclidean ones. One of the advantages of such estimator is that it

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 126 – 132, 2005.

should allow identification of 'linear' clusters along networks and a more precise surface pattern identified of network related phenomena. NDE could find application in 'traditional' environment of KDE, as population analysis, crime studies, retail analysis, and network studies.

2 Point Pattern Analysis: Density Estimation

When dealing with a point pattern, authors like Bailey e Gatrell ([4], [5]) consider *events*, referred to observed phenomena over a point distribution, and *points*, referred to all the other places in the study area. The simple observation of events' distribution over space can provide initial information on the structure of the distribution, but more refined analytical instruments are needed for more in depth analysis, and particularly to identify clusters or regularity in the distribution.

Quadrat analysis is one of the means of ordering the pattern of a distribution of events within a region *R*. The procedure involves dividing the study region in sub-regions having equal and homogeneous surfaces, or quadrats[1]. The following step consists of counting the number of events falling in each sub-region (quadrat) in order to simplify and group the spatial distribution.

The number of events therefore becomes an attribute of the quadrat. It is then possible to represent the spatial distribution by means of homogenous and easy comparable areas. Density analyses become possible using an easy to use and to compute method ([6], [7]).

The method has some disadvantages, as the loss of information from original data, as well as different levels of arbitrariness deriving from:

- The choice of quadrat dimension;
- The orientation of the grid;
- The origin of the grid.

Different analyses could be computed, changing a grid's origin or the quadrats' dimensions. One of the solutions involves considering the number of events for each area unit within a mobile 'window'. A fixed radius is chosen and we hypothesise to centre it in a number of places in the space, where events are organised in a grid superimposed to the study region R. An estimate of the intensity in each point of the grid is therefore provided. This generates an estimate of the variation of the intensity smoother than that obtained from a fixed grid of square cells superimposed.

Such method is very close and at the basis of the procedure called Kernel Density Estimation (KDE).

The kernel consists of 'moving three dimensional functions that weights events within its sphere of influence according to their distance from the point at which the intensity is being estimated' ([6]).

The general form of a kernel estimator is

$$\hat{\lambda}(s) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{s - s_i}{\tau}\right) \tag{1}$$

[1] Other tessalations of space are also possible, as triangles, hexagons or other polygonal shapes.

where $\hat{\lambda}(s)$ is the estimate of the intensity of the spatial point pattern measured at location *s*, s_i the observed i[th] event, *k( )* represents the kernel weighting function and τ is the bandwidth.

For two-dimensional data the estimate of the intensity is given by

$$\hat{\lambda}(s) = \sum_{d_i \leq \tau}^{n} \frac{3}{\pi\tau^2}\left(1 - \frac{d_i^2}{\tau^2}\right)^2 \quad (2)$$

where d_i is the distance between the location *s* and the observed event point s_i. The kernel values therefore span from $\frac{3}{\pi\tau^2}$ at the location *s* to zero at distance τ ([6]).

The kernel density estimation function creates a surface representing the variation of density of point events across an area. The procedure can be organized in three steps ([8])

- A fine grid is placed over the study region and the point distribution
- A moving three-dimensional function visits each cell and calculates weights for each point within the function's radius (threshold).
- Grid cell values are calculated by summing the values of all circle surfaces for each location.

3 Network Density Estimation (NDE): The Algorithm

Kernel Density Estimation allows finding out clusters in point pattern distributions over a study area particularly highlighting 'circular' clusters. However clusters can appear also following different distribution schemes as network-led spaces. Here a procedure for considering network spaces in a point distribution is presented.

The algorithm foresees in particular the modification of the searching kernel function from a circular to a network-based service area.

Steps in the algorithm:

1. selection of a point process (i.e., population, ATM, robberies, services' locations) ;
2. generation of a regular grid over study area;
3. generation of centroids of cells belonging to regular grid overlapped to study area;
4. definition of a bandwidth; (same process as Kernel Density Estimation)
5. selection of a network;
6. computation of service area analysis from cells' centroids over network
7. overlay (= spatial intersection) of service areas and point process;
8. count of point (events) per service area;
9. assignment of count (= weight; relative density, etc.) to cell centroid;

9a. if necessary, before visualising density surface a further interpolation could be performed between cell's centroids in order to smooth the density surface]

10. Visualisation of density surface.

The derived density function reflects the network structure of the space. Point processes are therefore analysed rejecting the hypotheses of homogeneity and isotropy of space, considering the network-driven structure of the pattern. The density function is therefore not the result of a circular search radius but of a network-shaped one. Doing so allows evaluating more precise densities of network based phenomena, as ATM locations, burglaries, etc. One of the advantages in using such modified kernel density estimation is that clusters can be more easily detected when phenomena group along a street or a road, what is not always perceivable in traditional, Euclidean KDE.

4 Network Density on Point Patterns: Applications

4.1 Kernel Density Estimation on Networks

An application of the Network Density Estimation procedure was carried out using GIS software and spatial statistical packages considering different phases and steps. The example considered starts from a research on network structures, where KDE algorithm was used to highlight network spaces starting from the spatial distribution of nodes ([2]). In that occasion a network density analysis was performed over a point distribution consisting of nodes of a road network. It was assumed that the network structure could be simplified by using nodes instead of arcs and therefore compute a density analysis. The aim of the network density analysis was the identification of network spaces and particularly centres of urban areas and settlements. As KDE allowed the identification of peak areas corresponding to higher concentration of nodes - junctions in the road network – centres and subcentres in urban areas and peri-urban settlements could be highlighted.

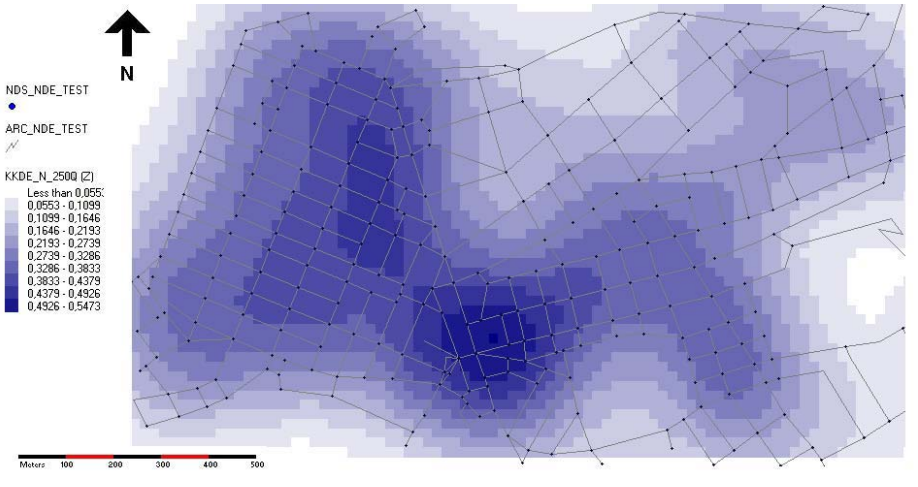


Fig. 1. Kernel Density Estimation computed over road network's nodes distribution in the Trieste (Italy) city centre. 250 m Bandwidth (Density surface produced using CrimeStat 1.1 – [9])

Figure 1 shows the results for a KDE performed over the road network of Trieste city centre (Italy). A 250 m bandwidth was chosen and nodes were weighted by the number of arcs converging to each node. The bandwidth of 250 m generates a surface that presents several peaks and therefore seems more suitable to highlight local interaction phenomena. Darker areas represent clusters of highly connected junctions in the road network. The density decreases as one moves from the 'centres', characterized by high values of road transport network density.

Although interesting results can be found in the analysis of network structure and shape, particularly when they are compared with the built environment's distribution, some limitations in the use of a 'pure' KDE to network point datasets exist. When dealing with network structures in fact the spatial configuration of point distribution over a network cannot be considered as lying onto a homogeneous and isotropic space.

Such a limitation become particularly evident when linear clusters appear. Using KDE, which relies on a circular searching function, such kind of clusters are not always detected. A network density estimator therefore appears as a more suitable solution for highlighting a network's structure and orientation.

4.2 Network Density Estimation

A initial version of the Network Density Estimation algorithm was implemented to adapt density analysis over network spaces. As suggested in paragraph 3, one of the steps of NDE involves the creation of service areas for each grid cell's centroid. Service areas were therefore used to simply count the number of nodes falling inside and then assigning such value to cell's centroid. In this initial application of the algorithm the proximity of the events to centroids into the service area was not considered.

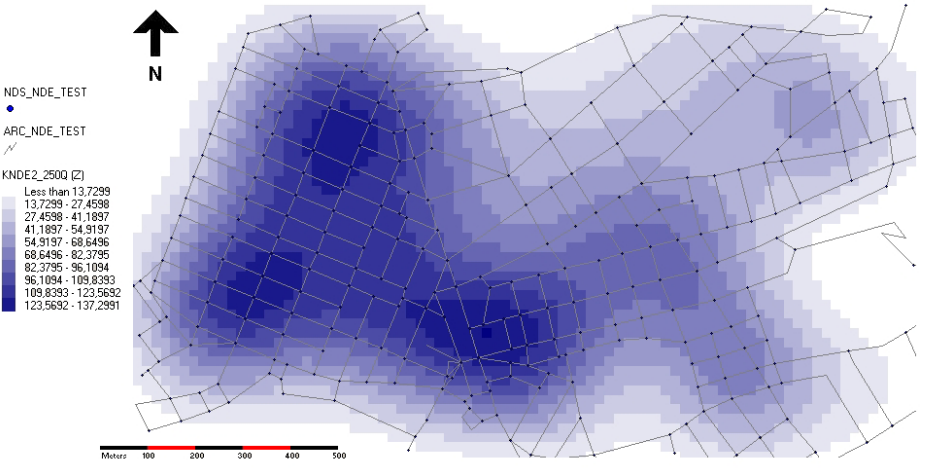


Fig. 2. Network Density Estimation computed over road network's nodes distribution in the Trieste (Italy) city centre. 250 m Bandwidth. (Density surface produced using CrimeStat 1.1 – [9])

Service areas were computed in a GIS environment, as well as the nodes' count and assignment to service areas' centroid. As not all of the cells' centroids fall onto the network or close to it, not all of the grid cells were filled with density values. To avoid the presence of 'holes' in the final graphical representation of the results, the values from the density analysis were interpolated to obtain a KDE-like density surface. Interpolation was performed using a spatial analytical package.

Figure 2 shows the results of the NDE performed over the road network of Trieste. The 250 m bandwidth was used to compute distances over the network rather than in the Euclidean space. Nodes – junctions – were counted and weighted by the number of arcs converging to each node.

Centres highlighted by higher values of network densities can be spotted in Figure 2. If we compare the results from the NDE with those previously obtained with the KDE we can notice different peak areas in the NDE analysis particularly in the western part of the map and a lighter one on the north-eastern part. Such peaks seem to be more consistent with the pattern drawn by the real road network.

5 Conclusions

The first applications of the Network Density Estimation procedure provided interesting results in highlighting network-driven distributions of events in space and therefore reducing the hypothesis of homogeneity and isotropy of space that to some extent is contained in the Kernel Density Estimation algorithm. NDE performed over road network's junction distribution highlights in particular clusters following the orientation of the network, showing network spaces and centres in a more proficient way.

Further research will involve tackling different open issues, both regarding network density analysis and concerning the study of other point-like phenomena distributed over a network space.

With reference to the 'pure' network density analysis there is the need to implement the NDE's steps within an algorithm in a complete GIS or spatial analytical environment. A distance-weighting function should be inserted to weight events according to the distance from the service areas' centroid.

With reference to the applications of the algorithm, when continuing the experiments on networks' spaces minor, linear settlements should be considered to test clustering along main roads. An application of the algorithm to real transport networks should also involve the directional constraints an urban network usually has, as one ways and limited access arcs. That should interestingly affect the shape of service areas obtained and offer different network density areas in an urban environment.

Finally the algorithm should be performed over other point datasets, as ATM, post office, retails, etc. in order to study their distribution along networks in urban and extra-urban areas.

References

1. Borchert J R.: The twin cities urbanized areas: past present and future. Geographical Review, (1961) 51: 47-70.

2. Borruso G.: Network Density and the Delimitation of Urban Areas. Transactions in GIS, (2003), 7 (2), pp. 177 – 191.
3. Okabe, A., Yamada, I.: The K-function method on a network and its computational implementation. Geographical Analysis (2001), Vol.33, No.3, pp.271-290.
4. Gatrell A., Bailey T, Diggle P., Rowlingson B.: Spatial Point Pattern Analysis and its Application in Geographical Epidemiology. Transactions of the Institute of British Geographers (1996) 21: 256-74.
5. Bailey T. C., Gatrell A. C.:.Interactive Spatial Data Analysis. Longman Scientific & Technical, Essex (1995)
6. Holm T.: Using GIS in Mobility and Accessibility Analysis, ESRI Users Conference, San Diego, http://www.esri.com/library/userconf/proc97/proc97/to450/pap440/p440.htm (1997)
7. Matti W.: Gridsquare Network as a reference system for the analysis of small area data. Acta Geographica Lovaniensia, (1972) 10: 147-163.
8. Chainey S., Reid S and Stuart N., When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime, in Kidner D., Higgs G. and White S. (eds.), *Socio-Economic Applications of Geographic Information Science*, Innovations in GIS 9, Taylor and Francis, (2002), pp. 21-36
9. Levine Ned, CrimeStat II: A Spatial Statistics Program for the Analysis of Crime Incident Locations (version 2.0). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. May 2002.
10. Ratcliffe, J. H., McCullagh, M. J.: Hotbeds of crime and the search for spatial accuracy, Journal of Geographical Systems (1999), Volume 1 Issue 4: 385-398.

Linking Global Climate Grid Surfaces with Local Long-Term Migration Monitoring Data: Spatial Computations for the Pied Flycatcher to Assess Climate-Related Population Dynamics on a Continental Scale

Nikita Chernetsov[1] and Falk Huettmann[2]

[1] Biological Station Rybachy, Zoological Institute, Russ. Acad. Sci., Rybachy, 238535, Kaliningrad Region, Russia
nchernetsov@bioryb.koenig.ru
[2] EWHALE lab, Inst. of Arctic Biology, Biology and Wildlife Dept., Univ. of Alaska-Fairbanks (UAF), Fairbanks AK 99775 USA
fffh@uaf.edu

Abstract. Bird populations are known to be affected by climate and habitat change. Here we assess on a continental scale the relationship of a bird population index for the Pied Flycatcher (*Ficedula hypoleuca*) with spatially explicit long-term climate data. For 1971-2001 and using multiple linear regression and AIC selection methods for candidate models we found that log-transformed 30 year long-term fall bird monitoring data from Rybachy Station (Russia), Baltic Sea, can be explained by 40% with monthly mean temperatures in the West African wintering grounds; the positive relationship suggest that increasing bird numbers are explained by increasing mean November temperatures. Precipitation, European fall, spring and breeding range temperatures did not show a strong relationship, nor with bird monitoring data from two other international stations (Pape and Kabli). Our findings help to improve hypotheses to be tested in the poorly known wintering grounds. However, due to various biases care has to be taken when interpreting international long-term bird monitoring data.

1 Introduction

For several decades, there has been an increasing interest in the trends shown by bird populations, including passerines [1]. The accuracy of such estimates, especially for common species, may be questionable and therefore ornithologists brought forward indices such as from Breeding Bird Surveys (BBS), Common Bird Censuses (CBC) or standardised long-term trapping projects. In Eurasia for instance, standardized bird trapping projects are run since the 1950s, and increasingly since the 1970s, e.g. Operation Baltic, Mettnau-Reit-Illmitz-Program of Vogelwarte Radolfzell, projects run by the Biological Station Rybachy in Russia and Chokpak Ornithological Station in south Kazakhstan [2],[3],[4],[5],[6].

The aim of this paper is to investigate whether the use of long-term datasets originating from standardized trapping projects is representative of populations in time and

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 133 – 142, 2005.

space, i.e. whether the dynamics of local trapping figures are a good indicator of the overall population dynamics which were already suggested to be determined by climate [7]. There are good reasons to believe that numbers of a migratory passerine population are related, among other factors, to their breeding performance and to survival rate in winter quarters [8], [9],[10]. Due to our currently incomplete knowledge of density dependence and songbird populations [11] we do not address this subject in this study.

Population size is also likely to be linked to habitat. However, on a continental scale we are lacking habitat data with a fine resolution in time and space. The best data available for these purposes is the widely known and used HYDE and SAGE data [12], which does not indicate relevant changes between 1970 and 1990 for our study area (Huettmann and Chernetsov unpublished). Therefore, we only deal with climate data here.

In order to address the questions outlined above, we related long-term trapping data from three different stations on the eastern Baltic coast on the same flyway to large scale climate information from the presumed breeding grounds, winter quarters and migratory stopover areas in fall and spring. Although climate is known to vary spatially, so far, no true spatially explicit climate link has been made for bird monitoring data. We selected the Pied Flycatcher *Ficedula hypoleuca* - a nocturnal migrant- as our model species because of all trans-Sahara passerine migrants, the best dataset was available for this species. It is known that fall numbers for this species are consisting mostly of young birds. Using the best available scientific long-term data for birds and climates to investigate the outlined question, we assumed that if abundance indices derived from trapping data are reliable, a relationship between regional-scale climate fluctuations and bird trapping numbers should be found, since climate is reported to be a proxy for survival [13]. An important outcome of this investigation is also to learn about the magnitude of the link between climate and bird monitoring data. Data we used are coming from three long-term migration Eastern European monitoring projects (Rybachy, Pape, and Kabli); and most probably sample the same passerine populations.

2 Methods

2.1 Bird Data

We used the results of long-term migration monitoring projects run in Rybachy in the Kaliningrad Region of Russia (55º05' N, 20º44' E, Fringilla field station of the Biological Station Rybachy), in Pape (56º11' N, 21º03' E, Latvia), and in Kabli (58º01' N, 24º27' E, Estonia). All three stations use comparable, big stationary funnel traps to capture birds during the fall migratory season and allow for a powerful long-term data set. The fall trapping session at Rybachy covered the whole migratory period of the Pied Flycatcher (starting 15-20 August in the Baltic area [13], whereas trapping at Pape and Kabli in some seasons could commence after this date. Rybachy data date back to 1957, which makes it one of the longest standardized bird monitoring programs in the world. Pape started in 1967. For this study, we used the results obtained in 1971-2000 because it was the first year when the Kabli station started standardized operations, so that all three stations were run in parallel and produced comparable

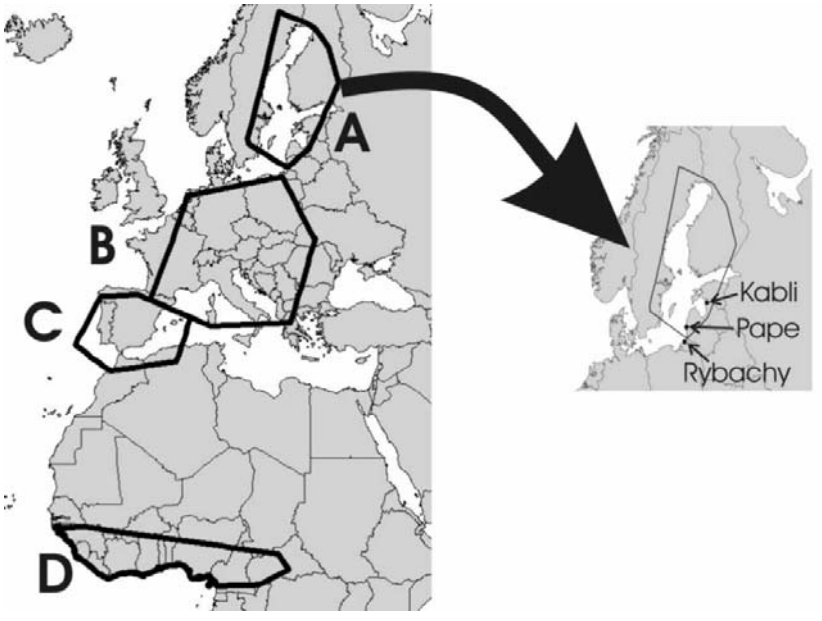


Fig. 1. Map of bird monitoring sites and spatial strata with climate data. A = Breeding Ground, B = Spring Migration, C = Fall Stop-Over, D = Wintering Ground

indices. In the analysis, annual totals were used, i.e. number of Pied Flycatchers captured during the specified fall time window. The vast majority of nocturnal passerine migrants captured at coastal banding stations in fall are hatching-year individuals (due to the so-called 'coastal effect').

2.2 Climate Data

In order to link the bird data with climate, we used the best source of spatially explicit large-scale data available. These data are provided by CRU (Climate Research Unit; CR2.0 [14]. Despite findings by [15] for spring migration we were not able to find a relevant relationship for our bird data with the North Atlantic Oscillation (NAO). NAO does not provide truly spatially explicit information, and we believe that the CRU grid data are of higher explanatory power due to their spatial-temporal resolution, and therefore contribute towards an improved inference how birds actually link with climate in space and time. We used climate data for the specific time windows and where they were of direct importance to the birds (breeding, migrating,wintering).

2.3 GIS and Spatial Approaches

We used ArcView 3.3 to plot and analyze data. Pre-processing was done in Excel. A map of the study area and locations mentioned in the text are shown in Figure 1. The probable breeding areas of Pied Flycatchers migrating along the eastern Baltic coast were identified on the basis of band recoveries of birds marked at Rybachy and controls at Rybachy of Pied Flycatchers banded elsewhere during the breeding period

(most of them were actually marked as nestlings [16]). This approach makes it possible to define an area relevant for the population sampled at the banding stations. Further, two widely applied methods have been used to estimate the breeding area: minimum convex polygon (9,360 0.5 degrees climate raster points) and 95% kernel (Animal Movement extension). These polygons were overlaid with climate data (mean monthly air temperature and precipitation for May, June and July each). The mean monthly air temperatures and precipitation values were nearly identical in breeding areas as indicated by minimum convex polygon and kernel ($r^2 > 0.99$ for temperature and $r^2 > 0.96$ for precipitation). Therefore, and for ease of interpretation, breeding area as estimated by minimum convex polygon was used in further analyses.

The potential source of bias in these breeding area estimates is that the reporting probability might be much lower in NW Russia (Leningrad Region and Russian Karelia) than in the adjacent Baltic countries and especially Finland. We however believe that if we underestimated the importance of these eastern potential recruitment areas for Pied Flycatchers migrating along the eastern Baltic coast, this bias is not very strong since only affecting a relatively small area from the overall contributing area. Therefore it is unlikely to affect our conclusions. Very few recoveries of Pied Flycatchers from African wintering grounds are available and they are still not sufficient to identify the area where Baltic Pied Flycatchers spend their winter. Therefore, we averaged the data from the whole wintering range of the species (22,754 0.5 climate raster points) as indicated by a local field guide [17]. Mean monthly air temperatures and precipitation from October throughout March (the time when birds are known to winter in Africa) were included in the analysis. Spatial stratification of the fall stop-over areas was made difficult due to insufficient information on the spatial and temporal schedule of Pied Flycatcher migration (but see [18]). We decided to analyze the data from Iberia and the adjacent parts of North-West Africa following [19] that this would be the area where Pied Flycatchers fatten up before crossing the Sahara (3,222 0.5 degrees climate raster points). We assumed that weather in August and September in this area might have an effect on the survival of Pied Flycatchers during their fall migration and thus on bird abundance during the next year, as measured by trapping data in the Baltic region. This assumption is based on the need of birds to gain considerable amount of fuel before crossing the Sahara.

It has been suggested that temperatures during spring passage, i.e. in April, could influence migratory arrival, timing of breeding and breeding performance in the Pied Flycatcher (e.g. [13], [20]). Therefore, we included April weather data averaged across much of central and southern Europe (an area estimated by us to be of importance during spring migration) to test whether they might contribute to Pied Flycatcher numbers next fall.

2.4 Statistical Analysis

For each year, we averaged the monthly climate CRU grid point data for the relevant spatial strata (breeding, stop-over and wintering grounds) in ArcView 3.3. Within the spatial strata these long-term yearly means were regressed in S-PLUS 6.2 against log-transformed long-term trapping numbers at the three stations. We log-transformed the fall catching data in order to obtain a normal distribution and a better fit. Monthly

climate data relating to each stratum were also regressed with each other to assess their correlations and independence. As a next step, linear correlations with the yearly bird monitoring data were performed. For each spatial stratum (breeding, wintering, spring and fall stopover) we selected one, or sometimes two, predictors and then performed a multiple linear regression to explain variation of bird numbers by climate. For the predictor set to be tested we followed [21] for model selection using candidate models and AIC values. AIC does not allow to be compared across data sets, and thus, our predictor sets varied. We decided to select models based on their lowest AIC values; the amount of explained variance was provided for additional information. This approach allowed us to infer from the most parsimonious models and their predictors explaining long-term bird trapping numbers with spatially explicit long-term climate data. For general comparison, we also included Null Models and computed as well the amount of variance explained by each candidate model. For Rybachy (see Table 1) and Pape we used 11 candidate models based on predictors that showed the highest partial correlations with bird data, and which were assumed to be independent judged by 5% significance levels. For Kabli we used seven candidate models. Linear regressions used these candidate models in order to explain Rybachy, Pape and Kabli bird data. For Rybachy, strongly correlated October and November temperatures competed to represent early winter temperatures, and for Pape, either June precipitation or air temperature, but not both, were included.

Table 1. Rybachy Models (T=Temperature, P=Precipitation, months abbreviated)

Candidate Model	d.f.	AIC	Variance explained, %
TNov+PMar+TSep	5	19.21	44.80
TNov+TSep	4	21.00	37.37
TNov	3	22.64	29.28
PMar+TSep	4	22.73	33.64
TOct+PMar	4	22.74	21.25
TOct+PMar+TSep	5	23.63	36.03
TSep	3	26.14	20.52
TOct+TSep	4	26.87	23.83
TNov+PMar	4	27.87	33.63
TOct	3	29.01	12.54
PMar	3	30.19	9.05
NULL Model	1	31.03	NA

3 Results

The overall set of candidate models is shown in Table 1 (Rybachy; Pape and Kabli are not shown for space reasons but follow similar structure), and the parameters of the model showing the smallest AIC values, are shown in Tables 2-4. In the Rybachy dataset, nearly 45% of variance could be explained by the predictors in the selected model which is shown in Tables 1 and 2.

Table 2. Parameters of the best model for Rybachy, as determined by AIC. Null Deviance: 4.32 on 29 degrees of freedom, Residual Deviance: 2.3879 on 26 degrees of freedom, Multiple r^2 = 0.44

Predictor	Coefficient	Std. Error	Deviance
Intercept	-7.5834	2.96	
TNov	0.2801	0.12	1.26
PMar	-0.0063	0.00	0.18
TSep	0.1092	0.04	0.48

Table 3. Parameters of the best model for Pape, as determined by AIC. Null Deviance: 11.71 on 29 degrees of freedom, Residual Deviance: 9.2134 on 27 degrees of freedom, Multiple r^2 = 0.21

Predictor	Coefficient	Std. Error	Deviance
Intercept	-0.8021	1.73	
PJun	-0.0135	0.00	1.57
TSep	0.1395	0.08	0.92

Table 4. Parameters of the best model for Kabli, as determined by AIC. Null Deviance: 6.81 on 29 degrees of freedom, Residual Deviance: 5.45 on 27 degrees of freedom, Multiple r^2 = 0.19

Predictor	Coefficient	Std. Error	Deviance
Intercept	8.9383	4.2918	
PJun	-0.0074	0.0045	0.72
TDec	-0.3081	0.1737	0.63

This model is based on the mean November temperature in the wintering area, mean March precipitation in the same area, and September temperature in the fall stopover area (September and November of the previous year). Most variance is explained by West African November temperature (Table 1: 29 % variance when used alone; see also Table 2 for strongest deviance) which however stands for overall winter months since temperature in different winter months was strongly correlated.

In the Pape and Kabli datasets, only 21% and 19 % of the variance could be explained (Tables 3, 4). We believe this is due to the fact that in some years parts of the migratory season were not surveyed adding noise to both datasets. At both stations, the model with the lowest AIC index included precipitation in June which represents weather conditions in the breeding area mainly during the nestling period.

Interestingly, the best model for Rybachy did not include weather parameters from the breeding area (Table 1). Models from Pape and Kabli included climate from the breeding grounds, but the overall models were very low in explanatory power.

4 Discussion

Our findings are in support that temperature in winter quarters would affect populations through survival and subsequent ability to produce offspring next summer (which is caught in fall at banding stations). However, using linear regression of bird trapping numbers with climate on a regional scale (entire wintering ground), we found a positive relationship with early winter temperature, and a negative relationship with late winter precipitation. This is not in full agreement with other studies suggesting that droughts in winter quarters are major contributor to declining numbers ([8], [9], [10], [22]). These mismatches are probably not due to scale, but due to species specific wintering ecology. The derived regression coefficients from Kabli and Pape are more in support of earlier findings. We think that these bird trapping data have a high variance, and subsequently the predictors have a low explanatory power, likely due to imperfect standardization. It was also suggested that the CRU precipitation data for Africa are of lower quality (D. McGuire pers.com). So far, there are no better data available to us, but are urgently needed to address climate issues for birds.

We tested three different bird datasets from the same flyway which could be explained in varying magnitude by weather predictors. Although not extremely clear defined, yet, we suggest that a relationship with wintering conditions in Africa does exist. Existing bias may be brought by (i) sampling procedure, (ii) use of indeces, (iii) relatively coarse climate data, (iv) spatial stratification, (v) use of a nocturnal migrant species and (vi) using relatively simple linear relationships. This has to be overcome in future studies. An example for (i) is presented by the late onset of trapping sessions in some years at Pape and Kabli (s. Methods section). (ii) for statistical reasons the use of indeces is not recommended (D. Anderson pers.com.). (iii): the CRU data set is currently the best globally available spatial climate information available to the research community; before improved data are published this item cannot be improved on a large scale. Examples for (iv) are the spatial uncertainties in determining accurate breeding and wintering areas; bird banding recovery data and distribution maps must be considered as being coarse and/or biased (but see [23] for statistically predicted wintering grounds). These biases might be especially large concerning African winter quarters, as we averaged weather variables for the known wintering range of the entire species. An example for (v) is that Pied Flycatchers, being nocturnal migrants, are captured in large numbers at daytime stopovers following large landfalls, i.e. when a large number of migrants cease migration. It had been indicated that day-to-day captures in funnel traps are only weakly related to the overall flow of nocturnal migrants, and there are no reasons to believe that the Pied Flycatcher forms an excep-

tion from this rule. Landfalls of birds are likely to be associated with certain meteorological conditions ([24]) which might occur more or less frequently in a particular fall season. Finally, the use of linear relationships (vi) is common in ecology and ornithology but might miss more detailed features in a regression linking complex data such as climate and populations. It is suggested to assess alternative relationships for a better understanding of the ecosystem.

Climate data from the breeding area, wintering area, and presumed fall stopover area were included into the most parsimonious models in various combinations. Prescreening excluded potential predictors that showed partial correlations with log-transformed bird numbers with $r^2 < 0.05$. All weather variables which showed stronger correlations were included into the set of candidate models. The best model for Rybachy, selected on the basis of AIC value, included all correlated variables except of October temperature which was strongly correlated with the best predictor, the November temperature. The best model for Pape did not include precipitation in November which showed a rather strong partial correlation ($r^2 = 0.11$). The best model for Kabli excluded precipitation in August (partial correlation: $r^2 = 0.11$). However, even with these variables included, this model was not parsimonious and variance explained was app. 24 % in both datasets which is much smaller than the 45 % of variance explained in the Rybachy dataset. Despite uncertainties where exactly to cut-off AIC model selection [21], our results appear to be robust.

Overall, temperature in the wintering area seems to be the most powerful predictor of bird numbers during the next fall. This could mean that the range of conditions experienced by Pied Flycatchers in the Baltic area basically allows this species to maintain the necessary level of breeding performance without major effects. It is worth noting is that the mean April temperature and precipitation in the large portion of central and southern Europe did not show strong partial correlations with any of bird datasets. The strongest relationship was between the mean April temperature and bird numbers in Rybachy but it was not significant ($r^2 = 0.090$, $p>0.10$). The local data for Kaliningrad showed a significant relationship between April temperatures and the timing of Pied Flycatcher arrival, breeding, and number of juveniles captured during the postfledging period. However, fall numbers of migrants were not related to local April temperatures ([20]). The averaged data on April temperatures from most of Europe south of Rybachy agree with the local pattern described here.

Our findings are mostly correlational, but besides the presented inference and a quantified population-climate formula, we also provide new predictions to be tested as hypotheses in the future: warm early winter temperatures in West Africa increases the survival rate for Pied Flycatchers. It is necessary to understand the underlying biological mechanism of the derived correlation. Further studies of winter ecology of Palaearctic migrants in Africa, specifically dealing with the temperature link, survival, and the exact location of wintering grounds, are also needed to understand how temperature in the wintering grounds exactly affects survival of migrants across relevant scales (global, flyway, regional and small). The Pied Flycatcher is even one of the better studied species in African winter grounds (e.g. [25]); however much more needs to be known regarding habitat and prey preference, and the distribution and abundance of both. In addition, it would be desirable to know the exact spatial links between breeding and wintering grounds.

Overall, climate seems to affect bird populations, e.g. in wintering grounds. Even with our coarse data (0.5 degrees resolution and for a nocturnal migrant), we can explain app. 45% of the variation in bird numbers in Rybachy. Still, caution is to be taken when analyzing bird trapping data; other data are supposed to be analyzed to backup findings even further. As shown here, long-term index data from a single banding station might be distorted due to a variety of factors, e.g. imperfect standardization of the banding protocol or trapping session period, or local vegetation succession over time as suggested by [26] for certain stations. This is in support of the argument that Breeding Bird Survey data, if available, could be used as a better index of passerine populations and should also be considered for the type of investigations shown here. In areas where such data do not exist, bird migration monitoring data may prove useful and efficient if well standardized and designed methods are followed.

Acknowledgements

This project was made possible by the free availability of climate data (CRU dataset); we are grateful to D.McGuire, J.Clein and M.Calef helping us with these data. Other support was provided by A.Kitaysky, J.and S.Linke. A. Leivits from Kabli and J. Baumanis from Pape kindly permitted to use their bird monitoring data. We are grateful to two reviewers, L. Strecker (FH), and all bird banders, professionals and volunteers, who worked at the three stations during the thirty years! We thank as well UAF for providing data, infrastructure and funding to carry out this collaborative research on internationally migratory birds. This is EWHALE publication No 10.

References

1. Peach, W.J., Baillie, S.R. and Balmer, D.E.: Long-term changes in the abundance of passerines in Britain and Ireland as measured by constant effort mist-netting. – Bird Study (1998) 257-275.
2. Hjort, C. and Pettersson, J.: Bird ringing figures and the monitoring of the bird populations. – Special report from Ottenby Bird Observatory (1986) N 7.
3. Busse, P. Population trends of some migrants at the southern Baltic coast – autumn catching results 1961-1990. – Ring (Proceed. Baltic Birds 1994) 115-158.
4. Berthold, P. et al. : 25-year study of the population development of Central European Songbirds: A general decline, most evident in long-distance migrants. – Naturwissenschaften (1998) 350-353.
5. Sokolov, L.V. et al. : Comparative analysis of long-term monitoring data on numbers of passerines in nine European countries in the second half of the 20th century. – Avian Ecol. Behav. (2001) 41-74.
6. Gavrilov, E.I., Gavrilov, A.E. and Kovshar, V.A.: Long-term dynamics of numbers of Swallows (*Hirundo rustica* L.) on seasonal migrations in the foothills of West Tien Shan (South Kazakhstan). – Avian Ecol. Behav. (2002) 1-22.
7. Saether, B.-E. et al.: Population dynamical consequences of climate change for a small temperate songbird. - Science (2000) 854-856.

8. Svensson, S.E.: Effects of changes in tropical environments on the North European avifauna. – Ornis fenn. (1985) 56-63.
9. Peach, W.J., Baillie, S.R. and Underhill, L.: Survival of British Sedge Warblers *Acrocephalus schoenobaenus* in relation to West African rainfall. – Ibis (1991) 300-305.
10. Jarry, G.: Incidence de plus de 25 années de désordre climatique en Afrique tropicale occidentale sur les habitats et les oiseaux migrateurs du paléarctique occidental. – Journée étude Aves 1996. - Aves (1997) 12-15.
11. Royama, T.: Analytical population dynamics. Chapman and Hall (1992).
12. Ramankutty, N., and Foley, J.A.: Estimating historical changes in global land cover: crop lands from 1700 to 1992. – Global Biogeochemical Cycles (1999) 997-1027.
13. Sokolov, L.V. : Spring ambient temperature as an important factor controlling timing of arrival, breeding, post-fledging dispersal and breeding success of Pied Flycatchers *Ficedula hypoleuca* in Eastern Baltic. – Avian Ecol. Behav. (2000) 79-104.
14. New, M., Lister, D., Hulme, M. and Makin, I.: A high-resolution data set of surface climate over global land areas. – Climate Research (2002) 1-25.
15. Hüppop, O. and Hüppop, K.: North Atlantic Oscillation and timing of spring migration in birds. – Proc. R. Soc. Lond. B (2003) 233-240.
16. Bolshakov, C.V., Shapoval, A.P. and Zelenova, N.P. :Results of bird ringing by the Biological Station "Rybachy": controls of bird ringed outside the Courish Spit in 1956-1997. Part 1. – Avian Ecol. Behav. Suppl. (2002) 1-106.
17. Borrow, N. and Demey, R.: Birds of Western Africa. – Christopher Helm (2001).
18. Bibby, C.J. and Green, R.E.: Foraging behaviour of migrant Pied Flycatchers, *Ficedula hypoleuca*, on temporary territories. - J. Anim. Ecol. (1980) 507-521.
19. Schaub, M. and Jenni, L.: Variation of fuelling rates among sites, days and individuals in migrating passerine birds. – Funct. Ecol. (2001) 584-594.
20. Sokolov, L.V.: Population dynamics in 20 sedentary and migratory passerine species of the Courish Spit on the Baltic Sea. – Avian Ecol. Behav. (1999) 23-50.
21. Anderson, D.R. and Burnham, K.P. :Avoiding pitfalls when using information-theoretic methods. – J. Wildlife Management (2002) 912-918.
22. Payevsky, V.A., Vysotsky, V.G. and Zelenova, N.P.:Extinction of a Barred Warbler *Sylvia nisoria* population: long-term monitoring, demography, and biometry. - Avian Ecol. Behav. (2003) 89-105.
23. Wisz, M.S. : Modelling potential distributions of sub-Saharan African birds: techniques and applications for conservation and macroecology. Unpublished PhD thesis, Darwin College, University of Cambridge (2004).
24. Richardson, W.J.: Timing of bird migration in relation to weather: updated review. – In: Gwinner, E. (ed.) Bird Migration: Physiology and Ecophysiology. Springer-Verlag, (1999) pp. 78-101.
25. Salewski, V., Bairlein, F. and Leisler, B.: Different wintering strategies of two Palearctic migrants in West Africa – a consequence of foraging strategies? – Ibis (2002) 85-93.
26. Sokolov, L.V. et al. : Comparative analysis of long-term monitoring data on numbers of passerines in nine European countries in the second half of the 20th century. – Avian Ecol. Behav. (2001) 41-74.

Classifying Internet Traffic Using Linear Regression

T.D. Mackay and R.G.V. Baker

School of Human and Environmental Studies,
University of New England,
ARMIDALE, NSW 2351, Australia
rbaker1@une.edu.au

Abstract. A globally weighted regression technique is used to classify 32 monitoring sites pinging data packets to 513 unique remote hosts. A statistic is developed relative to the line of best fit for a 360^o manifold, measuring either global or local phase correlation for any given monitoring site in this network. The global slope of the regression line for the variables, phase and longitude, is standardised to unity to account for the Earth's rotation. Monitoring sites with a high global phase correlation are well connected, with the observed congestion occurring at the remote host. Conversely, sites with a high local phase correlation are poorly connected and are dominated by local congestion. These 32 monitoring sites can be classified either globally or regionally by a phase statistic ranging from zero to unity. This can provide a proxy for measuring the monitoring site's network capacity in dealing with periods of peak demand. The research suggests that the scale of spatial interaction is one factor to consider in determining whether to use globally or locally weighted regression, since beyond one thousand kilometres, random noise makes locally weighted regression problematic.

1 Introduction

The Internet forms the physical network of connectivity (such as, optical cables and phone wires) where there are nodes or 'routers' that navigate packets of data from one computer to another [3, 4, 5]. The time taken to transmit this data between computers is termed 'latency'. Packet loss is a proxy for peak demand, since in heavy traffic periods, too many packets arrive and the routers then hold them in buffers until the traffic decreases. If these buffers fill during these times of congestion, the routers drop packets and this is measured as 'packet loss'. Sites that are not well connected in the network suffer high 'packet loss'. Internet traffic within and between time zones can be described by a set of equations for the exchange rate $E(x,t)$, continuity conditions for monitoring sites (origins) and remote hosts (destinations) and traffic flow as [1]:

$$E(x,t) = -\frac{1}{M}\frac{\partial \phi_0}{\partial t} + \eta\phi_0 \tag{1}$$

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 143–151, 2005.

$$\frac{\partial \phi_0}{\partial x} = -\frac{\partial E}{\partial t} \tag{2}$$

$$\frac{\partial \phi_0}{\partial x} = \frac{1}{M}\frac{\partial^2 \phi_0}{\partial t^2} - \eta \frac{\partial \phi_0}{\partial t} \ . \tag{3}$$

where $\phi_0(x,t)$ is the site density for a time distance p between pairs assumed to be equal for remote hosts and the i^{th} monitoring site (located at x and having meridians of longitude as time bands between $t_i = ip$); M is the transfer capacity of the network infrastructure; and η as a deviation constant from interference generated by demand circumnavigating the Earth every 24-hours. The ping latency Δt and the partitioning of distance Δx are constrained by the inequality [2]:

$$2M\Delta x \leq \Delta t^2 \ . \tag{4}$$

This inequality states that there will never be instantaneous Internet interaction and the 'death of distance' hypothesis is void. These equations summarise Internet interaction characterised by a demand wave of traffic, with regional and global contributions, circumnavigating the Earth every 24-hours.

The determination of the source of packet loss (and hence congestion) from regional traffic from the same time zone, or from global congestion from other time zones (both forwards and backwards, relative to Greenwich Mean Time) comes from considering the time operator components of Eq. 3. The second order derivative $\frac{\partial^2}{\partial t^2}$ calibrated to the first order spatial operator $\frac{\partial}{\partial x}$ yields time gaussians across the same time zone. The second time derivative $\frac{\partial}{\partial t}$ describes the deviation in the wave from congestion originating from different time zones.

What is then required is a statistic to measure the phase correlation of origin-destination pairs with the local time zone and the remote time zones. This will be shown using the year 2000 data set compiled by SLAC. The method used was a linear regression plot of the phase of the wave (for 5% periodicity) against longitude. This had to be constructed for points distributed relative to a 360^o manifold with a boundary. The slope of the regression line was then calibrated for the global phase of unity from the rotation of the Earth and a slope of zero for congestion in phase with the local time zone.

2 Classifying Internet Sites Using Linear Least Squares Regression

Let θ and ϕ be longitude and phase respectively. Let (θ_i, ϕ_i) be the longitude and phase of the i^{th} data point. Let $\phi_{\mathrm{G}}(\theta)$ and $\phi_{\mathrm{L}}(\theta)$ be the linear least squares regression of the n data points (θ_i, ϕ_i) where $i \in \{1...n\}$ such that:

$$\phi_{\mathrm{G}}(\theta) = \theta + \delta_{\mathrm{G}} \tag{5}$$

$$\phi_{\mathrm{L}}(\theta) = \delta_{\mathrm{L}} \ . \tag{6}$$

For some $\delta_{\mathrm{G}}, \delta_{\mathrm{L}} \in [-\pi, \pi] \subset \Re$. Since we must have $\frac{d\phi}{d\theta} \in \{0, 1\}$ to satisfy sensible boundary conditions of continuity over the 24 hour boundary.

We wish to find δ_{G} and δ_{L} such as to minimise the normalised sum of the squares of the residuals R_{G} and R_{L}.

Where:

$$R_{\mathrm{G}} = \frac{1}{n}\sum_{i=1}^{n} \min\{(\phi_{\mathrm{G}}(\theta_i) - \phi_i)^2, (|\phi_{\mathrm{G}}(\theta_i) - \phi_i| - 2\pi)^2\} \,, \tag{7}$$

$$R_{\mathrm{L}} = \frac{1}{n}\sum_{i=1}^{n} \min\{(\phi_{\mathrm{L}}(\theta_i) - \phi_i)^2, (|\phi_{\mathrm{L}}(\theta_i) - \phi_i| - 2\pi)^2\} \,. \tag{8}$$

We also wish to scale R_{G} and R_{L} so as to produce a useful statistic for comparison. To this end we multiply by a scale factor so that R_{G} and R_{L} take values on the interval $[0, 1] \subset \Re$. The maximum sum of squares of angular residuals $R_{\max}$ occurs when the data points are uniformly distributed in the direction perpendicular to the regression line. So that residuals r are uniformly distributed on the interval $[-\pi, \pi] \subset \Re$. Thus

$$\begin{aligned} R_{\max} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} r^2 dr \\ &= \frac{\pi^2}{3} \,. \end{aligned} \tag{9}$$

The general statistic is therefore defined as:

$$\begin{aligned} \chi^2 &= \frac{R}{R_{\max}} \\ &= \frac{3R}{\pi^2} \,. \end{aligned} \tag{10}$$

This statistic can either measure global or local connectivity. If $\chi_{\mathrm{regional}} \to 0$, then much of the congestion comes from periodic traffic generated in the local time zone. It is not well-connected regionally or globally and the deviations usually follow a horizontal line. Sites where $\chi_{\mathrm{global}} \to 0$, show good connectivity with the observed packet loss caused by periodic congestion in the remote time zone. The points line up closely to the standardised global phase line with a slope of unity. This statistic gives some proxy of the site connectivity and its ability in dealing with regional and global traffic relative to its network capacity.

3 Methodology

The Stanford Internet experiments have been running from 1998 to 2004 with various monitoring and remote hosts distributed in a global network `http://www-iepm.slac.stanford.edu/pinger/`The year 2000 had the greatest connectivity between a network of sites and this year presents the best opportunity to develop the statistic. It featured 32 global monitoring sites and some 513 remote hosts, mainly distributed between the US and Europe (Fig. 2). These monitoring sites pinged this network every hour and Stanford Linear Accelerator

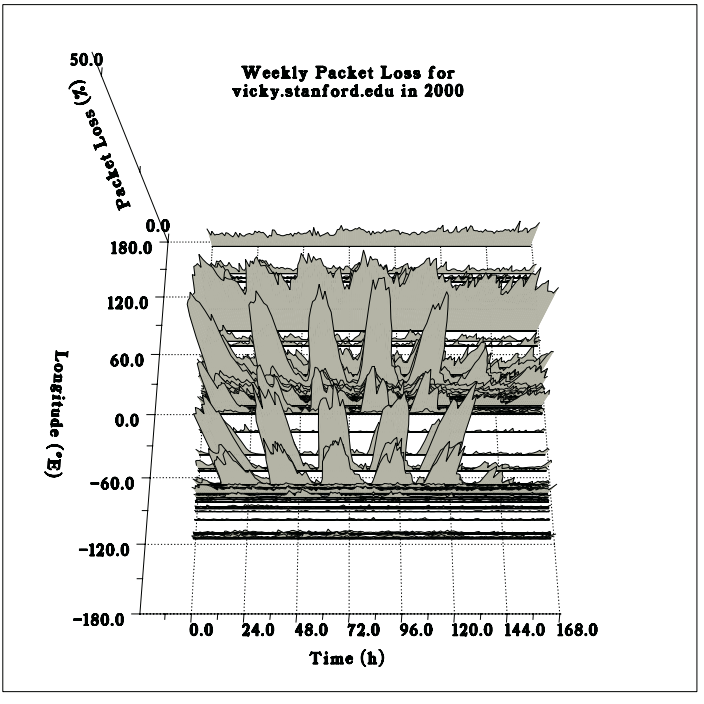
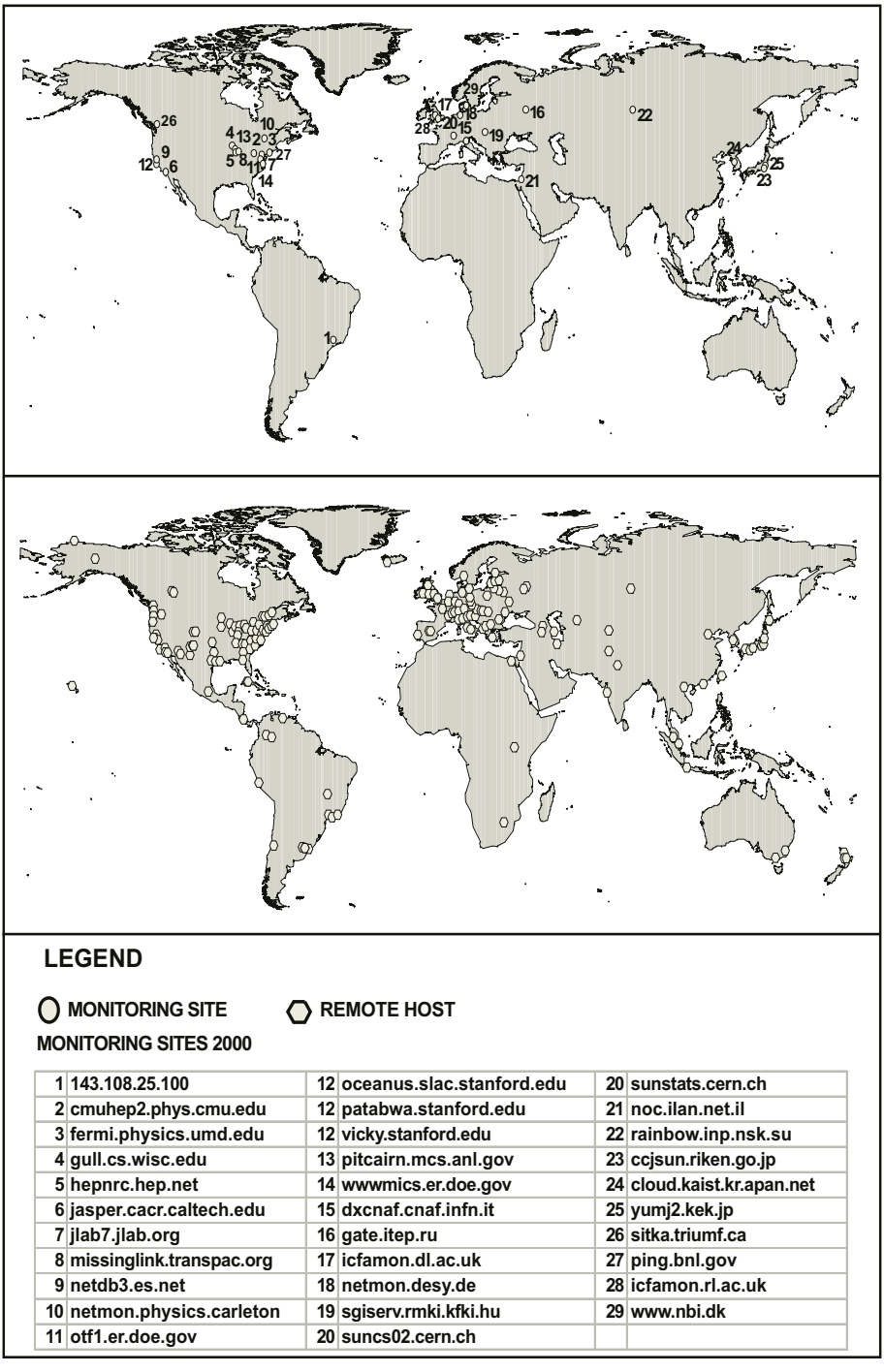


Fig. 1. The congestion wave for the `vicky.stanford.edu` monitoring site, 2000

Centre (SLAC) measured the ping latency and packet loss between each origin-destination pairs. The raw hourly packet loss data was collated for a given period and averages performed for each hour. This was plotted for the week (per hour) for monitoring site-remote host pairs (Monday 00:00 to Sunday 24:00 GMT).

The packet loss can observed to follow a periodicity either from regional or global traffic (Fig. 1). This packet loss is a proxy for peak demand and the capacity of the network to deal with congestion. The aim is to use a global and regional statistic to identify the capacity of each monitoring site in dealing with the traffic. However, the periodic connections for each monitoring site and its remote hosts need first to be identified. A Fourier analysis was performed for each pair (Monday 00:00 to Friday 24:00 remote host local time, as weekends show much less congestion) to determine periodicity and phase of the packet loss data. Each pair is filtered for periodicity of more than 5% in the 24-hour component. This subset is tested for regional and global phase correlation by globally weighted regressions. An algorithm then decided which sites were dominated by locally periodic packet loss (horizontal lines standardised with zero slope for the local time zone) or globally periodic packet loss from remote hosts in different time zones (standardised with a slope of unity relative to the rotation of the Earth). Those monitoring sites that were well-connected suffered little packet loss due to local congestion and were defined by a χ_{global} statistic approaching zero.



1	143.108.25.100	12	oceanus.slac.stanford.edu	20	sunstats.cern.ch
2	cmuhep2.phys.cmu.edu	12	patabwa.stanford.edu	21	noc.ilan.net.il
3	fermi.physics.umd.edu	12	vicky.stanford.edu	22	rainbow.inp.nsk.su
4	gull.cs.wisc.edu	13	pitcairn.mcs.anl.gov	23	ccjsun.riken.go.jp
5	hepnrc.hep.net	14	wwwmics.er.doe.gov	24	cloud.kaist.kr.apan.net
6	jasper.cacr.caltech.edu	15	dxcnaf.cnaf.infn.it	25	yumj2.kek.jp
7	jlab7.jlab.org	16	gate.itep.ru	26	sitka.triumf.ca
8	missinglink.transpac.org	17	icfamon.dl.ac.uk	27	ping.bnl.gov
9	netdb3.es.net	18	netmon.desy.de	28	icfamon.rl.ac.uk
10	netmon.physics.carleton	19	sgiserv.rmki.kfki.hu	29	www.nbi.dk
11	otf1.er.doe.gov	20	suncs02.cern.ch		

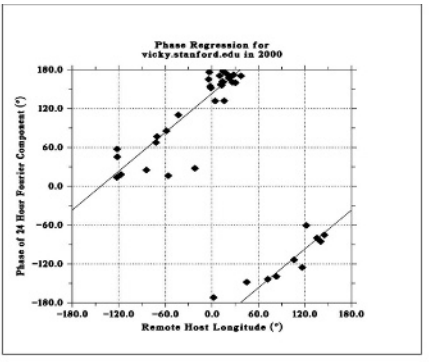
Fig. 2. Location of monitoring sites and remote hosts in the 2000 data set

4 Results

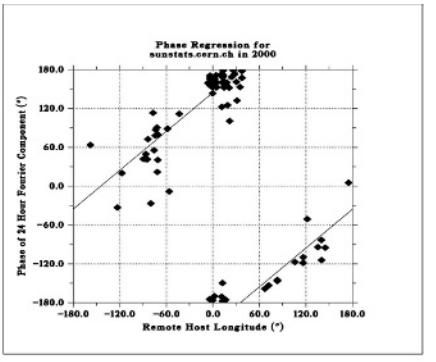
The calculations of the indexes are presented in Table 1. The US monitoring sites, in general, had low χ_{global} statistics, suggesting that the network had the capacity to deal, not only with local traffic, but peak demand generated from different time zones across Europe and Asia. For example, `vicky.stanford.edu` in San Francisco had a χ_{global} value of 0.059 and the regression plot, standardised to the Earth's rotation, showed most pairs periodic and in phase with the rotation of the Earth (Fig. 3(a)). The majority of the packet loss was from remote hosts

Table 1. Monitoring sites showing latitude and longitude, periodicity, phase, phase coordinate, regional and global phase statistics

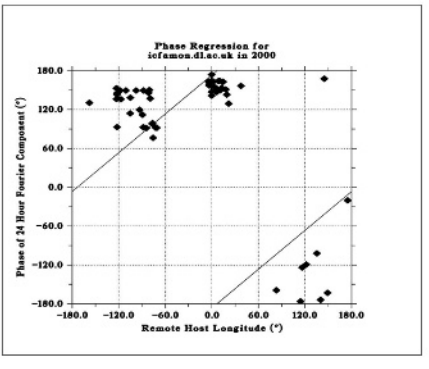
Monitoring Site	Latitude	Longitude	No	Periodic (%) (at 5%)		Phase Intercept	χ_{global}	Phase Intercept	χ_{regional}
143.108.25.100	-22	-46	73	70	96	152.9	0.316	112.5	0.086
ccjsun.riken.go.jp	35.7	139.767	99	61	62	150.0	0.424	-135.3	0.432
cloud.kaist.kr.apan.net	37.55	126.967	96	77	80	137.1	0.695	-78.4	0.172
cmuhep2.phys.cmu.edu	40.437	-80	47	40	85	142.1	0.134	79.5	0.576
dxcnaf.cnaf.infn.it	44.5	11.33	66	49	74	-167.8	0.421	174.7	0.106
fermi.physics.umd.edu	38.99	-76.93	42	25	60	141.8	0.068	142.0	0.533
gate.itep.ru	55	37	73	65	89	-169.1	0.502	169.7	0.042
gull.cs.wisc.edu	43.073	-89.401	60	38	63	149.4	0.283	106.0	0.575
hepnrc.hep.net	41.85	-88.31	175	90	51	127.2	0.184	113.5	0.564
icfamon.dl.ac.uk	53	-2	65	63	97	173.4	0.318	146.2	0.144
icfamon.rl.ac.uk	51.75	-1.25	63	57	90	169.8	0.265	145.3	0.297
jasper.cacr.caltech.edu	34.147	-118.133	68	43	63	128.8	0.161	141.2	0.450
jlab7.jlab.org	37	-76	217	134	62	136.8	0.183	125.8	0.360
missinglink.transpac.org	41	-87	49	30	61	154.5	0.074	173.4	0.521
netdb3.es.net	38	-122	74	47	64	134.5	0.186	140.0	0.587
netmon.desy.de	53	9	60	58	97	-173.8	0.425	178.0	0.100
netmon.physics.carleton.ca	45.42	-75.7	49	46	94	137.0	0.216	67.1	0.340
noc.ilan.net.il	31.783	35.233	49	36	73	170.7	0.222	163.5	0.220
oceanus.slac.stanford.edu	37.41	-122.2	224	152	68	131.3	0.223	108.5	0.445
otf1.er.doe.gov	38.9	-77.04	55	43	78	174.2	0.476	107.4	0.180
patabwa.stanford.edu	37.442	-122.15	85	39	46	146.5	0.063	131.8	0.528
ping.bnl.gov	40.76	-72.9	112	72	64	138.7	0.156	122.8	0.433
pitcairn.mcs.anl.gov	41.858	-88.017	222	182	82	131.4	0.161	84.9	0.483
rainbow.inp.nsk.su	55.06	83.08	58	45	78	-132.2	0.620	-145.8	0.094
sgiserv.rmki.kfki.hu	47.43	19.25	48	42	88	-159.3	0.550	167.5	0.090
sitka.triumf.ca	49.25	-123.08	45	32	71	126.5	0.187	-65.7	0.720
suncs02.cern.ch	46.23	6.06	25	2	08	147.5	0.023	-175.5	0.034
sunstats.cern.ch	46.23	6.06	147	85	58	144.5	0.078	144.0	0.444
vicky.stanford.edu	37.442	-122.15	87	41	47	143.3	0.059	152.2	0.523
www.nbi.dk	55	11	45	12	27	157.8	0.151	-174.7	0.324
wwwmics.er.doe.gov	38.9	-77.04	72	57	79	159.0	0.229	124.0	0.248
yumj2.kek.jp	36.08	140.25	51	51	100	-109.2	0.797	-118.3	0.143
Avg						158.0			



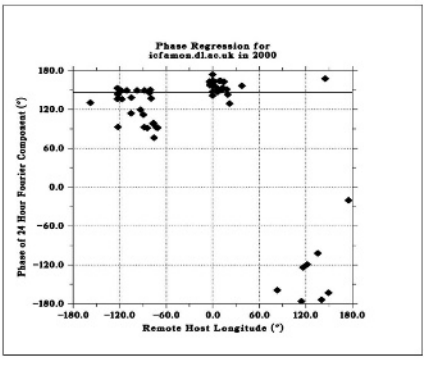
(a) $\chi_{\text{global}} = 0.059$ for the `vicky.stanford.edu` monitoring site



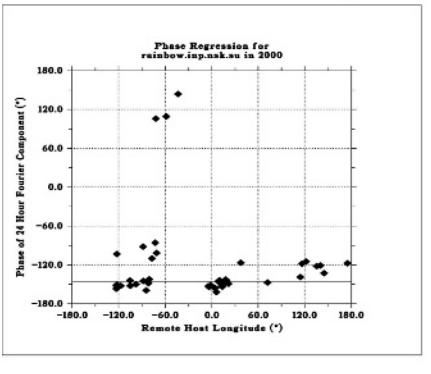
(b) $\chi_{\text{global}} = 0.078$ for the `sunstats.cern.ch` monitoring site



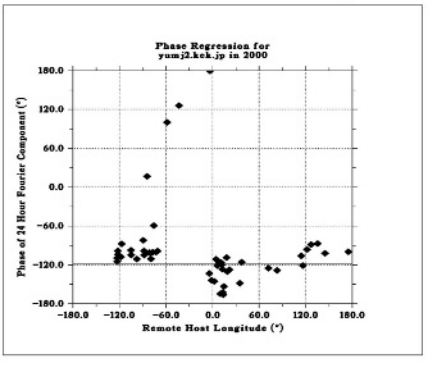
(c) $\chi_{\text{global}} = 0.318$ for the `icfamon.dl.ac.uk` monitoring site



(d) $\chi_{\text{regional}} = 0.144$ for the `icfamon.dl.ac.uk` monitoring site



(e) $\chi_{\text{regional}} = 0.094$ for the `rainbow.inp.nsk.su` monitoring site



(f) $\chi_{\text{regional}} = 0.143$ for the `yumj2.kek.jp` monitoring site

Fig. 3. The distribution of periodic origin-destination pairs relative to global or local regression lines

outside the US, and in phase with the remote time zone. The value of the χ_{regional} statistic was 0.523 showing little correlation with the local time zone of the monitoring site. Note that the global and local statistics are not complementary, because the estimation is dependent on the clusters of pairs relative to the global and regional regression lines.

For Europe, some sites were similar to the `vicky.stanford.edu` site. For example, the `sunstats.cern.ch` site in Geneva, Switzerland had a χ_{global} value of 0.078 and a χ_{regional} value of 0.444 (Fig. 3(b)). However, others struggled and appeared to be dominated by packet loss from regional traffic. For example, the `icfamon.dl.ac.uk` site in Liverpool, UK had a χ_{global} value of 0.318 and a χ_{regional} value of 0.144. There is a noticeable cluster of periodic pairs of packet loss from the local time zone (Fig. 3(c)). Indeed, the line of best fit is most appropriate for the horizontal phase line with a 145.3^o intercept (Fig. 3(d)). The statistics identify regional capacity problems. This situation in the UK had been previously reported by [6], where link speeds there contributed a longer latency than in France. All UK hosts with hop counts of 10 had a longer delay than other hosts with hop-counts of over 20 in France. The global and regional statistics above support the identification of network problems in the UK and the benefit of the above statistic is that any improvements can be gauged in the [0,1] value range.

For Asia, there were network problems identified in the monitoring sites. For example, the `rainbow.inp.nsk.su` site in Novosibirsk in Kazakhstan had a χ_{global} value of 0.620 and a χ_{regional} value of 0.090 (Fig. 3(e)). It had substantial periodic packet loss from local regional traffic. The `yumj2.kek.jp` site in Japan had a similar problem with a χ_{global} value of 0.797 and a χ_{regional} value of 0.143 (Fig. 3(f)). This could be a result of an 'island effect' from the monitoring site being surrounded by the Pacific Ocean and consequently a lack of network development. It is a possibility that such sites could benefit from a locally weighted regression to see whether the poor connectivity is an artefact of geography rather than network investment.

5 Conclusion

A new application of globally weighted regression has been presented that has been applied to classifying 32 monitoring sites pinging data packets within a network of remote hosts. The congestion proxy is packet loss and the variables used are phase and longitude for periodic pairs. Sites can be viewed either globally against a regression line with the slope standardised to unity or regionally where the phase line is horizontal (with a slope of zero). In doing so, significance testing has been sacrificed for the physical constraint of the Earth's rotation. The global line looks at peak demand contributions from different time zones. Those monitoring sites that are well-connected are in phase with the Earth's rotation and suffer minimal local interference. Those sites that are not well-connected have problems navigating periods of local peak demand due to regional traffic from the same time zone. The consequence is deviations relative to a horizontal

phase line for the particular time zone. These points can be regressed against a regional line of best fit which is horizontal and an index similarly calculated. The US monitoring sites are generally well-connected with small global deviations ($\chi_{\text{global}} \to 0$ and $\chi_{\text{regional}} \to 1$). The global and regional statistics are not complementary because the values depend on the degree of clustering of points relative to the global and regional regression phase lines. European monitoring sites are a mixture of connectivity and ability to handle peak loads with the UK identified as having problems in 2000. Asia had the worst connectivity for the monitoring sites. It is surprising that the Japanese site struggled with the global phase, but this could be a problem with the global weighting statistics rather than network capacity and this is further complicated by the geography of the country and its isolation. Perhaps in this case, a locally weighted regression would be more appropriate. For the global phase regression, where contributions are sourced from different time zones some thousands of kilometres away, the use of a locally weighted regression would be problematic because of random noise in the system.

Acknowledgments

We would like to thank Brett Carson and Rajanathan Rajaratnam for their data reduction and cartographic services. This project was funded by Australian Research Council Discovery Grant DP0343265.

References

1. Baker R.G.V. (2001) Modelling Internet transactions as a time-dependent random walk: an application of the RASTT model. Geojournal, 53, 417-418 and reprinted in Boots B., Okabe A., & Thomas R. (Eds) (2003) Modelling Geographical Systems: Statistical and Computational Applications. Ch.15. 295-316, Kluwer, Amsterdam (2003).
2. Baker R.G.V. (2005) The gravity inequality and Internet traffic. Journal of Geographical Systems, 7, (forthcoming)
3. Cohen R., Erez K., Be-Avraham D., & Havlin. (2000). S. Resilience of the Internet to random breakdown, Physical Review Letters, 85, 4626-4628.
4. Padmanabhan V.N. & Subramanian L. (2001) An investigation of geographic mapping techniques for Internet hosts. Proceedings of ACM SIGCOMM, San Diego, USA.
5. Toroczkai Z. and Bassler K.E. (2004) Network dynamics: jamming is limited to scale-free systems. Nature, 428, 716.
6. Wang Y., Lai P., and Sui D. (2003) Mapping the Internet using GIS: The death of distance hypothesis. Journal of Geographical Systems, 5, 381-405.

Modeling Sage Grouse: Progressive Computational Methods for Linking a Complex Set of Local, Digital Biodiversity and Habitat Data Towards Global Conservation Statements and Decision-Making Systems

Anthonia Onyeahialam[1], Falk Huettmann[2], and Stefania Bertazzon[1]

[1] Department of Geography, University of Calgary, Calgary Canada
t_onyeahialam@yahoo.com, bertazzs@ucalgary.ca
[2] EWHALE lab, University of Alaska-Fairbanks, Fairbanks, USA
fffh@uaf.edu

Abstract. Modern conservation management needs to link biological questions with computational approaches. As a global template, here we present such an approach from a local study on sage grouse breeding habitat, leks, in North Natrona County, Wyoming, using remote sensing imagery, digital datasets, spatial statistics, predictive modelling and a Geographic Information System (GIS). Four quantitative models that describe sage grouse breeding habitat selection were developed for multiple scales using logistic regression and multivariate adaptive regression splines (MARS-Salford Systems). Based on candidate models and AIC, important habitat predictor variables were elevation, distance to human development, slope, distance to roads, NDVI and distance to water, but not Sagebrush. Some predictors changed when using different scales and MARS. For the year 2011, a cumulative prediction index approach is presented on how the population viability of sage grouse can be assessed over time and space using Markov chain models for deriving future landscape scenarios and MARS for species predictions.

1 Introduction

Complex computations and advanced statistics play an important role for our daily lives. Biodiversity and habitats contribute to human well-being, but as well to economics and wealth [1]. Often, it is not clear to the general public that these complex subjects are linked, and how computing, quantitative methods, biodiversity and habitat data sets, biology and geography are connected towards a sustainable future [2], [3], [4]. We use the sage grouse (*Centrocerus urophasianius*) – a species of North American conservation concern - as an example how these research disciplines, with each method used here being at the forefront of the individual research discipline, can get merged at a local scale. As an outlook we show how such an approach can get applied globally. We believe that the methods presented here are of relevance to global management issues of biodiversity, wildlife and habitats, as well as to the well-being of humans.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 152 – 161, 2005.

2 Methods

Making informed decisions on the conservation of sage grouse requires the full understanding of its life history characteristics, preferences, limiting factors as well as its dependence on sagebrush vegetation. It also requires an evidence and an understanding of its habitat selection patterns and any observed underlying processes. This will aid in the identification and modelling of its habitats and how landscape change can affect this species with the goal of providing adequate and timely information for the species management. Approximately 736,000 km^2 of sagebrush vegetation types existed in North America [5], making it one of the most widespread habitats in North America. Unfortunately, much of this habitat has been lost or degraded over the last 100 years [6].

2.1 Biology Data

Sage grouse is an endangered species in North America. It can be found in sage (*Artemisia spp*) habitat, which usually is rangeland used by cattle but which can be increasingly diminished through other land uses such as real estate, industrial activities and road construction. Sage grouse mate at lek sites, which are crucial habitats for their reproduction; nests are usually found related to lek sites [7],[8],[9], [10]. For the study area, yearly lek site surveys were carried out during the years 1979-2001, resulting into a maximum of 11 detected and subsequently geo-referenced leks.

2.2 Remote Sensing and Other Habitat Data

A Landsat TM image was obtained 1985 for the study area. A second image, ETM+ Landsat 7, was obtained for summer July 2001. This image was processed and classified using approaches described in [11], [12], and resulting into an overall accuracy of 85%. Computations for deriving NDVI and Greenness were applied. Other data such as DEM (including derived slope and aspect), roads and water features were also available to the project (see [11] for earlier work).

2.3 Spatial Statistical Analysis

All data were imported into the Geographic Information System (GIS) ArcView 3.3. and ArcGIS 8.2 for further processing and analysis.
Dynamic population level processes with spatial characteristics can create patterns relating locational attributes to them. These spatial patterns are the result of underlying processes and these processes can be described, measured, and evaluated using various spatial descriptive statistical methods. The observed patterns can then be related back to the ecology of the phenomenon in question and be used for predictive purposes.Point pattern analysis and spatial autocorrelation [13],[14],[15], were used to explore the distribution of observed sage grouse breeding habitats in North Natrona County, Wyoming. We examined if an underlying pattern exists in sage grouse breeding habitats.

The nearest neighbourhood index was estimated for describing the sage grouse breeding habitat patterns using a nearest neighbourhood index script in the Arcview

3.2 wGIS software. The analysis is based on the calculated average of all distances between each pair of points representing sage grouse breeding habitats using Pythagoras theorem. For the purposes of comparison, the calculated average nearest neighbourhood distance is compared to an expected average distance between nearest neighbours. The expected average distance between nearest neighbours is standardised to account for area coverage in a random point pattern.

According to [14] autocorrelation is a general statistical property of ecological variations observed across geographic space. [15] define spatial autocorrelation to measure the degree to which a spatial phenomenon is correlated to itself. Moran I and Geary's C indices are types of spatial autocorrelation measures. Moran's I represents the overall agglomerative patterns (i.e. are events clumped or dispersed ?) whereas Geary's C explains the similarity or dissimilarity of events [14]. The spatial autocorrelation of sage grouse breeding habitats was estimated using the above-mentioned metrics provided by the S-PLUS extension in the Arcview 3.2 GIS software.

2.4 Modeling Habitat Relationships

Based on known lek sites, we created 'presence/absence' locations. Following established methodology this was done using pseudo-absences, random locations. As outlined in [2] first we used Generalized Linear Models (GLM) to build a model [16],[17]. For predictors we choose Land use, Distance to water, Distance to roads, Distance to human development, NDVI, Greenness, Elevation, Slope and Aspect.

Model selection was done with a set of 83 candidate models and AIC; this method follows [18], [19] and was based on a modified S-PLUS code by [20].

Non-linear models are powerful inference and prediction tools. They provide great alternatives, and often improvements, when compared with GLMs. We decided to use the MARS-Salford (Multivariate Adaptive Regression Splines) algorithm due its speed, convenience and general accuracy [21]; similar algorithms such as Cart (Classification and Regression Trees), Neural Networks and TreeNet could also have been used [22], [23], [24].

We build models on two scales since it is known that the scale of study can affect the inference [25] and prediction. We choose the point scale, as well as the home range scale for each lek site in order to assess the scale effects. Home range size was determined from the literature and centered on lek sites.

2.5 Future Landscapes and Sage Grouse Predictions

Using the Markov [26], [27], [28] module in the IDRISI 32 software, multiple iterations were run on land use maps classified from Landsat 1985 and Landsat 2001 imageries in order to predict land cover changes for the next ten years, 2011 (see also [29] for other approaches to obtain cumulative effects landscape scenarios of the future). From the predicted distribution map of lek sites 2001 we computed a cumulative prediction index of lek site occurrence in the study area. This cumulative index for the study area was equaled with the number of known breeding sites. We used MARS since it allows for a convenient, fast and reliable modeling of the future distribution of sage grouse. Using the future landscape of 2011 as input into the 2001 MARS model, we then predicted lek sites spatially. Finally, we computed the cumula-

tive prediction index for 2011, as an indicator of how many breeding sites would exist in the future in the study area for the future landscape.

3 Results

3.1 Landuse Classification

Figure 1 shows a reclassified result of a hierarchical image classification approach. Overall, thirty seven vegetation and land use classes exist. The results of the reclassification with 11 classes were used for simplicity in the following model development.

3.2 Spatial Patterns

The results of the nearest neighbourhood and quadrat analysis, Geary's C and Moran's I indices indicate that the location of sage grouse breeding habitat followed distinct patterns. Thus, sage grouse does not seem to select the location of its breeding habitat randomly, but certain characteristics that are favourable to its breeding habitat are considered at every location. The results of the exploratory process set the stage for further analysis using other statistical techniques for modelling and predicting the state of its present and future habitats.

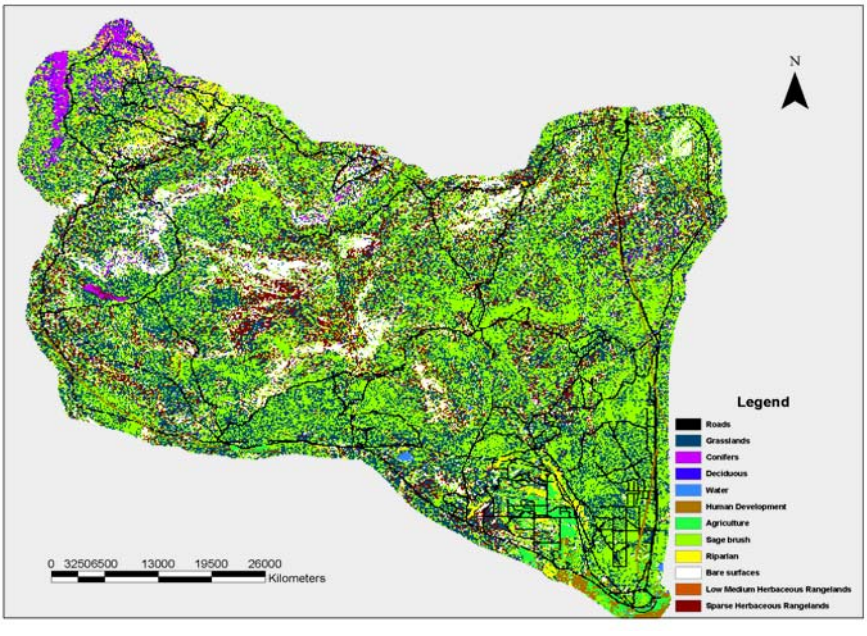


Fig. 1. Landcover classification from using 11 vegetation classes [7]

3.3 Spatial Prediction

For the GLMs, from the 83 candidate models we identified a final model which was within 5 ±AIC units. This was done for two scales, lek site and home ranges. Based on expert knowledge, we identified the homerange model as more accurate due to better predictive accuracies and higher variances explained. Model coefficients are presented in Equation 1.

$$\ln(1/1\text{-}p) = -3.70503 - 0.56858\,(\text{Slope}) + 0.00007\,(\text{Dist. from Human development}) + 0.0030287\,(\text{Elevation}) + 0.0000101\,(\text{Dist. to water}) - 0.000901\,(\text{Dist. to Roads}) - 4.96345\,(\text{NDVI})$$ **Equation 1**

When using MARS on the home range scale, only three predictors are selected (Table 1). However, it should be remembered that same as with the GLM, these importance values are driven by the characteristics of the algorithm and GIS, but not necessarily by the true biological needs of the sage grouse [2].

Table 1. Importance of predictor variables in MARS (Home range scale)

Variable	Cost of Omission	Importance	Scale of Importance
Slope	0.078	100.000	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Distance to Human Develop.	0.077	82.111	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Distance to Roads	0.074	37.210	IIIIIIIIIIIIIIIIII
Distance to water	0.074	0.000	
NDVI	0.074	0.000	
Vegetation Classes	0.074	0.000	
Greenness	0.074	0.000	
Elevation	0.074	0.000	
Aspect	0.074	0.000	

Sagebrush vegetation has been reported to be prominent in the life of the sage grouse. However, the results of the most appropriate model developed for predicting sage grouse breeding habitats using logistic regression as well as MARS and on multiple scales show that sage brush vegetation as such is not the only and key predictor variable that determined the location of sage grouse breeding habitats in North Natrona County. Terrain characteristics, proximity to sources of water, roads and human development play prominent roles in determining the occurrence of sage grouse breeding habitats in the area. This provides new insights into habitat features relevant for lek sites, and likely for related nest sites. NDVI, which describes the richness and vigour of the vegetation, was the only predictor variable in the model on multiple scales from logistic regression that gave information about the characteristics of the vegetation cover. Candidate models developed with logistic regression where sagebrush vegetation played a role as a predictor variable ranked between the 10^{th} and 20^{th} models across scales. This suggests that other identified habitat predictor variables should be more important in predicting the species habitat in North Natrona or other areas where sage brush is the major land cover type.

Applying the above home range formula with MARS for a prediction of lek sites to the overall study area for the year 2001 indicates that 28 of these breeding sites would exist.

3.4 Landcover Change and Future Scenario

Figure 2 represents the future distribution of lek sites in the landscape of North Natrona County, Wyoming, for the year 2011, based on the combined Cellular automata / Markov chain land cover prediction procedure. The results indicate a more fragmented landscape scenario for 2011, and the extent and distribution of sagebrush and herbaceous rangeland vegetation is lower as opposed to human development (table 2; map not shown here). Without ground-truthing information, the reliability of this classification cannot be statistically ascertained but likely presents the general trend.

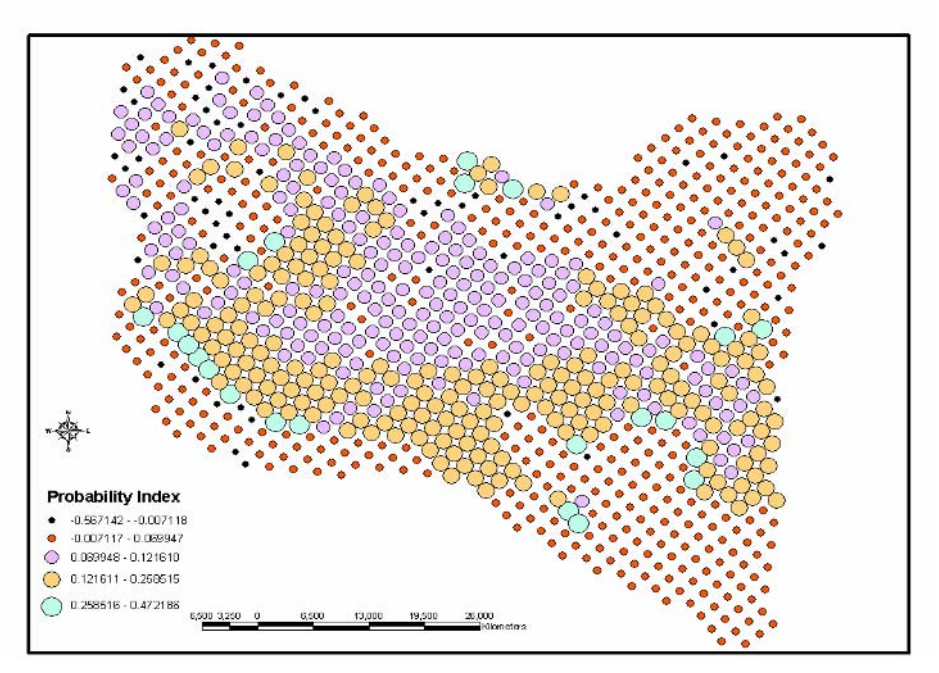


Fig. 2. Predicted future distribution of sage grouse lek sites for 2011

For the year 2011 a viability of the sage grouse population was estimated (compare [30]). Results suggest that an increase in the cumulative predicted probability index was experienced between 2001 (28 predicted lek sites) and 2011, increasing from 73.79 to 79.47. At a cumulative predicted index value of 79.47, an estimated value of 30 sage grouse breeding sites can be supported. However, the spatial distribution of these predicted sites in Figure 2 shows that they are mostly concentrated in the central portion of the study area, which coincides with predicted vegetation types such as sagebrush and rangelands reported to be favoured by sage grouse. Such a distribution might facilitate contact among individuals but also increases competition; it can pose

the threat that all relevant areas are concentrated into one single area making populations vulnerable if this one habitat would be majorly disturbed.

Table 2. Changes in the Class Area Landscape Metric of Vegetation and Land Use Classes from 2001 to 2011

Vegetation and Land Use Classes	Class Area (2001) (ha)	Class Area (2011) (ha)	Change (ha)
Roads	20522.97	23719.05	3196.08
Grasslands	115747.1	68422.86	-47324.3
Conifers	6773.31	29805.75	23032.44
Deciduous	4166.28	18944.55	14778.27
Herbaceous Rangelands	40155.57	25304.13	-14851.4
Water	399.42	36790.65	36391.23
Human Development	6234.21	32495.4	26261.19
Agriculture	2495.25	8562.87	6067.62
Sage brush	148235	100193.58	-48041.5
Riparian	17374.14	22941.36	5567.22
Bare surfaces	42060.78	36994.95	-5065.83

4 Discussion

Assessing biodiversity components and carrying out progressive and pro-active conservation research and management can require intensive digital data preparation and computationally demanding applications. Usually, this is done with the extensive use of modeling and statistical soft- and hardware and advanced (spatial) datasets; see [31], [32],[33] for change detection modeling and [29] for a future Landscape example. Much more multidisciplinary expertise and international projects are needed to address issues relevant for sustainable landscape management on larger scale. This localized project and its methodology can be applied on a larger, continental and global scale, as well as with other species. It is made possible through the advent of large scale data (Sagemap USGS website: http://sagemap.wr.usgs.gov.htm, Wyoming State clearing house: http://www.wygisc.uwyo.edu/clearinghouse/county.html) as well as through global data sets such as Landcover and Human Populations (SAGE and HYDE [34]), Climate Data (CRU[35]), Digital Elevation (etopo2 [36]), and others. These global aspects to species and biodiversity modeling should be enforced much stronger. First approaches are coming forward for the globe as well as for world-wide marine and terrestrial applications (e.g. Neubauer and Huettmann unpublished, Meyer and Huettmann unpublished).

It is well known that habitats are modified, transferred and lost globally at an increasing rate [37]. For the study area, table 2 shows that roads, conifers, deciduous, human development, agriculture, riparian cover classes, which are not favorable for the preservation of sage grouse breeding habitats, increased. This trend is in contrast

to herbaceous rangelands, grasslands, sagebrush and bare surface classes which sage grouse is more associated with and favorable to the species [38].

The presented methods are affected by error propagation and accuracy issues [39]. In this study, the classification result from the Landsat 1985 image could not statistically get assessed for accuracy and reliability. Secondly, the Markov chain model used in predicting future land use scenario is based on a large assumption that the rate and type of change in a land use at a given time will be the same for another time period [26], [27]. Also, the reliability of spatial models developed using logistic regression and MARS have not been statistically ground-truthed and its use and success in predicting sage grouse breeding habitats has not been quantified. Unfortunately, future models cannot be ground-truthed with real data neither since still in the future. More research is needed to overcome these problems across scales, but the overall methodology is established now and gets further developed [29]. We think the future will see an increased use of wildlife habitat resources, but also the data availability and computer power will increase making methods like the ones presented here more feasible providing opportunities for decisions support systems on a large and global scale (see for instance [40]). We suggest to further emphasize these biology-based computer models, their data and applications in governments, for environmental impact studies and in policy towards a sustainable future of the globe (for more details on Biodiversity Informatics see also http://jbi.nhm.ku.edu/viewissue.php?id=2004].

Acknowledgments

This international project was supported by the University of Calgary, Ford Foundation International Fellowship Fund, German Academic Exchange Service known as Deutscher Akademischer Austausch Dienst (DAAD), Wyoming Fish and Game Department, Wyoming. US, Digital Environmental Management Canada and Stuttgart University of Applied Sciences, Germany. We wish to thank M. Hall-Beyer, all committee members, J. and S. Linke, G.C and E.C. Onyeahialam and family, D. Jacobson, D.J. Hagan, L. Strecker, J. and C. Finnigan and all other project contributors and supporters. This is EHWALE publication no 15.

References

1. Cogan C.B.: Biodiversity Conflict Analysis at Multiple Spatial Scales : In Scott J. Michael, Patricia J. Heglund, Michael L. Morrison, Jonathan B. Haufler, Martin G. Raphael, William A. Wall, Fred B. Samson (eds) Predicting Species Occurrences Issues of Accuracy and Scale. (2002) 229-239.
2. Manly B., F.J, Lyman L. McDonald, D.L. Thomas, T.L. McDonald, W.P. Erickson:Resource Selection by Animals: Statistical Design and Analysis for Field Studies. Kluwer Academic Publishers. (2002).
3. Smallwood, K.S., and Morrison, M.L.: Using the best scientific data for endangered species conservation. - Environ. Manage. (1999):421-435
4. Huettmann, F.: Databases and science-based management in the context of wildlife and habitat: towards a certified ISO standard for objective decision-making for the global community by using the internet. - J. of Wildl. Manag. (in press)

5. McArthur E.D. and Ott, J.E.: Potential natural vegetation in the 17 conterminous western United States. in Barrow, J.R., McArthur, E.D., Sosebee, R.E. and Tausch, R.J. (compilers) Proceedings: Shrubland ecosystem dynamics in a changing environment. - General Technical Report INT-GTR-338, (Odgen, UT: Intermountain Research Station) (1996) 16-28
6. Connelly J. W., Schroeder, M. A. Schroeder, Sands, A.R., and Braun, C.E.: Sage Grouse Management Habitat And Management: Guidelines to manage sage grouse populations and their habitats. - Wildlife Soc. Bull. (2000) 967–985.
7. Bent A.C: Life histories of North American gallinaceous birds. - Smithsonian Inst., U.S. Nat. Mus. Bull.162. (1932)
8. Beck T. D. I., and C. E. Braun: The strutting ground count, variation, traditionalism, management needs. - Proc. Ann. Conf. West. Assoc. Fish Wildl. Agencies 60 (1980) 558-566.
9. Schroeder M. A., Young, J.R., and Braun, C.E.: Sage grouse (*Centrocercus urophasianus*). *In* A. Poole and F. Gill, editors. - The Birds of North America, No. 425. The Birds of North America, Philadelphia, Pennsylvania, USA.(1999) 1-28
10. Hupp J. W. and Braun, C.E.: Topographic distribution of sage grouse foraging in winter. - J. of Wildl. Manage.(1989) 823-829.
11. Onyeahialam A. I.: Integration of Classification Techniques for Vegetation and Land Use Classes using Landsat 7 ETM+ and GIS data. M.Sc Thesis submitted at Stuttgart University of Applied Sciences, Stuttgart. Germany (2003).
12. Fisher C., Gustafson, W. and Redmond, R.: Mapping Sagebrush/Grasslands FromLandsat TM-7 Imagery: A Comparison Of Methods. Prepared for the USDI, Bu-reau of Land Management Montana/Dakota State Office (2002).
13. Goodchild M.F.: Spatial Autocorrelation. Geo Books Norwich, CATMOG (1986): 47
14. Legendre, P.: Spatial Autocorrelation: Trouble or New Paradigm? - Ecology (1993) 1659-1673.
15. Cliff A. D. and Ord, J.K: Spatial Autocorrelation, Pion (1973).
16. Hosmer D.W. and Lemeshow, S.: Applied Logistic Regression. John Wiley and Sons Inc. New York. (1989)
17. McCullagh, P. and Nelder, J.A.: Generalized Linear Models. Chapman and Hall, London (1983).
18. Burnham K.P. and D.R. Anderson: Model Selection and Inference: a practical information-theoretic approach. Springer-Verlag, New York (1998).
19. Frank E. Harrell, Jr.: Regression modelling strategies with application to linear models, Logistic regression and survival analysis. Springer (2001).
20. Huettmann, F. and J. Linke (2003): An automated method to derive habitat preferences of wildlife in GIS and telemetry studies: A flexible software tool and examples of its application. - European J. of Wildl. Res. (2003) 219-232.
21. Yen, P., F. Huettmann, and F. Cooke: Modelling abundance and distribution of Marbled Murrelets (*Brachyramphus marmoratus*) using GIS, marine data and advanced multivariate statistics. - Ecol. Mod. (2004) 395-413
22. Hastie.T, R.Tibshirani and J. Friedman: The elements of statistical learning Data Mining, Inference, and Prediction. - Springer Series in Statistics (2001).
23. Lek, S., A. Belaud, P. Baran, I. Dimopoulos, J. Lauga and S. Aulagnier : Application of neural network to modelling non-linear relationships in ecology. - Ecol. Mod. (1996) 39-52.
24. Lusk J., Guthery, F.S., and DeMase, S. J.: A Neural Network for Predicting Northern Bobwhite Abundance in the Rolling Red Plains of Oklahoma. I*n:* Scott J. M., P. J. Heglund, M. L. Morrison, J.B. Haufler, M.G. Raphael, W. A. Wall, F.B. Samson (eds): Predicting Species Occurrences Issues of Accuracy and Scale. (2002) 345-356.

25. Huettmann F. and A.W. Diamond: Seabirds and Ecological Scale Questions for the North Atlantic. Landsc. Ecol. (in review).
26. Urban D.L and Wallin, D.O.: Introduction to Markov models in Learning Landscape Ecology. A practical guide to concepts and techniques by Gergel S.E and M.G Turner. Chapter 4 Springer Verlag NY. Inc (2002) 35-47.
27. Hogweg P.: Cellular automata as a paradigm for ecological modeling. Appl. Math. Comput. (1988) 81–100.
28. Drton. M, Marzban C. Guttorp .P and Schaefer, J.T.: A Markov Chain Model of Tornadic Activity. Department of Statistics, University of Washington, Seattle. June 8, (2003).
29. Huettmann, F., Franklin, S. and Stenhouse, G. : Deriving Future Landscape Scenarios of Grizzly Bear Habitat. - Forestry Chronicle (in press).
30. LaMontage J.M., Irvine, R.L. and Crone,E.E.: Spatial Patterns of population regulation in sage grouse (*Centrocercus* spp.) population viability analysis. -J. of Anim. Ecol. (2002) 672-682.
31. Riebsame.W. E., Parton, W. J., Galvin, K. A., Burke, I. C., Bohren, L., Young, R. and Knop, E. : Integrated modeling of land use and cover change. -Bioscience (1994) 350–356.
32. Jenerette G. D and Jianguo Wu: Analysis and simulation of land-use change in the central Arizona – Phoenix region, USA**. -**Landsc. Ecol. (2001): 611–626.
33. Brondizio E., E. Moran, P. Mausel, and Y.Wu: Land cover in the Amazon Estuary: Linking of the Thematic Mapper with botanical and historical data. -Photogrammetric Engineering and Remote Sensing, (1996) 921-929.
34. Ramankutty, N., and Foley, J.A.: Estimating historical changes in global land cover: crop lands from 1700 to 1992. – Global Biogeochemical Cycles (1999) 997-1027.
35. New, M., Lister, D., Hulme, M. and Makin, I.: A high-resolution data set of surface climate over global land areas. – Clim. Res. (2002) 1-25.
36. National Geophysical Data Center (NGDC): ETOPO2. NOAA/NGDC E/GC3, 325 Broad way, Boulder, CO USA http://www.ngdc.noaa.gov/mgg/fliers/01mgg04.html
37. Forman, R.T.T.: Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press (1995)
38. Braun C. E.: Sage grouse declines in Western North America: what are the problems? - Proceedings, Western Association of Fish and Wildlife Agencies (1998) 139-156.
39. Wisdom M.J., Wales, B.C., Rowland, M.M., Raphael, M.G., Holthausen, R.S., Rich, T.D. and Saab, V.A.: Performance of Greater sage grouse models for Conservation Assessement in the Interior Columbia Basin, USA. - Conservation Biology (2002) 1232-1242.
40. Refish, J., Kone, I. and Huettmann, F.: Review article: Hunting monkeys in tropical forests and its implications for biodiversity and human health. Frontiers in Ecology and the Envi ronment. (submitted).

Local Analysis of Spatial Relationships: A Comparison of GWR and the Expansion Method

Antonio Páez

Centre for Spatial Analysis/School of Geography and Geology,
McMaster University,
1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1
paezha@mcmaster.ca

Abstract. Considerable attention has been paid in recent years to the use and development of local forms of spatial analysis, including the method known as geographically weighted regression (GWR). GWR is a simple, yet conceptually appealing approach for exploring spatial non-stationarity that has been described as a natural evolution of the expansion method. The objective of the present paper is to compare, by means of a simulation exercise, these two local forms of spatial analysis. Motivation for the exercise derives from two basic research questions: Is spatial non-stationarity in GWR an artifact of the way the model is calibrated? And, how well does GWR capture spatial variability? The results suggest that, on average, spatial variability in GWR is not a consequence of the calibration procedure, and that GWR is sufficiently flexible to reproduce the type of map patterns used in the simulation experiment.

1 Introduction

In recent years there has been, as documented by Fotheringham and Brunsdon [1], an increased interest in the application and development of local forms of spatial analysis – methods that produce local, mappable results, as opposed to the single, one-size-fits-all results of traditional global methods. Evidence of this interest is the attention that local measures of spatial association have aroused, as for example the $G_i(d)$ [2] and LISA [3] statistics. In a multivariate framework, the spatial variability of relationships, or what has been termed spatial non-stationarity, has also received attention. Early developments that consider the possibility of spatially-varying relationships include the expansion method, a form of analysis of the covariance proposed in the early 1970s by Casetti [4]. More recent research includes the development of geographically weighted regression, or GWR for short [5]. Among the techniques used to study spatial non-stationarity, the expansion method and GWR are conceptually very attractive because they recognize the continuous nature of space in many settings. In addition, by working with elemental spatial units, these methods avoid the problems that arise when using modifiable areal units.

The roots of GWR can be found in local regression in statistics, and methods developed there based on the simple yet powerful idea of using sub-samples of data around specific points in attribute space [6]. More recently, GWR has been described

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 162 – 172, 2005.

as a natural evolution of the expansion method [7]. Given the interest that the idea of developing local models GWR-style has generated (e.g., [8]), it is important to address some issues that have so far tended to remain in the background. An important question relates to the situation, often encountered in practice, of finding extremely high coefficient variability, including sign reversals (see for example [9], and discussion in [10]). This situation naturally stirs up a concern that GWR results may be misleading. Underlying this concern is a more fundamental question, namely whether the variability observed is somehow built into the model by its own calibration and estimation mechanisms. The objective of this paper is to compare, using of a simulation exercise, GWR and the expansion method. Motivation for this exercise derives from two research questions: Is spatial non-stationarity in GWR an artifact of the way the model is calibrated? And, how well does GWR capture spatial variability? The results reported suggest that, on average, spatial variability in GWR is not an artifact of the calibration procedure, and that GWR is sufficiently flexible to reproduce the type of map patterns used in the simulation experiment.

2 The Models

2.1 The Expansion Method

The expansion method was proposed as a tool to model contextual variations, based on the principle of expanding the coefficients of an initial model as functions of contextual variables. The process of expanding the coefficient reflects, it has been argued, the way analysts prioritize their knowledge of spatial processes [11]. The result is a terminal model that incorporates "main" effects and "contextual" effects, or interactions. The following is a generic initial model that relates a variable of interest y to explanatory variables x, and a small amount of random variation ε.

$$y_i = \sum_{k=1}^{K} x_{ki}\beta_k + \varepsilon_i \quad . \tag{1}$$

The value of variable x_{1i} is usually set to 1 for all i (i.e., to give a constant term). The coefficients of this model can be expanded based on the coordinates (u_i,v_i) of location i to take into account the effect of spatial context. Consider as an example the following linear expansion of a simple bivariate model (note the use of subscript i in the expanded coefficients):

$$\beta_{1i} = \beta_{11} + \beta_{12}u_i + \beta_{13}v_i \tag{2}$$

$$\beta_{2i} = \beta_{21} + \beta_{22}u_i + \beta_{23}v_i \quad . \tag{3}$$

The terminal model becomes:

$$y_i = \beta_{1i} + \beta_{2i}x_{2i} + \varepsilon_i = \beta_{11} + \beta_{12}u_i + \beta_{13}v_i + \left(\beta_{21} + \beta_{22}u_i + \beta_{23}v_i\right)x_i + \varepsilon_i \quad . \tag{4}$$

Polynomial expansions of higher degrees (i.e.. quadratic, etc.) can be used. Making suitable assumptions, the terminal model can be estimated using ordinary least squares (OLS) and the significance of the coefficients can be tested in the usual way.

The expansion method is a conceptually appealing and technically simple approach to modeling non-stationarity that, moreover, does not involve discontinuities of the estimated coefficient surface. It has been argued, however, that depending on the form of the expansion, the method might fail to capture more complex variation patterns [7]. It is within this context that GWR has been described as a natural evolution of the expansion method, and a more flexible approach to capture non-stationarity.

2.2 Geographically Weighted Regression

GWR is defined as a semi-parametric model of the following form:

$$y_i = \sum_{k=1}^{K} x_{ki}\beta_k(u_i, v_i) + \varepsilon_i \ . \tag{5}$$

The error terms in the above expression assume the usual conditions of independence and constant variance (see Brundson et al. [12], p. 502). The difference with the global model is that the coefficients $\beta_k(u_i,v_i)$, instead of being constant, are unspecified (i.e., non-parametric) functions of the geographical coordinates of point i (u_i,v_i). As Brunsdon et al. note, if the function $\beta_k(u_i,v_i)$ is reduced to a constant, the model becomes the ordinary global, least squares specification [12]. To the extent that the coefficients of the model are continuous functions of space, it could be said that the task of the analyst is to evaluate (i.e. obtain a specific value of) the unspecified function at a given point in space.

Specific values of the set of functions $\hat{\boldsymbol{\beta}}(u_i, v_i)$ in model (5) are obtained using:

$$\hat{\boldsymbol{\beta}}(u_i, v_i|\gamma) = \left[\mathbf{X}^T\mathbf{W}(u_i, v_i|\gamma)\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{W}(u_i, v_i|\gamma)\mathbf{y} \tag{6}$$

where $\mathbf{W}(u_i, v_i|\gamma)$ is a diagonal matrix ($n\times n$) of geographical weights defined by a geographical weighting (or kernel) function such as:

$$w_{ij} = \exp\left(-\gamma d_{ij}^2\right) \ . \tag{7}$$

Clearly, estimation of the spatial coefficients depends, in addition to the distance d_{ij} between observations i and j, on an unknown quantity, called a kernel bandwidth γ. This quantity controls the steepness of the kernel function, and thus the values of the weights assigned to individual observations. Estimation requires that this kernel bandwidth be calibrated, something typically achieved by means of a procedure known as cross-validation, in which a value of γ is sought to minimize the score:

$$C(\gamma) = \sum_{i=1}^{N}\left(y_i - \hat{y}_{\sim i}(\gamma)\right)^2 \ . \tag{8}$$

The cross-validation score $C(\gamma)$ is simply the sum of squared errors between the observed values and the values predicted by the model at different points in space as a function of γ. The score can thus be seen as an overall measure of goodness-of-fit. In order to avoid perfect "fits" that disregard all but one observation, a leave-one-out procedure is adopted to obtain the bandwidth (i.e., observation y_i is not used to obtain estimated value $\hat{y}_{\sim i}$). Once a value of the kernel bandwidth has been found that

minimizes the cross-validation score, this value can be plugged into the geographical weighting function [equation (7)], which is in turn used to estimate the value of the spatial coefficients at a given location using equation (6). Other technical details of this procedure can be found in (among other papers) Brunsdon et al. [5].

Note that when the value of the kernel bandwidth in equation (7) is 0, the weights become unity for all pairs of observations. Since all observations receive identical unitary weights regardless of location, the model produces spatially invariant coefficients (i.e., a global model). When the bandwidth becomes very large, on the other hand, the weights tend to produce large amounts of coefficient variability in order to simply "connects-the-dots". To what extent if at all, then, is non-stationarity in GWR a consequence of the calibration procedure? The literature does not offer helpful guidance because work to date has tended to use empirical data for which the true underlying process is not known. Simulation of known processes is expected to shed light on this issue. Design of the simulation experiment is described next.

3 Simulation Design

Three models with known properties were defined for the simulation experiment, corresponding to a global process, a linear expansion (LE) and a quadratic expansion (QE). The global model was defined as a bivariate regression equation as follows:

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \ . \tag{9}$$

with spatially constant coefficients β_1=1.8 and β_2=1.2.

Draws of the independent variable x_i were taken from a uniform distribution bounded between 2 and 5, and the error terms ε_i were drawn from a standard normal distribution [N(0,1)]. The independent variable and the error terms are uncorrelated across space. Four hundred observations were simulated and placed in space at regular intervals (coordinates u_i and v_i), using a unit square area to avoid the potential for badly scaled matrix calculations. Scaling the distance to the unit does not affect the results of the model, as scale is absorbed by the coefficients, which are preset by design in the simulation. The process of randomly drawing values for x_i and ε was repeated 1000 times to produce as many replications for the simulation experiment.

The variables used to produce the global model were also used to produce the two models with locally-varying coefficients. The general form of the local models was given by the following expression:

$$y_i = \beta_{1i} + \beta_{2i} x_i + \varepsilon_i \ , \tag{10}$$

The local coefficients of the linear expansion (LE) were defined as follows:

$$\beta_{1i} = 1 + 0.7u_i + 0.3v_i \tag{11}$$

$$\beta_{2i} = 1 + 0.3u_i - 0.2v_i \tag{12}$$

to produce the coefficient surfaces shown in Figure 1.

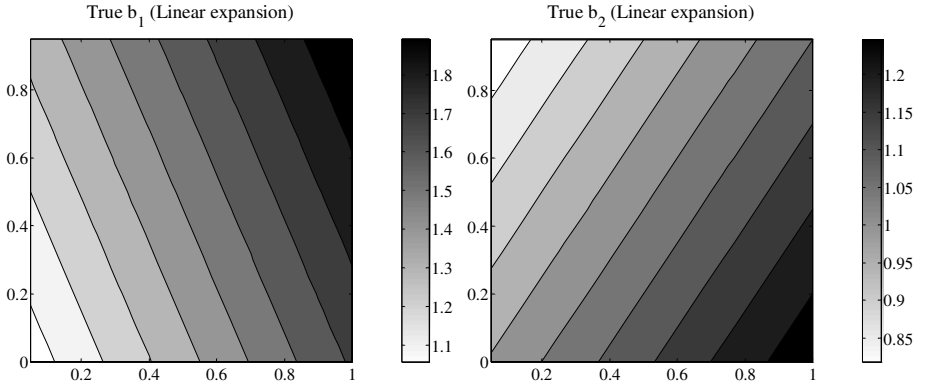


Fig. 1. True coefficient surfaces for model 1 (Linear expansion)

The quadratic expansion (QE) model, on the other hand, was defined using the following expressions for the two coefficients:

$$\beta_{1i} = 1 - 0.5u_i^2 + 0.1u_i v_i + v_i^2 \tag{13}$$

$$\beta_{2i} = 1 + u_i^2 - 0.5u_i v_i + 1.5v_i^2 \ . \tag{14}$$

These expressions produce the coefficient surfaces shown in Figure 2.

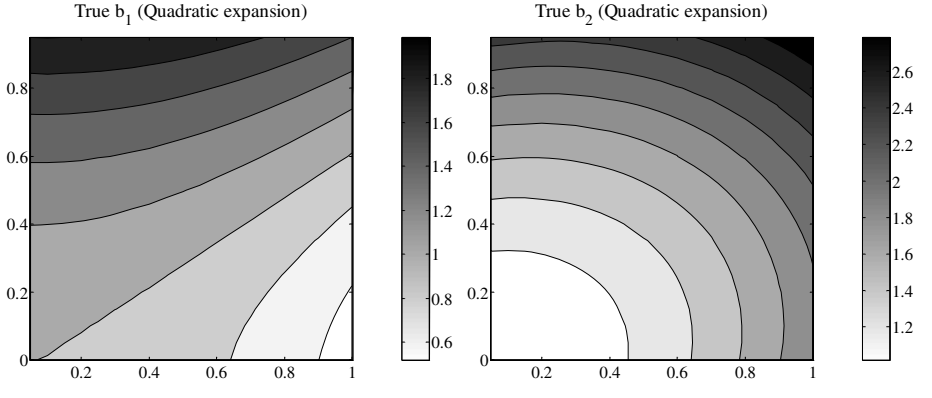


Fig. 2. True coefficient surfaces for model 2 (Quadratic expansion)

4 Results

After data was simulated, three models (LE, QE, and GWR) were estimated for each replication. Of these, the two expansions match one of the three known processes. The following figures and tables summarize the main findings of the simulation experiment.

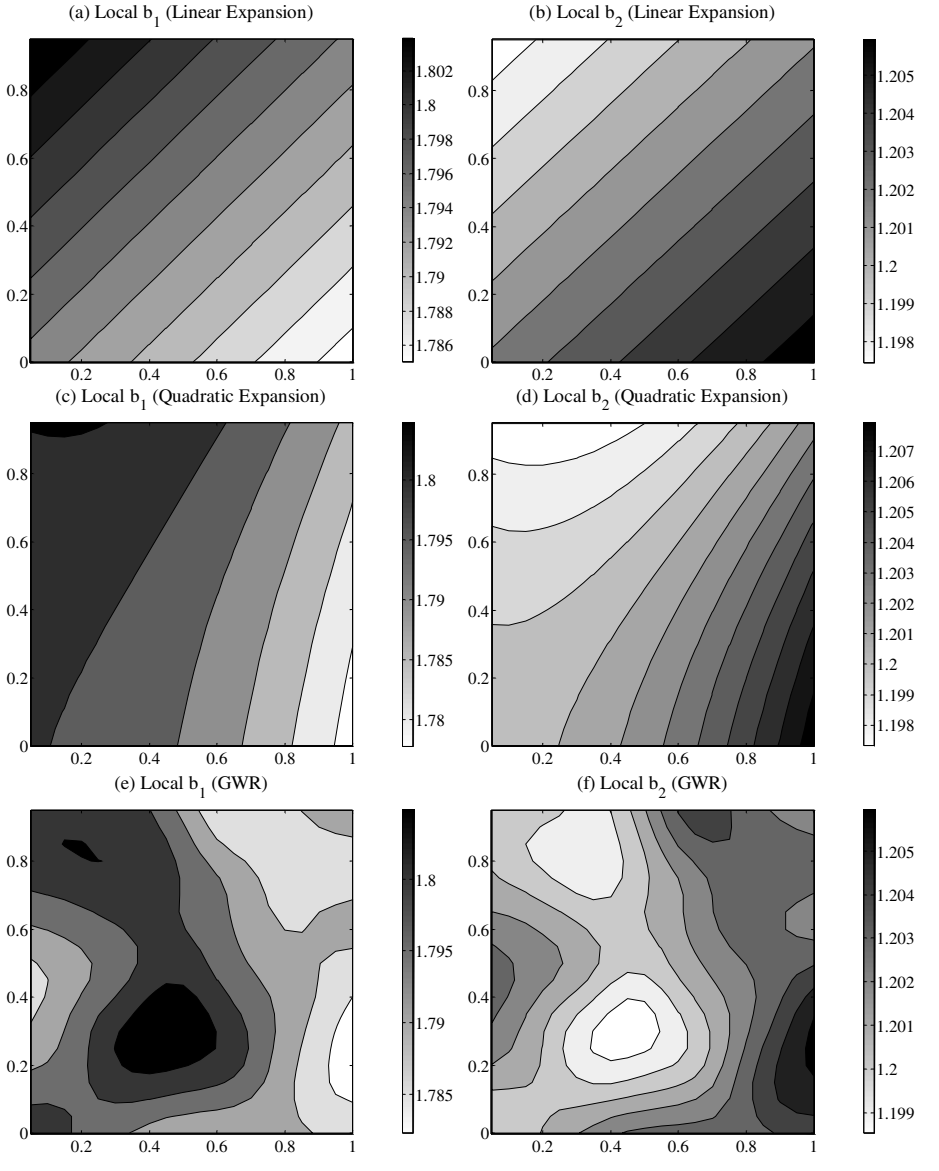


Fig. 3. Coefficient surfaces (β_1 and β_2) for three models: Figure 3 *(a)* and *(b)* are linear expansion coefficients; Figure 3 *(c)* and *(d)* quadratic expansion coefficients; and Figure 3 *(e)* and *(f)* are GWR coefficients. The results are the mean of 1000 simulations. True values in the global model are β_1=1.8 and β_2=1.2

Figure 3 shows the coefficient surfaces that result from the average of 1000 replications, when the true model is global [equation (9)]. Although there is some variability, the surfaces are essentially flat. GWR is not artificially producing more variability than exists in the underlying process. The average kernel bandwidth was close to 0.

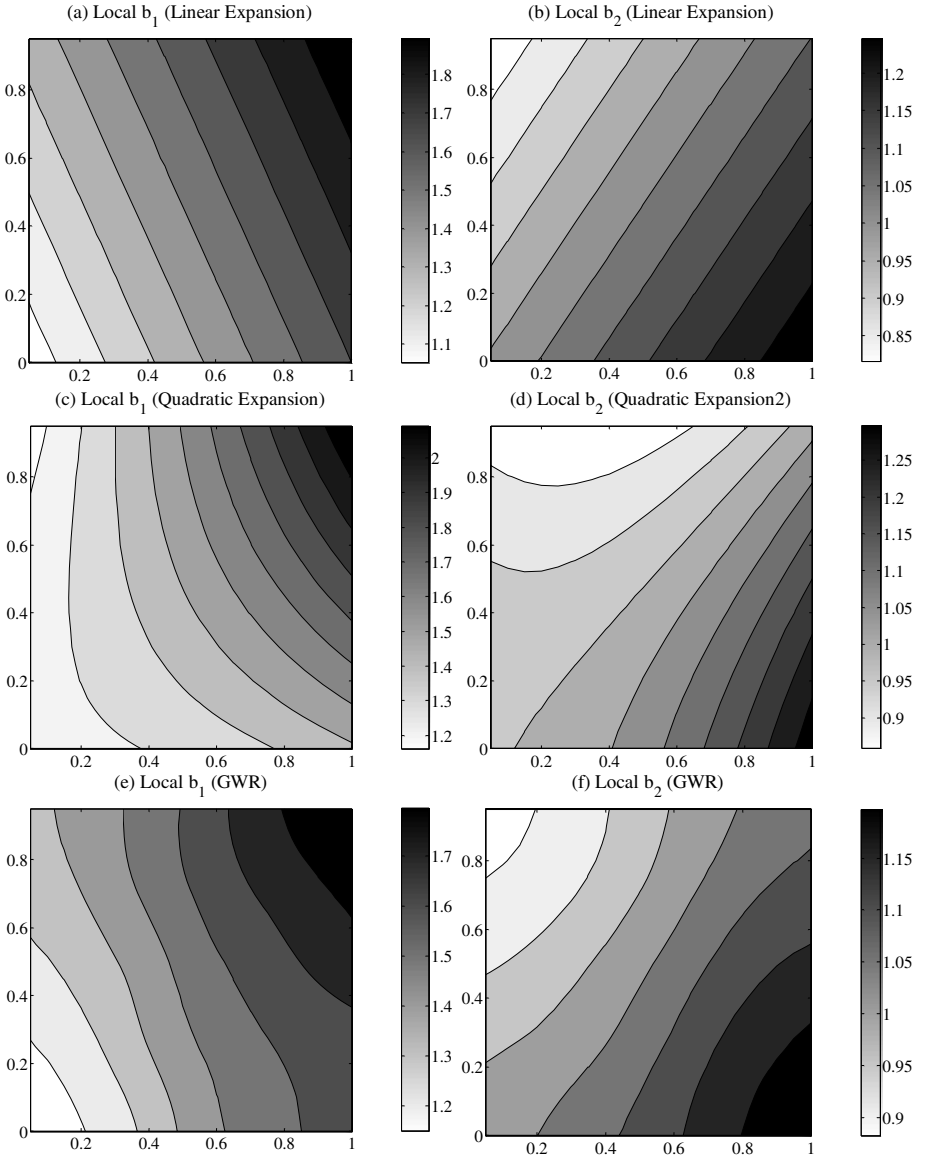


Fig. 4. Coefficient surfaces (β_1 and β_2) for three models: Figure 4 *(a)* and *(b)* are linear expansion coefficients; Figure 4 *(c)* and *(d)* quadratic expansion coefficients; and Figure 4 *(e)* and *(f)* are GWR coefficients. The results are the mean of 1000 simulations. True values are as shown in Figure 1

Next, the true model is a linear expansion. As seen in Figure 4, all three models retrieve (on average) the underlying process, with some distortions, produced by the parametric definition of the expansion, in the case of the quadratic expansion. These distortions appear to be less serious in the case of GWR.

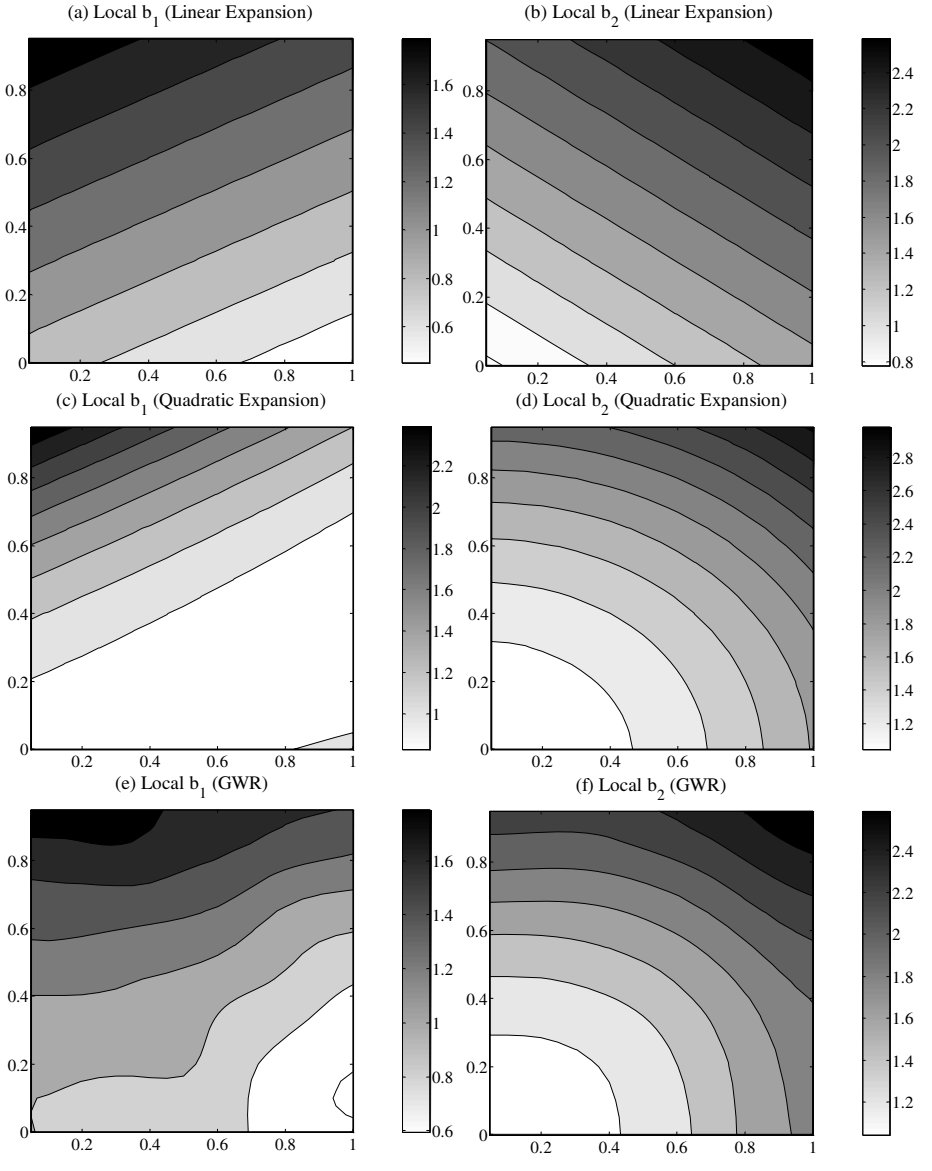


Fig. 5. Coefficient surfaces (β_1 and β_2) for three models: Figure 5 *(a)* and *(b)* are linear expansion coefficients; Figure 5 *(c)* and *(d)* quadratic expansion coefficients; and Figure 5 *(e)* and *(f)* are GWR coefficients. The results are the mean of 1000 simulations. True values are as shown in Figure 2

Figure 5 shows the coefficients when the true model is a QE. Again, all three models capture the general trend, but the linear expansion is limited by its parametric

definition. The quadratic expansion successfully captures the general shape of the true surface, but GWR may be closer to the true one, as suggested by Table 1 below.

In addition to the figures, some preliminary analysis was conducted on the estimated coefficients and model performance. Table 1 shows the correlation between the estimated coefficient surfaces and the known values. The correlation provides an indication of the ability of the different models to replicate the known surface, or in other words, of their ability to retrieve the underlying process. The estimated surface used to calculate the correlation is the average of 1000 replications. All models perform reasonably well in this regards. In the case of GWR, it is interesting to note that the correlations are, for all practical purposes, as high, or higher, than those corresponding to the expansion models – despite the fact that the true underlying models were derived based on coefficient expansions.

Table 1. Correlation analysis between true coefficients and estimated coefficients (surface compared in each case is the mean of 1000 replications)

		Estimated coefficients					
		LE		QE		GWR	
True coefficients		$\hat{\beta}_{1i}$	$\hat{\beta}_{2i}$	$\hat{\beta}_{1i}$	$\hat{\beta}_{2i}$	$\hat{\beta}_{1i}$	$\hat{\beta}_{2i}$
LE	β_{1i}	0.9998	-	0.9393	-	0.9944	-
	β_{2i}	-	1.0000	-	0.9452	-	0.9959
QE	β_{1i}	0.9722	-	0.9175	-	0.9941	-
	β_{2i}	-	0.9533	-	0.9957	-	0.9970

Note.- LE: linear expansion, QE: quadratic expansion

Table 2. Indicators of goodness of fit: absolute sum of errors and squared sum of errors (minimum, maximum and mean of 1000 replications)

		Absolute sum of errors ($\times 10^3$)			Squared sum of errors ($\times 10^4$)		
True model	Estimated model	Min	Max	Mean	Min	Max	Mean
Global	LE	2.298	2.491	2.401	1.3663	1.6007	1.4856
	QE	2.195	2.569	2.401	1.2360	1.7105	1.4855
	GWR	2.298	2.493	2.401	1.3661	1.6003	1.4855
LE	LE	1.984	2.168	2.085	1.0326	1.2311	1.1318
	QE	1.848	2.207	2.044	0.8852	1.2754	1.0871
	GWR	1.986	2.172	2.085	1.2265	1.2265	1.1285
QE	LE	2.831	3.064	2.947	2.2160	2.5998	2.3962
	QE	2.733	3.118	2.934	2.0819	2.7264	2.3875
	GWR	3.051	3.051	2.939	2.1970	2.5691	2.3793

Note.- Global: no expansion, LE: linear expansion, QE: quadratic expansion

Table 2, finally, shows two indicators of goodness-of-fit (absolute sum of errors, and squared sum of errors). The errors are calculated by comparing the known value of y_i to the value estimated using the model. The table thus provides indication of the ability of different models to replicate the data – or in other words, of their predictive accuracy. The differences appear to be minimal.

5 Discussion and Concluding Remarks

The results of the simulation experiment suggest that, on average, the spatial variability of the coefficients is not an artifact of the cross-validation procedure used to calibrate GWR models. For example, the coefficient surfaces were virtually flat and the kernel bandwidth was on average very close to 0 when the true model was a global model. In the case of other simulated processes, the average coefficient surfaces closely resemble the true values set in the simulation. In fact, correlation analysis of the estimated and known values suggests that GWR is sufficiently flexible to capture the kind of map patterns used in the simulation – in some cases even gaining a slight advantage over the expansion method. In terms of predictive accuracy, the difference in performance between the different models seems to be marginal at best. The results presented in this paper provide qualified support for the idea that GWR is not artificially producing spatial variability, at least on average. Further research is required to explore to what extent this remains so at the level of individual replications. Also, in this paper, only global and expansion type of processes were simulated. Coefficient surfaces of higher complexity should be considered in future research.

References

1. Fotheringham,A.S., Brunsdon, C.: Local forms of spatial analysis. Geographical Analysis Vol. 31. (1999) 340-358
2. Getis,A., Ord, J. K.: The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis Vol. 25. (1993) 276-276
3. Anselin,L.: Local Indicators of Spatial Association - LISA. Geographical Analysis Vol. 27. (1995) 93-115
4. Casetti,E.: Generating Models by the Expansion Method: Applications to Geographic Research. Geographical Analysis Vol. 28. (1972) 281-298
5. Brunsdon,C., Fotheringham, A. S., Charlton, M. E.: Geographically weighted regression: A method for exploring spatial nonstationarity. Geographical Analysis Vol. 28. (1996) 281-298
6. Cleveland,W.S.: Robust Locally Weighted Regression and Smoothing of Scatterplots. Journal of the American Statistical Association Vol. **74**. (1979) 823-836
7. Fotheringham,A.S., Charlton, M. E., Brunsdon, C.: Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environment and Planning A Vol. 30. (1998) 1905-1927
8. Páez,A.: Anisotropic Variance Functions in Geographically Weighted Regression Models. Geographical Analysis Vol. 36. (2004) 299-314

9. Foody,G.M.: Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. Global Ecology and Biogeography Vol. 13. (2004) 315-320
10. Jetz,W., Rahbek, C., Lichstein, J. W.: Local and global approaches to spatial data analysis in ecology. Global Ecology and Biogeography Vol. 14. (2005) 97-98
11. Casetti,E.: Mixed Estimation and the Expansion Method: An Application to the Spatial Modelling of the AIDS Epidemic. (1997) 15-34
12. Brunsdon,C., Fotheringham, A. S., Charlton, M.: Some notes on parametric significance tests for geographically weighted regression. Journal of Regional Science Vol. 39. (1999) 497-524

Middleware Development for Remote Sensing Data Sharing and Image Processing on HIT-SIP System

Jianqin Wang[1], Yong Xue[1,2,*], Chaolin Wu[1], Yanguang Wang[1], Yincui Hu[1], Ying Luo[1], Yanning Guan[1], Shaobo Zhong[1], Jiakui Tang[1], and Guoyin Cai[1]

[1] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Institute of Remote Sensing Applications, Chinese Academy of Sciences, P. O. Box 9718, Beijing 100101, China

[2] Department of Computing, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, UK

Tian1.wang@163.com, y.xue@londonmet.ac.uk

Abstract. Sharing spatial data derived from remote sensing is a very significant thing. Grid computing and Web Service technology provides fundamental support for it. In this paper we mainly discuss architecture and middleware of sharing spatial data derived from remote sensing and processing. Because middleware of automatically transferring and task execution on grid is the key of the architecture, we study the middleware. It can effectively protect the owner of data and middleware's property through giving users their required result not just simply copying data and codes resource to them. Based on this sharing architecture and middleware technology, a data and middleware transferring example is showed.

1 Introduction

The Grid (http://www.globus.org/about/faq/general.html#grid) refers to an infrastructure that enables the integrated, collaborative use of high-end computers, networks, databases, and scientific instruments owned and managed by multiple organizations. Grid applications often involve large amounts of data and/or computing and often require secure resource sharing across organizational boundaries, and are thus not easily handled by today's Internet and Web infrastructures. Ian Foster *et al.* [1][2][3] offered several definitions of grid: "A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities". Grid computing is concerned with "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations (VO). … [A VO is] a set of individuals and/or institutions defined by [some highly controlled] sharing rules."

There are several famous grid projects today. Access Grid (ww.fp.mcs.anl.gov/fl/access_grid) lunched in 1999 and mainly focused on lecture and meetings-among

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 173 – 180, 2005.

scientists at facilities around the world. European Data Grid sponsored by European union, mainly in data analysis in high-energy physics, environmental science and bioinformatics. Grid Physics Network (GriDPhyN) [4] lunched in 2000 and is sponsored by US National Science Foundation (NSF) mainly in data analysis for four physics projects: two particle detectors at CERN's Large Hadron Collider, the Laser Interferometer Gravitational Wave Observatory, and the Sloan Digital Sky Survey. Information Power Grid [5] is the NASA's computational support for aerospace development, planetary science and other NASA research. International Virtual DataGrid Laboratory (iVDGL) [6] was sponsored by NSF and counterparts in Europe, Australia, and Japan in 2002. Network for Earthquake Engineering and Simulation labs (NEESgrid) (www.neesgrid.org) intended to integrate computing environment for 20 earthquake engineering labs. TeraGrid (www.teragrid.org) is the general-purpose infrastructure for U.S. science: will link four sites at 40 gigabits per second and compute at up to 13.6 teraflops. UK National Grid is sponsored by U.K Office of Science and Technology[7]. Unicore (www.unicore.de) is a seamless interface to high-performance Education and Research computer centers at nine government, industry and academic labs. The famous Grid focused on spatial information includes SpaceGrid, EnvirGrid and EarthObsevation Grid. ESA's SpaceGrid is an ESA funded initiative (http://sci2.esa.int/spacegrid).

EnvirGrid main goals are generalization of Earth Science application infrastructure to become GRID-aware, extend GRID access to European Environmental and Earth Science application to large science communities, to value adding and commercial communities, …, and demonstrate collaborative environment for Earth Science.

Dozens of satellites constantly collecting data about our planetary system 24 hours a day and 365 days a year. Large scale of satellite data needed to be processed and stored in real time or almost real time. So far real time processing in remote sensing confronts much difficulties in one single computer, or even impossibility. Computing grid that is integrated by series of middleware provides a way to solve this problem [8].

Serials of middleware [9] of remote sensing image processing on grid platform have been developing by Telegeoprocessing and Grid study group in Institute of Remote Sensing Applications, Chinese Academy of Science. But as common users, they can have neither remote sensing data nor processing middleware except requirement of result. Automatically transferring and task execution (ATTE) middleware on grid can meet the requirement through farthest sharing remote sensing data and processing middleware.

The paper mainly describes the mechanism and process of the middleware. We describe the High throughput Computing for Spatial Information Processing System on Grid platform (HIT-SIP) developed in our research group in IRSA, CAS in Section 2.

2 Remote Sensing Data Processing Middleware on HIT-SIP Grid Platform

Remote sensing data processing is characterized by magnitude and long period of computing. Real time or almost real time processing is impossible on simple personal computer. But on the other hand, large numbers of PCs waste computing and restoring resources in certain time when they are free. So how to use the free

computer. But on the other hand, large numbers of PCs waste computing and restoring resources in certain time when they are free. So how to use the free geographically distributed resources to process remote sensing data in cooperative type seems very significant and practicable. Luckily, emerging grid technology can provide large computing power in the Internet as the security type. The grid computing nodes resource distribute loosely. They may be supercomputers or common PCs with heterogeneous architecture. The grid computing nodes resource distribute loosely. They may be supercomputers or common PCs with heterogeneous architecture.

The HIT-SIP system on Grid platform has been developed in Institute of Remote Sensing Application, Chinese Academy of Science. It is an advanced High-Throughput Computing system specialized for remote sensing data analysis using Condor. Heterogeneous computing nodes including two sets of Linux computers and WIN 2000 professional computers and one set of WIN XP computer provide stable computing power. The grid pool uses java universe to screen heterogeneous characters. The configuration details of PCs in the pool are shown in Table 1. The structure of HIT-SIP system is shown in Figure 1. Figure 2 show the interface of front page of HIT-SIP system, which can be access by web service.

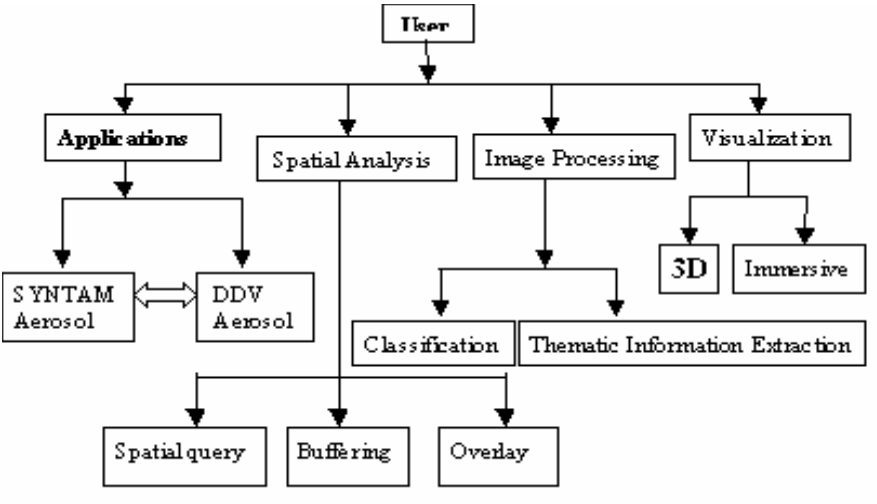


Fig. 1. The Structure of HIT-SIP system

The HIT-SIP system is a fundamental bare computing platform and can not run remote sensing image processing program. So remote sensing data processing middleware running on grid pool is inevitable. Common users can use the heterogeneous grid and share its strong computing power to process remote sensing data with middleware as if on one supercomputer. Till now, aerosol middleware, thermal inertia middleware and image classification middleware have been developed in HIT-SIP system.

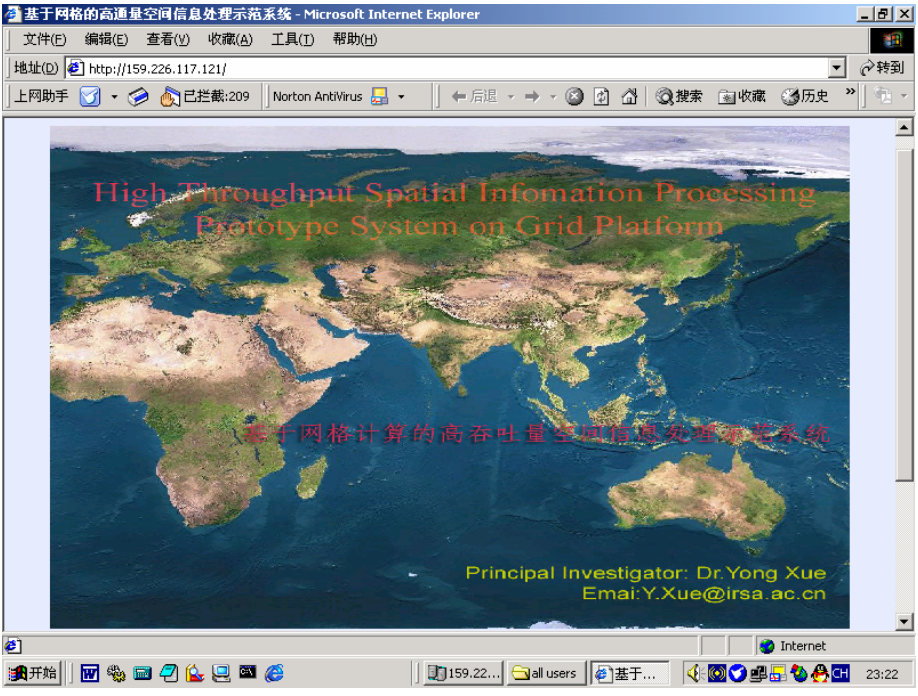


Fig. 2. The interface of front page of HIT-SIP system

Table 1. The configuration of computers used by HIP-SIP in the Grid pool

IP	Name	Arch-OS	Role
192.168.0.5	Manager.linux	Itel/Linux	Manager/client
192.168.0.3	Wjq.linux	Intel/WIN50	Client
192.168.0.102	Tele1.linux	Intel/WIN50	Client
192.168.0.111	Tele2.linux	Intel/WIN51	Client
192.168.0.6	Client2.linux	Intel/Linux	Client

3 Share of Remote Sensing Data and Image Processing Middleware

3.1 Sharing Requirement

Remote sensing data sharing became more and more significant with the digital information era coming. Grid technology also provides powerful tool for remote sensing data sharing .In the mean time, sharing remote sensing data processing middleware on grid naturally become very important. There still has a sharp contradiction between data and middleware owners and common users: as data and middleware owners, they don't just want to simply copy their production to other

people because of property, but as common users, they also don't want to take long time to develop middleware. We have developed Middleware prototype of automatically transferring and task execution on grid to solve this contradiction. It can both meet users' sharing requirement and protect owner's property.

3.2 Sharing Architecture of Remote Sensing Data and Image Processing

We bring forward sharing architecture of remote sensing data and image processing middleware codes based on grid technology and Web Service.

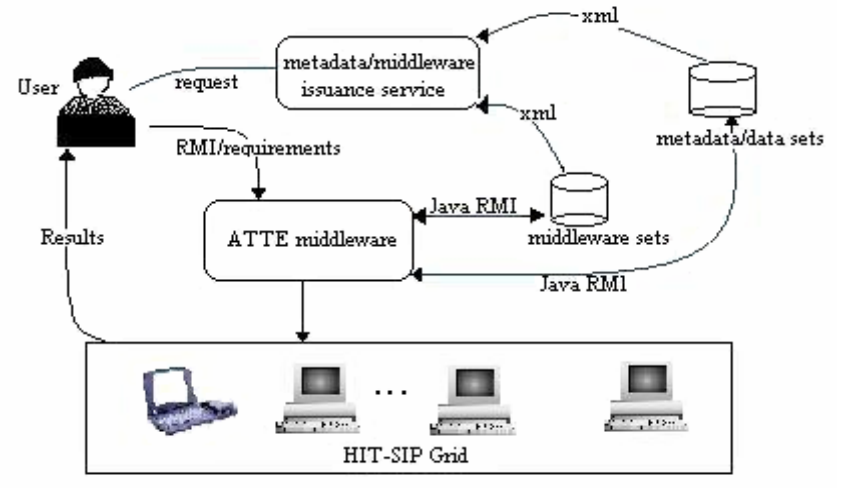


Fig. 3. Architecture of Remote Sensing Data and Image Processing middleware sharing

Geographically distributed remote sensing data sets and middleware codes data sets can be involved in this architecture. XML (extensible Markup Language) acting as a new standard for data interchanging on Internet provides a powerful tool for Web Service. Metadata of remote sensing and middleware codes will be registered at clearinghouse in certain suited type. So common users can find required remote sensing data and middleware through metadata matching module. Nowadays, there exists several metadata standards about geo-spatial data derived from remote sensing, of which "Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata" is the most famous one made by Federal Geographic Data Committee (FGDC). Metadata on the grid must be different with the FGDC's standards but with reference of it. Metadata of remote sensing data processing middleware must including task description files and middleware description files that must match the metadata format of remote sensing data.

Through module of metadata describing middleware and remote sensing matching service, user can easily find matched data location. Java Remote Method Invocation (RMI) is an advanced java technology used for distributed application program. Middleware of automatically transferring and task execution (ATTE) on grid can

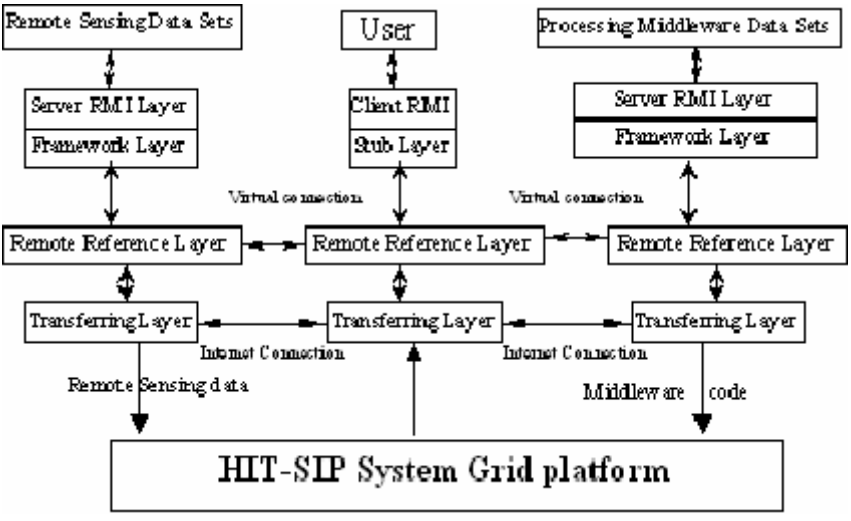


Fig. 4. Architecture of ATTE Middleware

long-distance transfer geospatial data derived from remote sensing satellite and remote sensing data processing middleware codes. When user gets the URL (including directory) of remote sensing data and processing middleware codes, he can easily starts ATTE. Matched data and middleware codes can be executed on grid platform and computed result be automatically transferred back to user.

3.3 Middleware of Automatically Transferring and Task Execution on Grid

Recently Java RMI has been popular. It is easier to use than CORBA because applications may only be in one language. Object can be easily serialized and transferred long-distance but not needed to execute on local computer.

The remote sensing data sets and processing middleware data sets nodes will be acted as server in the system architecture. Accordingly, user node will be acted as client. Application program on user node includes interface with server transferring module and task execution module on grid platform. RMIServer object will be serialized into Sub; so executing client sub is the same as executing server.

3.4 A Practical Example Using ATTE Middleware and Processing Middleware on Grid

On HIT-SIP grid platform, we test this sharing mechanism. Unsupervised Classification of Remotely Sensed Images middleware has been chosen as processing middleware codes resource. Overall resource and Internet configuration as following: The Table 2 shows remote sensing data sets resource located at Tele2 (IP: 192.168.0.110) and unsupervised classification image middleware codes at TGP (IP:192.168.0.110). Client (IP:192.168.0.3) can be the node in the grid pool or not. In this test we choose it as the grid node and share its computing resource.

Table 2. Remote sensing aata and processing arrangement in the Grid pool

IP	Name	Resource	Role	Grid node (yes/no)
192.168.0.6	Condordlient2	Computing		yes
192.168.0.3	WJQ	Computing	Client	yes
192.168.0.110	Tele2	RS data & Computing	Server	yes
192.168.0.1	TGP	Processing middleware	Server	No

Following is the remote sensing data and middleware description: In this test, we use IKONOS satellite dada as data resource, which IKONOS satellite images were acquired on 26 April 2002 with approximately 1 m and 4 m spatial resolution in the visible (panchromatic) and relatively broad bands in the blue, green, red, and near-infrared portions of the spectrum, respectively. We use bands of blue, green, red, and near infrared for about 4m spatial resolutions. Test ground surrounds Olympic park in Beijing. The middleware [9] is the no-alternative unsupervised calcification algorithm. We image that each job partition data should be computed individually and be submitted only once. After computing they can be fused and displayed. Through deep analysis we find that we can try to solve this problem for several reasons. Object can be distinguished and classified whatever in one whole image or in divided parts according to brightness of every pixel. The most important issue we care about is which type every object can be classified and the difference among different objects .We can do not care about the brightness of the whole core center. We set the classified number of each divided part. The problem may be that each divided part must be classified into given class number but it may be less the number in just one divided part. Through amalgamating dynamically class number through critical threshold we can adjust the class number in each divided part. Then classifying effect in each part can be almost the same as in one whole image after properly matching colors.

4 Conclusions

Architecture of sharing remote sensing data and image processing middleware codes provides a integrated reference framework for resource sharing. Grid computing and Web Service technology is really a very effective method for sharing remote sensing data and processing middleware codes. Metadata of Remote sensing and middleware will be issued at cleaning house. So user can find the correct URL of matched data resource. Through ATTE middleware that we have developed, the owner of data and middleware need not just simply copy the data to client because of property. User just gets the required result through ATTE

Fig. 5. Classification Result on Client Node

middleware on grid instead of getting raw data. It is very convenient not only for user but also for owner.

Acknowledgement

This publication is an output from the research projects "CAS Hundred Talents Program" and “Monitoring of Beijing Olympic Environment” (2002BA904B07-2) and “Remote Sensing Information Processing and Service Node” funded by the MOST, China and "Aerosol fast monitoring modeling using MODIS data and middlewares development" (40471091) funded by NSFC, China.

References

[1] I. Foster, “What is the Grid: A Three-Point Checklist”, *Grid Today*, 1 (6), 2002.
[2] I. Foster, C. Kesselman (Eds), *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers, 1998.
[3] I. Foster, C. Kesselman, S. Tuecke, “The Anatomy of the Grid: Enabling Scalable Virtual Organizations”, *Inter. Journal of Super-computer Applications*, 15(3), 2001, 200-222.
[4] Grid Physics Network.. http://www.griphyn.org/index.php.
[5] W. E. Johnston, IPG Chief Architect. Information Power Grid. http://www.ipg.nasa.gov/
[6] IVDGL Grid . http://www.ivdgl.org/
[7] UK e-Science Grid Support Center. http://www.grid-support.ac.uk/
[8] Jianqin Wang, Yong Xue, and Huadong Guo, 2003, A Spatial Information Grid Supported Prototype Telegeoprocessing System. In Proceedings of 2003 IEEE/IGARSS’2003 held in Toulouse, France on 21-25 July 2003, v 1, 345-347.
[9] Jianqin Wang, Xiaosong Sun, Yong Xue, Yanguang Wang, Ying Luo, Guoyin Cai, Shaobo Zhong, and Jiakui Tang, 2004, Preliminary Study on Unsupervised Classification of Remotely Sensed Images on the Grid. *Lecture Notes in Computer Science,* Vol. 3039, pp.995-1002.

A New and Efficient K-Medoid Algorithm for Spatial Clustering

Qiaoping Zhang and Isabelle Couloigner

Department of Geomatics Engineering, University of Calgary
2500 University Drive N.W.
Calgary, Alberta Canada T2N 1N4
{qzhang, couloigner}@geomatics.ucalgary.ca

Abstract. A new *k*-medoids algorithm is presented for spatial clustering in large applications. The new algorithm utilizes the TIN of medoids to facilitate local computation when searching for the optimal medoids. It is more efficient than most existing *k*-medoids methods while retaining the exact the same clustering quality of the basic *k*-medoids algorithm. The application of the new algorithm to road network extraction from classified imagery is also discussed and the preliminary results are encouraging.

1 Introduction

Clustering is the process of grouping a set of objects into classes or clusters so that objects within a cluster have similarity in comparison to one another, but are dissimilar to objects in other clusters [Han *et al* 2001]. In simpler words, clustering is a process of finding natural groupings in a set of data. It has been widely used in the applications such as marketing, city planning, insurance, medicine, chemistry, remote sensing, and so on. A complete review of the current state-of-the-art of spatial clustering can be found in [Han *et al*, 2001] or more recently, in [Guo *et al*, 2003].
In the field of automatic object extraction from remotely-sensed imagery, Doucette *et al* (1999; 2001) provided a self-organizing road map (SORM) approach to road centerline delineation from classified high-resolution Multi-Spectral Imagery (MSI). The SORM is essentially a spatial clustering technique adapted to identify and link elongated regions. This technique is independent from a conventional edge definition, and can meaningfully exploit multispectral imagery. Therefore, it has some promising advantages over other existing methodologies.

Spatial clustering techniques enable the identification and link of elongated road regions. Unfortunately, traditional *k*-means, Kohonen learning approaches, which were used by Doucette *et al* (1999, 2001), are sensitive to the noises in classified images. The *K-Medoids* approach is more robust in this aspect, but it is very time-consuming. We have to investigate more efficient spatial clustering algorithms in order to apply these techniques to a large image. This is the main motivation of this paper.

The remaining parts of this paper are organized as follows: The existing *k*-medoids methods are briefly introduced in the next section. The new efficient *k*-medoids algorithm is presented afterwards with some experiment results. The application of

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 181 – 189, 2005.

the proposed algorithm in road network extraction is discussed and finally some conclusions are given.

2 *K*-Medoids Clustering Methods

The k-medoid method is one of the partitioning methods. Because it uses the most centrally located object (*medoids*) in a cluster to be the cluster centre instead of taking the mean value of the objects in a cluster, it is less sensitive to noise and outliners compared with the *k*-means approach [Han *et al*, 2001]. Therefore, the *k*-medoids method should be more suitable for spatial clustering purpose than the *k*-means method because of the better clustering quality it can achieve. However, it is well known that a *k*-medoids method is very time-consuming. This motivates us to investigate the possibility to improve the efficiency of the *k*-medoids method in the context of road network extraction.

2.1 PAM Algorithm

An early *k*-medoids algorithm called Partitioning Around Medoids (PAM) was proposed by Kaufman and Rousseeuw (1990). The PAM algorithm can be described as follow [Ng and Han, 1994]:

1. Select k representative objects arbitrarily.
2. Compute total cost TC_{ih} for all pairs of objects O_i, O_h where O_i is currently selected, and O_h is not.
3. Select the pair O_i, O_h which corresponds to $min_{Oi, Oh}$ (TC_{ih}). If the minimum TC_{ih} is negative, replace O_i with O_h, and go back to Step (2).
4. Otherwise, for each non-selected object, find the most similar representative object. Halt.

Because of its complexity, PAM works effectively for small data sets (e.g., 100 objects in 5 clusters), but is not that efficient for large data sets [Han *et al*, 2001]. This is not too surprising if we perform a complexity analysis on PAM [Ng and Han, 1994]. In Step 2 and 3, there are altogether k (n-k) pairs of [O_i, O_h], where k is the number of clusters and n is the total number of objects. For each pair, computing TC_{ih} requires the examination of (n-k) non-selected objects. Thus, Step 2 and 3 combined is of $O(k(n-k)^2)$. This is the complexity of only one iteration. Thus, it is obvious that PAM becomes too costly for large values of n and k [Ng and Han, 1994].

2.2 CLARA Algorithm

To deal with larger data sets, a sampling-based method, called Clustering LAR Applications (CLARA) was developed by Kaufman and Rousseeuw (1990). CLARA draws multiple samples of the data set, applies PAM on each sample, and returns its best clustering as the output. The complexity of each iteration now becomes $O(ks^2 + k(n-k))$, where s is the size of the sample, k the number of clusters, and n the total number of objects.

The effectiveness of CLARA depends on the sampling method and the sample size. CLARA cannot find the best clustering if any sampled medoid is not among the

best k medoids. Therefore, it is difficult to determine the sample size. Experiments reported in [Kaufman and Rousseeuw, 1990] indicate that 5 samples of size $40 + 2k$ gave satisfactory results. However, this is only valid for a small k. In the case of road network extraction, we will have hundreds of medoids (i.e., k>100). To assure the quality, the "optimal" sample size has to be about 10 times of k, which will require too much computational time for performing PAM on each sample set.

2.3 CLARANS Algorithm

To improve the quality and scalability of CLARA, another clustering algorithm called Clustering Large Applications based upon RANdomized Search (CLARANS) was proposed in [Ng and Han, 1994]. When searching for a better centre, CLARANS tries to find a better solution by randomly choosing object from the other (n-k) objects. If no better solution is found after a certain number of attempts, the local optimal is assumed to be reached. CLARANS has been experimentally shown to be more efficient than both PAM and CLARA. However, its computational complexity is still about $O(n^2)$ [Han *et al*, 2001], where n is the number of objects. Furthermore, its clustering quality is depending on the two predefined parameters: the maximum number of neighbors examined (*maxneighbor*) and the number of local minima obtained (*numloc*).

In summary, in terms of efficiency and quality, none of the existing k-medoids algorithms is qualified for spatial clustering for road network extraction. We will propose a new efficient k-medoids which can overcome part of the computational problems and is suitable for clustering large data sets.

3 *CLATIN: An Efficient k-Medoids Clustering Method*

When we take a close look at the PAM algorithm, it is found from **Fig.1** that the replacement of the current medoid O_i with the non-medoid object O_h will only affect the small portion of the data set, i.e., the right bottom portion and the left top portion. Therefore, in Step 2 of the PAM clustering, to calculate the total cost of this replacement, we only need to take into account this small portion. The question remaining is then: how to determine efficiently this subset of the original data set?. Thankfully, we can use the triangular network of the medoids to help us find the subset (see **Fig. 2**). This leads us to develop a new efficient k-medoids clustering algorithm. We call it Clustering Large Applications with Triangular Irregular Network (CLATIN).

The main steps of our CLATIN can be described as follows:

1. Initialization.
   1) Select k representative objects arbitrarily as initial medoids.
   2) Construct the TIN of these k medoids.
2. Compute total cost TC_{ih} for all pairs of objects O_i, O_h where O_i is currently selected medoid, and O_h is one of the non-medoid object.
   1) Determine the affected object subset S through a link analysis in the medoid-TIN.

2) Calculate the total cost TC_{ih} over the neighboring object subset S.

3. Select the pair O_i, O_h which corresponds to $min_{Oi, Oh}$ (TC_{ih}). If the minimum TC_{ih} is negative,
   1) replace O_i with O_h,
   2) update the TIN and clustering results locally,
   3) go back to Step 2.
4. Otherwise, for each non-selected object, find the most similar representative object. Halt.

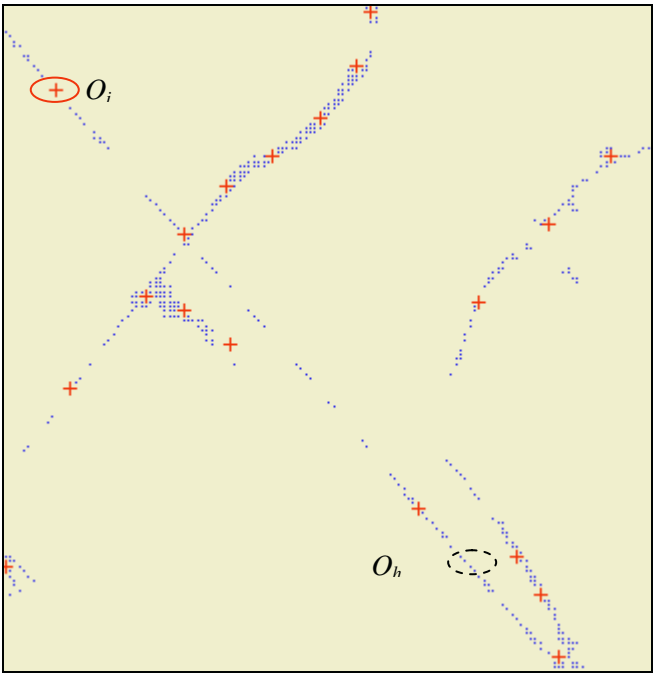


Fig. 1. Illustration of the replacement of current medoid O_i with a non-medoid O_h in a PAM Clustering: Red crosses are current selected medoids, blue points are non-medoids

If we compare the new algorithm with the previous one, there are three additional steps: Step 1.1, Step 2.1 and Step 3.2. These three steps consume a small amount of computational time. In Step 1.1, the TIN can be constructed with a complexity of O($k$log k), where k is the number of nodes; in our case, the number of clusters. Step 2.1 is almost of linear complexity. Step 3.2 is also of linear complexity because we can update the TIN and cluster locally. So in overall, the computation complexity is O(k (n-k) s), where s is the average number of the affected neighbors, which is around 1/10 of n.

There are many existing algorithms that can be chosen for the construction of the TIN of medoids. For example there is the incremental insertion algorithm of Lawson (1977), the divide-and-conquer algorithm of Lee and Schachter (1980), or the plane-

sweep algorithm of Fortune (1987). In our current implementation, we use the incremental insertion algorithm of Lawson (1977), which is also described in [Bourke, 1989]. This algorithm is not worst-case optimal as it achieves $O(n^2)$ time complexity. However, its expected behavior is much better. In average, $O(n \log n)$ time complexity is achieved.

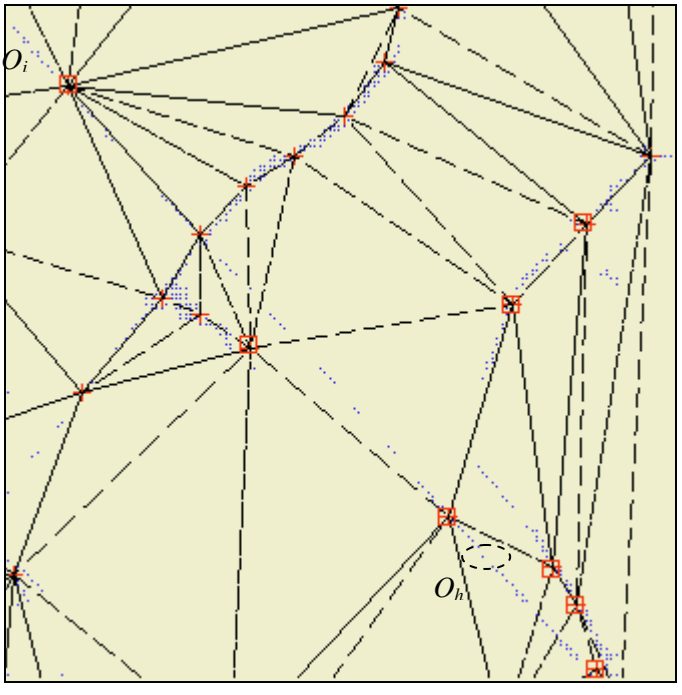


Fig. 2. Illustration of the replacement of current medoid O_i with a non-medoid O_h in a CLATIN clustering: Red crosses are current selected medoids, blue points are non-medoid points, black dash lines are the TIN of current medoids, red rectangles are the possible affected medoids

Table 1 shows the computation time comparison between the new CLATIN approach and other existing *k*-medoids method on different data sets. The tests were performed in a PC with P4 2.02GHz CPU and 512 MB RAM. For CLARA algorithm, we set the number of sample set to 5 and the size of each sample set is 10**k*, where *k* is the number of clusters.

It can be seen that the CLATIN is faster than the basic PAM in all the cases, while retaining exactly the same clustering quality. The larger the number of clusters, the more time saving is achieved (see Fig 3). CLATIN is faster than the CLARA in the cases of small average cluster population (e.g. less than 50 objects per cluster) only (see Fig. 4). This is not surprising because we fixed the sample size as 10 times of the number of clusters for all the cases, and the computation time of CLARA is less sensitive to the number of objects. However, the clustering quality from CLARA usually is not comparable to that of PAM. To improve the clustering quality, we have

Table 1. Computation time comparison

Case	#Sample	#Cluster	Average population	Computation time (seconds)		
				PAM	CLARA	CLATIN
1	1353	9	150.33	13	<1	11
2	3531	9	392.33	111	1	111
3	6000	9	666.67	351	1	364
4	1353	42	32.21	57	27	29
5	3531	42	84.07	411	29	184
6	6000	42	142.86	1,447	35	633
7	3531	120	29.43	1,100	693	229
8	6000	120	50.00	3,251	675	900
9	13531	120	112.76	18,742	707	3,311
10	13531	270	50.11	45,207	7,424	4,261
11	12859	270	47.63	38,778	6,982	2,685
12	10228	270	37.88	25,165	5,780	2,438

to either increase the number of sample sets or increase the size of each sample set. Both cases will result in a dramatic reduction in computation efficiency. This makes it justifiable to choose CLATIN over CLARA for large applications (e.g. spatial clustering for road network extraction).

4 Applications

In this paper, a similar methodology to the SORM proposed by Doucette *et al* (2001) is used. However, we use the new efficient *k*-medoids clustering algorithm: CLATIN.

The proposed approach starts with an image segmentation using spectral clustering techniques. The road cluster is identified automatically using a fuzzy classification based on a set of predefined membership functions for different landscape types (e.g.

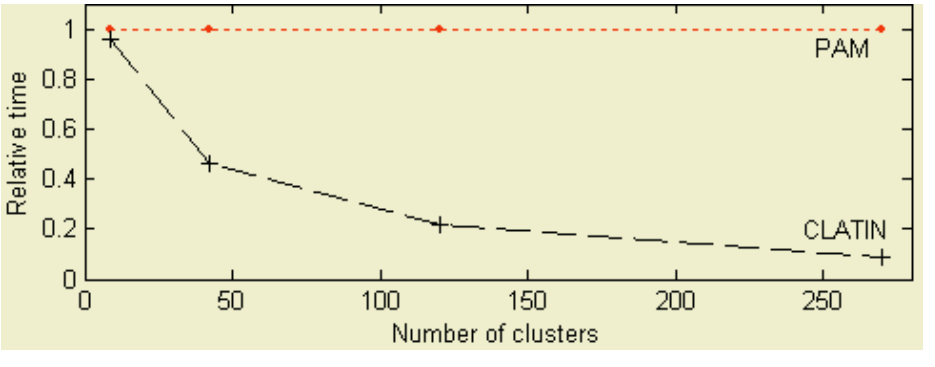


Fig. 3. Comparison of computation time: CLATIN (black dash line) vs. PAM (red dotted)

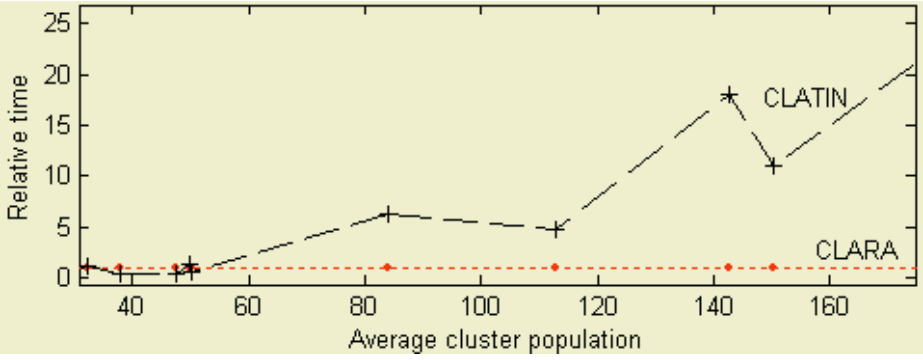


Fig. 4. Comparison of computation time: CLATIN (black dashed) vs. CLARA (red dotted)

water bodies, grasslands, roads). These membership functions are established based n the general spectral signature of each landscape type as well as the characteristics of the image scene under consideration. An angular texture index is used to further reduce the misclassifications between roads, buildings and parking lots. The whole road network is finally extracted by performing a spatial clustering on the road pixels. In this step, the new clustering algorithm is used for spatial clustering purpose and on the road pixels only. The neighboring cluster centers will be linked together to create the corresponding road centerlines and road junctions and to form the road network topology.

The proposed clustering approach for road network extraction has been tested on several IKONOS MS images (4m spatial resolution) covering

Fig. 5. Road center pixels (white) found by CLATIN algorithm (Test area 1)

Fig. 6. Road center pixels (white) found by CLATIN algorithm (Test area 2)

the suburban area of the city of Fredericton, Canada. Some results are shown in Fig. 5 and Fig.6. The results are quite encouraging in the sense that the topology of the whole road network has been precisely captured by the resultant road center pixels (in white in Fig 5 and 6). Note that in our current implementation, there is still a large misclassification among roads, buildings and parking lots which makes the results not that pleasing. However, with the introduction of angular texture index and thus the improvement in image classification, we expect much better results from spatial clustering.

5 Conclusions

Spatial clustering has been a useful tool for Geographical Data Mining. Experiments have also demonstrated that spatial clustering approach for linear feature extraction from remotely-sensed imager has some promising advantages over other existing methodologies because it is independent from a conventional edge definition, and can meaningfully exploit multispectral imagery.

K-medoids methods are very robust to the existence of outliers. This makes it an ideal spatial clustering technique for road network extraction from classified imagery. To overcome the computational issue of existing *k*-medoids methods, we have introduced a new *k*-medoids algorithm, which is much faster than the basic PAM algorithm while retaining exactly the same clustering quality. The new CLATIN algorithm is also more efficient than CLARA in the case of large number of clusters.

The future work includes the full implementation of the CLATIN-based road network extraction from multi-spectral imagery, particularly the improvement of image classification accuracy and the implementation of automatic link of clustering centers to form the whole road network. The applicability of clustering approach to road extraction in the urban area will also be one of our next steps.

Acknowledgements

Financial support from the Canadian NCE GEOIDE research program "Automating photogrammetric processing and data fusion of very high resolution satellite imagery with LIDAR, iFSAR and maps for fast, low-cost and precise 3D urban mapping" is much acknowledged.

References

Bourke, P., 1989. Efficient Triangulation Algorithm Suitable for Terrain Modeling, http: //astronomy.swin.edu.au/~pbourke/terrain/triangulate/, accessed on December 10, 2004.

Doucette, P., Agouris, P., Musavi, M., and Stefanidis, A., 1999. Automated Extraction of Linear Features from Aerial Imagery Using Kohonen Learning and GIS Data. In Agouris, P. and Stefanidis, A. (Eds.), 1999. *Integrated Spatial Databases: Digital Images and GIS*, Lecture Notes in Computer Science, Vol. 1737, pp.20-33, Springer-Verlag: Berlin Heidelberg, 1999.

Doucette, P., Agouris, P., Stefanidis, A., Musavi, M., 2001. Self-Organised Clustering for Road Extraction in Classified Imagery. *ISPRS Journal of Photogrammetry & Remote Sensing*, 55, pp.347-358.

Fortune, S. 1987. A Sweepline Algorithm for Voronoï Diagrams. *Algorithmica*, 2(2), pp153-174.

Guo, D., Peuquet, D., and Gahegan, M., 2003. ICEAGE: Interactive Clustering and Exploration of Large and High-dimensional Geodata, *GeoInformatica*, 7(3), pp229 – 253.

Han, J., Kamber, M. and Tung, A., 2001. Spatial clustering methods in data mining: A survey. In *Geographic Data Mining and Knowledge Discovery* [Miller, H.J, and Han, J., Eds]. London: Taylor & Francis Inc.

Kaufman, L. and Rousseeuw, P.J., 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wilsy & Sons.

Lawson, C.L., 1977. Software for C^1 Surface Interpolation. Mathematical Software III (John R. Rice, editor), pages 161-194. Academic Press: New York.

Lee, D.T. and Schachter, B.J., 1980. Two Algorithms for Constructing a Delaunay Triangulation. *International Journal of Computer and Information Sciences*, 9(3), pp219-242.

Ng, R. and J. Han, 1994, Efficient and Effective Clustering Methods for Spatial Data Mining, *Proc. 20th International Conference on Very Large Databases*, Santiago, Chile.

Security Management for Internet-Based Virtual Presentation of Home Textile Product*

Lie Shi[1], Mingmin Zhang[1], Li Li[2], Lu Ye[1], and Zhigeng Pan[1,2]

[1] School of Computer Science and Technology, Zhejiang University, Hangzhou, 310027, China
sl@insigma.com.cn
yelue03@yahoo.com.cn
[2] Institute of VR and Multimedia, HZIEE, Hangzhou, 310037, China
{lili, zgpan, zmm}@cad.zju.edu.cn

Abstract. Internet and E-commerce technology are two fast developing fields. Companies who adopt these new techniques have gained good economic profit. In this paper, we will focus on the application of these techniques in the virtual presentation and sale of textile product. The paper will address the following technical problems in detail: Internet-based garment CAD/electronic garment fitting, virtual presentation and design of home textile product, security problems in electronic commerce, and copyright protection of textile patterns on Internet.

1 Introduction

The garment industry in China has a very long history. In 1960's, with the development of computers, industrial automation has found application in garment and textile manufacturing. USA, England, France, Italy, Spain, and Japan et al employed the new technique and gained a lot. In 1980's, a lot of garment companies were established all over China. In Seventh-five Plan, garment CAD was regarded as the key developing project, which had a great promotion to the garment design, manufacturing, et al. Computers have been applied to various aspects of garment industry [1],[2],[3]. In the beginning of 1990's, with the promotion of human life level and the pursuit of high living quality, various home textile products appeared in the market. And a lot of companies producing textile product were established in last decade. Beyond Company is a typical example.

With the development of software technology, hardware technology, and Internet, there will be certainly more application systems in the filed of textile and garment industry. And these new techniques will bring a lot of economic profits for companies in textile and garment fields.

* This research is supported by NSF in Zhejiang Province(grant no.KYZ070304005).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 190–196, 2005.

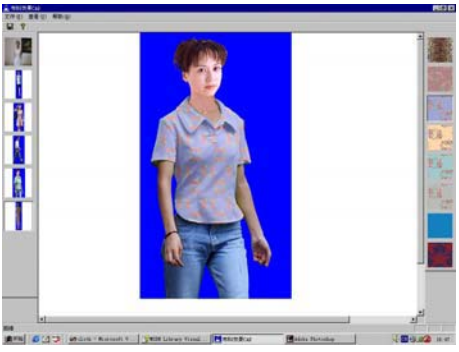


Fig. 1. User Interface of our EasyShow system

Fig. 2. The Replacement Effects of Materials for Skirts

Several Universities and institutions have do some research work in electronic garment fitting [4],[5],[6],[7],[8]. Even there are some electronic garment fitting systems in the market, but they were not widely used. There are four reasons:

- The user interface of the systems is not intuitive and friendly, and it also lacks intelligence,
- The effect for garment fitting is not as good as expectedtexture feature and color feature can not reach realistic feeling,
- They lack efficient body-varying tool, since each individual has its own body parameters.

To solve these problems, we develop a new generation electronic garment-fitting system (Fig. 1 shows the user interface). The system can make replacement to the cloth materials, cloth color. And in the replaced picture, the shadow changes and crapy effect can be preserved very well (See Fig. 2).

EasyShow is a visual presentation and computer aided design system for home textile, developed by State Key Lab. of Zhejiang University and Pioneer Graphics Software Corp. It can display the layout of a room with all textile

Fig. 3. Visual presentation of sofa slipcover

adornment (such as slipcover of sofabedspreadcarpetcurtainrag doll). For each textile adornment, the system can change the color of cloth with the design style unchanged. In addition, users can input their own image patterns into image library, change the color of style with their own patterns(as Fig. 3).

According to our investigation, many textile mills and color separation press factories have a lot of cloth styles and image patterns, which are useful to the design of new textile products. It is very important to create a cloth style and image pattern library, compress the library data, and save the information in computer database. Then the users can search for a cloth style and image pattern quickly and easily. The system is being developed, which can be used soon.

Firstly, electronic commerce over the Internet fundamentally depends on the communication security services. In the commercial world, security problems manifest themselves in a number of ways, such as eavesdropping, password sniffing, data modification, spoofing and repudiation. However, various simple holes exist in our software and systems and basic flaws exist in the Internet's infrastructure. Several studies have independently shown that many individuals and companies are abstaining from joining the Internet simply because of security concerns.

Secondly, although the network distribution of multimedia information offers the promise of improved efficiency and much lowered distribution costs, the undesirable side effects are the enormous business threats to information provider: the unauthorized redistribution of the copyrighted multimedia data. Protecting content against misappropriation or abuse is a key step toward a comprehensive information commerce infrastructure.

Thirdly, fraud exists in current commerce systems: cash can be counterfeited, checks altered, credit card numbers stolen, even our identities can be forged. There is an urgent need for technologies and infrastructures that can forestall spoofing and guarantee confidentiality in Web commerce. There are several schemes now available to solve these problems to some extent: firewalls, one-time passwords, data encryption, digital signatures, public-key certificates, cryptographic protocols etc. However, a robust security solution satisfies all the fundamental requirements for Internet-based e-commerce is still unavailable. In this following, we propose a possible solution based on digital steganography and watermarking for transaction. And we present our method for copyright protection on clothing styles, flower pattern and design models.

2 Security Management in EasyShow

2.1 A Secure Transition Model Based on Steganography and Watermarking

Steganography deals with the hiding of messages so the potential monitors do not even know that a message is being sent. It is different from cryptography where they know that a secret message is being sent. The latter might consist of an unreadable message like FOKLYA, while the former might use secret ink[9].

According to the consensus on the First International Information Hiding Workshop, a slightly changed model of core components in a typical steganographic system, also called as stegosystem, is shown in Fig. 4. In such systems, data is whatever types of data that's appropriate, such as various multimedia data. Embedded-data represents messages to be hidden. Cover-data is the input that will be the carrier of the embedded-data. Stego-data, which has the embedded message hidden in it, is the output of the hiding process.

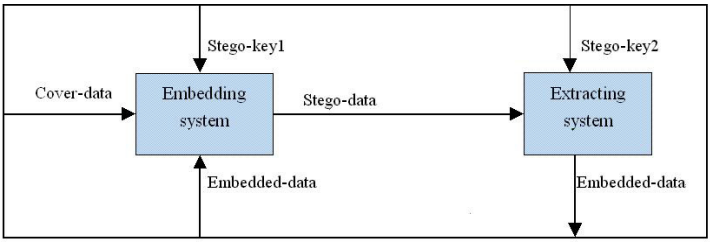


Fig. 4. Model of core components of stegosystems

Stego-key is the additional secret data that may be needed in the steganographic process. Stego-key1 is used in the embedding process while stego-key2 is used in the extracting process.

Compared with the cryptographic terminology, if stego-key2 is the same as stego-key1, we call the system symmetric, otherwise asymmetric. In a symmetric system, the stego-key is also referred to as private key. And in an asymmetric system, one key is called private, the other is called public.

Steganography conceals the very existence of the secret message. It's therefore broader than cryptography and can tackle some problems that are not solvable by cryptography. Especially it provides an effective way to implement secure communications, or in another word, secret communications[10]. Digital watermarking stems from steganography, with robust embedded labels as a key feature. The label, called as a digital watermark, is a piece of hidden information within digital data (such as audio, video, still images, text and 3D models). To integrate the above ideas we propose a model that incorporates digital steganography and watermarking into an online transaction. The model is depicted in Fig. 5.

Fig. 5. Watermark image

Our model meets the following requirements for secure transactions:

Security: The transaction data is embedded into an ordinary-looking multimedia document being transmitted. It achieves the secure transaction desired because ordinary documents are unlikely to be eavesdropped or to stimulate the desire of spies to attack. Concealment is the first layer of security. The transaction data should be robustly embedded, in another word, it is difficult to be extracted and damaged without the proper key and survives common transformations and transmission errors. And only Bob can extract the transaction data from the stego-document with his private key. Robust steganography is the second layer of security.

Robust: The watermark must be difficult (hopefully impossible) to remove or destroy. Acts of removal can be either unintentional, like common signal processing operations, or intentional, like deliberate and malicious attacks such as collusion and forgery.

Secure: The digital watermark is secure, i.e., the locations where the watermark is embedded and how much modification occurs in order to embed the watermark are secret and undetectable without secret elements.

Our blind digital watermarking algorithm is based on FFT frequency domain. The original image is color image of 512×512. The watermark image is binary image of 64×64. Watermark embedment is taken as formula (1). Watermark extraction is also described as following.

(1) Embedding watermarks

Step 1. Take Fourier Transform of R,G,B components of original color image of size 512×512 and get Fourier magnitude matrixes A_i and angle matrix $Z_i, i = 1, 2, 3$. If the size of the original color image is more than 512×512, we only select one part of it to 512×512. The watermark image is binary and every pixel is 0 or 1. We take every pixel of the watermark image and get a sequence of 0 and 1. The sequence is divided into three groups and we can get $B_i, i = 1, 2, 3$, The total 0 in B_i is C_i,$i = 1, 2, 3$.

Step 2. Embed B_i into some fixed positions of A_i and get $D_i, i = 1, 2, 3$. The watermark embedment is taken as following formula (1):

$$v'_j = v_j + \alpha w_j, \qquad \alpha \gg v_j \tag{1}$$

$w_j \in B_i$, w_j=0 or 1, v_i is chosen data from,$i = 1, 2, 3$. strength parameter $\alpha = 12000$.

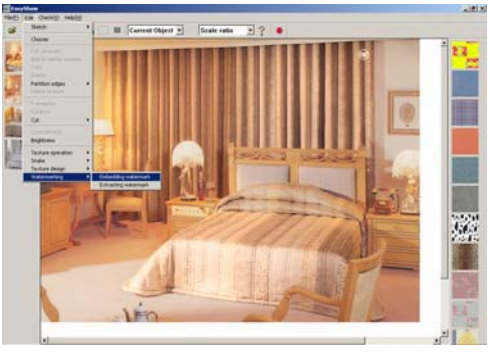


Fig. 6. Embedding watermark interface of Easy Show system

Step 3. Invert FFT transform to D_i and Z_i and get $\ddot{A}_i$, $\ddot{A}_i$ is R,G,B components of the image-signed, $i = 1, 2, 3$. So, we get the image-signed I_W.

(2) Extracting watermarks

Step 1. Take R,G,B components of the image-signed I_W and get A_i, magnify the noise of A_i and get $\ddot{A}_i$ using method mentioned in Section 2.1, $\beta = 30$, $i = 1, 2, 3$.

Step 2. Take Fourier Transform of A_i and get B_i,$i = 1, 2, 3$. Find watermark information from the corresponding positions of B_i and get sequence $E_i, i = 1, 2, 3$. Turn C_i minmum of E_i into 0 and turn others into 1 and get sequence $F_i, i = 1, 2, 3$.

Step 3. Let $F = \{F_i\}, i = 1, 2, 3$. Turn F into 64×64 blocks and get the watermark image covered W.

(3) Experimental Results

The watermark image is the binary image (see Fig. 5). We embed it into Easy Show system to protect copyright of textile products (see Fig. 6).]

3 Conclusion

The introduction of new information technology will accelerate the development of textile and clothing industry, and it is a very good chance for the companies in this field to have revolution on traditional manual design and manufacturing methods. The Internet-based electronic commerce provides good chance for the fast development for the companies in textile field. Besides establishing web site, the companies can also set-up a virtual product presentation studio, supporting electronic clothing-fitting of human beings on Internet and the interactive design and customizing of home-textile product through Internet.

Acknowledgements

This work is co-supported by National NSF Project (grant no: 60473111), the Excellent Young Teacher Awarding Fund of MOE in China.

References

1. Pan Y.H., et al.: CAD— Techinques and PrincipleZhejiang University Press(1996)
2. Wang J.: Costumed/Cloth Effect Simulation Based on Image, Master Thesis, Zhejiang University(1999)
3. Wang J., Pan Z.G., Ye C.Q.: An Image Morphing Algorithm based on Contour Polygon, Computer Research and Development, **10** (1999) 1358–1363
4. Wang J., Pan Z.G.: An New Algorithm Implementing Natural Texture Mapping. Journal of CAD and Graphics. **3** (2000) 161–165
5. Zhan S.Y., Jiang F.Y.: Intelligent Flower Edge CADMechanical Electronic Engineering**5** (1999) 144–145
6. Hing N.Ng.et al, and Grimsdate R.,L, Computer graphics techniques for modeling cloth, IEEE Computer graphics and applications, Sept 1996 .28-41
7. Volino P., et al.: An Evolving system for simulating clothes on virtual actors. IEEE Computer graphics and applications. (1996) 42–51
8. Yang Y.,Thalmann N.M.: An improved algorithm for collision deletion in cloth animation with human body. In: Proc of Pacific Graphics, Singapore.(1993) 237–251
9. Yin K.K., Pan Z.G., Shi J.Y.: A Secure Transaction Model Based on Digital Stenography and Watermarking. CAD/Graphics'99. (1999)
10. Li, L., Zhang, D.P., Shi, J.Y., et, al.: Watermarking 3D Mesh by Spherical Parameterization, Computer & Graphics, Vol.28, No.6, (2004) 981–989

An Efficient Approach for Surface Creation

L.H. You and Jian J. Zhang*

National Centre for Computer Animation,
Bournemouth University, UK
{lyou, jzhang}@bournemouth.ac.uk

Abstract. In this paper, we present an efficient method for generation of free from surfaces using the solution to a fourth order partial differential equation. In the interest of computational efficiency, the surface function is taken to be the combination of the boundary functions modulated by some unknown functions. Making use of the properties of boundary functions, the fourth order partial differential equation is transformed into a fourth order ordinary differential equation. To solve this equation, we further convert it to a set of one-dimensional finite difference equations where the number of unknowns is reduced significantly allowing fast surface generation.

1 Introduction

Surface generation is an important topic in computer graphics and CAD. Many methods have been developed. Among current techniques, Bézier, B-spline and NURBS are the most popular [1].

Alongside the purely geometric methods in surface modelling, complementary approaches are being developed. For example, Sederberg and Parry proposed a free form deformation (FFD) tool for shape manipulation of surfaces [2]. Coquillart extended free form deformations to EFFD [3].

PDE (partial different equation) based surface generation is another effective approach, which was originally proposed by Bloor and Wilson [4], [5]. Later on, this approach was extended and applied to various developments, including vase surface design [6], blending surface generation [7] and surface modelling [8]. Most recently, Monterde and Ugail investigated harmonic and biharmonic Bézier surfaces [9].

Fourth order partial differential equations used in surface representation are often difficult to solve with analytical methods, if not impossible. In many cases numerical methods, such as finite differencing and finite element methods, are the only option. As two independent variables are necessary for surface representation, the resolution of fourth order PDEs is a two dimensional problem involving many unknowns to be solved. Consequently, it is expensive both in terms of processing speed and memory consumption. In order to address this

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 197–205, 2005.

problem, in this paper, we present an efficient resolution method to solve fourth order PDEs and employ it to generate free form surfaces.

2 Finite Difference Algorithm

Subject to suitably defined boundary conditions, the solution to a vector-valued fourth order PDE can be employed to represent a free form surface [6], [7]. In this paper, we use the following partial differential equation

$$\frac{\partial^4 \mathbf{x}}{\partial u^4} + \mathbf{b}\frac{\partial^4 \mathbf{x}}{\partial u^2 \partial v^2} + \mathbf{c}\frac{\partial^4 \mathbf{x}}{\partial v^4} = \mathbf{0} \tag{1}$$

where $\mathbf{b} = [b_x\ b_y\ b_z]^T$ and $\mathbf{c} = [c_x\ c_y\ c_z]^T$ are vector-valued shape parameters, and $\mathbf{x} = [x\ y\ z]^T$ is a vector-formed positional function.

Given two non-coincident 3D curves, we can create a surface which passes these two curves. The slope of the surface can be described with the tangents of the surface at these two curves. Therefore, the boundary conditions for the surface can be written as

$$\begin{array}{l} u = 0\ \mathbf{x} = \mathbf{g}_0(v)\ \frac{\partial \mathbf{x}}{\partial u} = \mathbf{g}_1(v) \\ u = 1\ \mathbf{x} = \mathbf{g}_2(v)\ \frac{\partial \mathbf{x}}{\partial u} = \mathbf{g}_3(v) \end{array} \tag{2}$$

where $\mathbf{g}_0(v)$ and $\mathbf{g}_2(v)$ are called the boundary curves, and $\mathbf{g}_1(v)$ and $\mathbf{g}_3(v)$ are called the boundary tangents.

The closed form solution of PDE (1) under the boundary conditions (2) usually does not exist. Among various numerical methods, the finite difference approach is a common one. However, it is very expensive if this method is to be directly applied to Eq. (1), as explained earlier. Below we present a new method to get around the problem.

Since the surface to be generated must satisfy the boundary conditions, we can construct a surface function which is a combination of some unknown functions and the boundary functions given in Eq. (2). It takes the form of

$$\mathbf{x}(u, v) = \sum_{i=0}^{3} \mathbf{f}_i(u)\mathbf{g}_i(v) \tag{3}$$

where $\mathbf{f}_i(u)(i = 0, 1, 2, 3)$ are unknown functions.

Substituting the constructed surface function (3) into PDE (1), the following differential equation results

$$\sum_{i=0}^{3} \left[\frac{d^4 \mathbf{f}_i(u)}{du^4}\mathbf{g}_i(v) + \mathbf{b}\frac{d^2 \mathbf{f}_i(u)}{du^2}\frac{d^2 \mathbf{g}_i(v)}{dv^2} + \mathbf{c}\mathbf{f}_i(u)\frac{d^4 \mathbf{g}_i(v)}{dv^4} \right] = \mathbf{0} \tag{4}$$

Substituting the constructed surface function (3) into boundary conditions (2), a series of new boundary conditions can be obtained.

For the boundary condition $\mathbf{x} = \mathbf{g}_0(v)$ at $u = 0$, the insertion of surface function (3) into (2) leads to

$$[\mathbf{f}_0(u)\mathbf{g}_0(v) + \mathbf{f}_1(u)\mathbf{g}_1(v) + \mathbf{f}_2(u)\mathbf{g}_2(v) + \mathbf{f}_3(u)\mathbf{g}_3(v)]_{u=0} = \mathbf{g}_0(v) \tag{5}$$

Eq. (5) can be satisfied when the following conditions hold

$$\begin{aligned} \mathbf{f}_0(u)\,|_{u=0} &= 1 \\ \mathbf{f}_1(u)\,|_{u=0} &= 0 \\ \mathbf{f}_2(u)\,|_{u=0} &= 0 \\ \mathbf{f}_3(u)\,|_{u=0} &= 0 \end{aligned} \tag{6}$$

Similarly, we can get the following boundary conditions.
For the boundary condition $\frac{\partial \mathbf{x}}{\partial u} = \mathbf{g}_1(v)$ at $u = 0$,

$$\left[\frac{d\mathbf{f}_0(u)}{du}\mathbf{g}_0(v) + \frac{d\mathbf{f}_1(u)}{du}\mathbf{g}_1(v) + \frac{d\mathbf{f}_2(u)}{du}\mathbf{g}_2(v) + \frac{d\mathbf{f}_3(u)}{du}\mathbf{g}_3(v)\right]_{u=0} = \mathbf{g}_1(v) \tag{7}$$

$$\begin{aligned} \tfrac{d\mathbf{f}_0(u)}{du}\,|_{u=0} &= 0 \\ \tfrac{d\mathbf{f}_1(u)}{du}\,|_{u=0} &= 1 \\ \tfrac{d\mathbf{f}_2(u)}{du}\,|_{u=0} &= 0 \\ \tfrac{d\mathbf{f}_3(u)}{du}\,|_{u=0} &= 0 \end{aligned} \tag{8}$$

For the boundary condition $\mathbf{x} = \mathbf{g}_2(v)$ at $u = 1$,

$$[\mathbf{f}_0(u)\mathbf{g}_0(v) + \mathbf{f}_1(u)\mathbf{g}_1(v) + \mathbf{f}_2(u)\mathbf{g}_2(v) + \mathbf{f}_3(u)\mathbf{g}_3(v)]_{u=1} = \mathbf{g}_2(v) \tag{9}$$

$$\begin{aligned} \mathbf{f}_0(u)\,|_{u=1} &= 0 \\ \mathbf{f}_1(u)\,|_{u=1} &= 0 \\ \mathbf{f}_2(u)\,|_{u=1} &= 1 \\ \mathbf{f}_3(u)\,|_{u=1} &= 0 \end{aligned} \tag{10}$$

For the boundary condition $\frac{\partial \mathbf{x}}{\partial u} = \mathbf{g}_3(v)$ at $u = 1$,

$$\left[\frac{d\mathbf{f}_0(u)}{du}\mathbf{g}_0(v) + \frac{d\mathbf{f}_1(u)}{du}\mathbf{g}_1(v) + \frac{d\mathbf{f}_2(u)}{du}\mathbf{g}_2(v) + \frac{d\mathbf{f}_3(u)}{du}\mathbf{g}_3(v)\right]_{u=1} = \mathbf{g}_3(v) \tag{11}$$

$$\begin{aligned} \tfrac{d\mathbf{f}_0(u)}{du}\,|_{u=1} &= 0 \\ \tfrac{d\mathbf{f}_1(u)}{du}\,|_{u=1} &= 0 \\ \tfrac{d\mathbf{f}_2(u)}{du}\,|_{u=1} &= 0 \\ \tfrac{d\mathbf{f}_3(u)}{du}\,|_{u=1} &= 1 \end{aligned} \tag{12}$$

After the above treatment, the resolution of PDE (1) subject to boundary conditions (2) is transformed to solving Eq. (4) subject to boundary conditions (6), (8), (10) and (12).

Observing Eq. (4), if the equation in the square bracket is zeroed, their sum is also zero. Thus a solution of Eq. (4) is found. In this way, we obtain four differential equations and their corresponding boundary conditions.

$$\begin{aligned} &\tfrac{d^4\mathbf{f}_i(u)}{du^4}\mathbf{g}_i(v) + \mathbf{b}\tfrac{d^2\mathbf{f}_i(u)}{du^2}\tfrac{d^2\mathbf{g}_i(v)}{dv^2} + \mathbf{c}\mathbf{f}_i(u)\tfrac{d^4\mathbf{g}_i(v)}{dv^4} = \mathbf{0} \\ &(i = 0, 1, 2, 3) \end{aligned} \tag{13}$$

$$\begin{aligned}
&\mathbf{f}_i(u)|_{u=0} = \mathbf{c}_{i1} \quad \left.\frac{d\mathbf{f}_i(u)}{du}\right|_{u=0} = \mathbf{c}_{i2} \\
&\mathbf{f}_i(u)|_{u=1} = \mathbf{c}_{i3} \quad \left.\frac{d\mathbf{f}_i(u)}{du}\right|_{u=1} = \mathbf{c}_{i4} \\
&(i = 0, 1, 2, 3)
\end{aligned} \tag{14}$$

In the above equation, the known constants $\mathbf{c}_{ij}$ is determined below

$$\mathbf{c}_{ij} = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases} \tag{15}$$

If the second and fourth derivatives of the boundary functions $\mathbf{g}_i(v)$ $(i = 0, 1, 2, 3)$ can be expressed with the functions themselves, we have

$$\begin{aligned}
&\frac{d\mathbf{g}_i^2(v)}{dv^2} = (-1)^m \xi_i^2 \mathbf{g}_i(v) \\
&\frac{d\mathbf{g}_i^4(v)}{dv^4} = (-1)^{2m} \xi_i^4 \mathbf{g}_i(v) \\
&(i = 0, 1, 2, 3)
\end{aligned} \tag{16}$$

where ξ_i $(i = 0, 1, 2, 3)$ are known constants determined by the differential operation, and m may take the value of 1 or 2 depending on the properties of the boundary function.

For those boundary functions whose second and fourth derivatives cannot be represented by the functions themselves, we can convert them into those which have the properties given by Eq. (16).

The insertion of Eq. (16) into (13) results in a fourth order ordinary differential equation as follows

$$\begin{aligned}
&\frac{d^4\mathbf{f}_i(u)}{du^4} + (-1)^m \mathbf{b} \xi_i^2 \frac{d^2\mathbf{f}_i(u)}{du^2} + (-1)^{2m} \mathbf{c} \xi_i^4 \mathbf{f}_i(u) = \mathbf{0} \\
&(i = 0, 1, 2, 3)
\end{aligned} \tag{17}$$

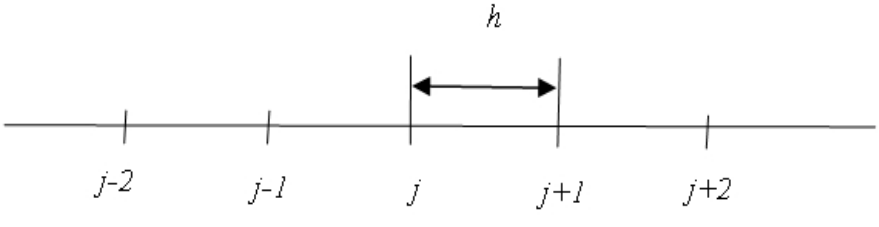


Fig. 1. Typical node j

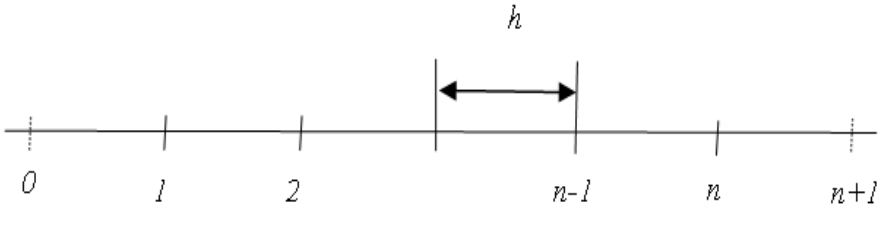


Fig. 2. Node collocation

Central finite difference approximation of the second and fourth derivatives in Eq. (17) can be written below using the nodes shown in Fig. 1.

$$\begin{aligned}\left(\frac{d^2\mathbf{f}_i}{du^2}\right)_j &= \frac{\mathbf{f}_i^{j+1}+\mathbf{f}_i^{j-1}-2\mathbf{f}_i^j}{h^2}\\ \left(\frac{d^4\mathbf{f}_i}{du^4}\right)_j &= \frac{6\mathbf{f}_i^j-4(\mathbf{f}_i^{j+1}+\mathbf{f}_i^{j-1})+\mathbf{f}_i^{j+2}+\mathbf{f}_i^{j-2}}{h^4}\end{aligned} \tag{18}$$

Substituting Eq. (18) into (17), and putting together the terms with the same node index, the finite difference equation at the typical node j is generated.

$$\begin{aligned}&\mathbf{f}_i^{j-2}+\left[-4+(-1)^m\mathbf{b}h^2\xi_i^2\right]\mathbf{f}_i^{j-1}+\left[6-(-1)^m 2\mathbf{b}h^2\xi_i^2+(-1)^{2m}\mathbf{c}h^4\xi_i^4\right]\mathbf{f}_i^j\\ &+\left[-4+(-1)^m\mathbf{b}h^2\xi_i^2\right]\mathbf{f}_i^{j+1}+\mathbf{f}_i^{j+2}=0\\ &(i=0,1,2,3)\end{aligned} \tag{19}$$

Assuming the resolution region in u direction to be $0 \le u \le 1$, and uniformly dividing this region into $n-1$ equal intervals, we get nodes shown in Fig. 2 where the numbers 1 and n are boundary nodes, 0 and $n+1$ are imaginary nodes outside the boundaries.

With the nodes shown in Fig. 2, boundary conditions (14) become

$$\begin{aligned}&\mathbf{f}_i^1=\mathbf{c}_{i1}\\ &\mathbf{f}_i^2-\mathbf{f}_i^0=2h\mathbf{c}_{i2}\\ &\mathbf{f}_i^n=\mathbf{c}_{i3}\\ &\mathbf{f}_i^{n+1}-\mathbf{f}_i^{n-1}=2h\mathbf{c}_{i4}\end{aligned} \tag{20}$$

For function $\mathbf{f}_i$, its value at each node is unknown. There are $n+2$ nodes in Fig 2. Therefore, the total unknown constants are $n+2$. For the inner nodes 2, 3, 4,$\cdots$, $n-1$, $n-2$ linear algebraic equations can be constructed from Eq. (19). Plus the four linear equations given by Eq. (20), there are $n+2$ linear algebraic equations in total whose resolution determines $n+2$ unknown constants. Then substituting these function values back into Eq. (3), the constructed surface function can be used to generate the resulting free form surface.

3 Numerical Applications

In this section, we will give a number of examples to demonstrate the applications of the proposed method in surface generation.

Firstly, using a example we demonstrate how different node numbers affect the surface quality. The boundary conditions for this example are

$$\begin{aligned}u=0\ x&=r_0\cos 2\pi v & \frac{\partial x}{\partial u}&=r_0'\cos 2\pi v\\ y&=r_0\sin 2\pi v & \frac{\partial y}{\partial u}&=r_0'\sin 2\pi v\\ z&=h_0 & \frac{\partial z}{\partial u}&=h_0'\\ u=1\ x&=r_1\cos 2\pi v & \frac{\partial x}{\partial u}&=r_1'\cos 2\pi v\\ y&=r_1\sin 2\pi v & \frac{\partial y}{\partial u}&=r_1'\sin 2\pi v\\ z&=h_1 & \frac{\partial z}{\partial u}&=h_1'\end{aligned} \tag{21}$$

The boundary functions are $g_{0x}(v) = r_0 \cos 2\pi v$, $g_{1x}(v) = r_0' \cos 2\pi v$, $g_{2x}(v) = r_1 \cos 2\pi v$ and $g_{3x}(v) = r_1' \cos 2\pi v$ for x component, $g_{0y}(v) = r_0 \sin 2\pi v$, $g_{1y}(v) = r_0' \sin 2\pi v$, $g_{2y}(v) = r_1 \sin 2\pi v$ and $g_{3y}(v) = r_1' \sin 2\pi v$ for y component, and $g_{0z}(v) = h_0$, $g_{1z}(v) = h_0'$, $g_{2z}(v) = h_1$ and $g_{3z}(v) = h_1'$ for z component.

Substituting these boundary functions into Eq. (16), we find that $m = 1$ and $\xi_i = 2\pi$ $(i = 0, 1, 2, 3)$ for x and y components, and $\xi_i = 0$ $(i = 0, 1, 2, 3)$ for z component.

Taking $r_0 = 0.6$, $r_0' = -2$, $r_1 = 0.8$, $r_1' = 1$, $h_0 = 3$ and $h_0' = h_1 = h_1' = 0$, solving the finite difference equations for x, y and z components, respectively, and using the solution to generate the surface, we obtain the images shown in Fig. 3 where the left, middle and right ones are from $n = 20$, $n = 40$ and $n = 60$, respectively. With the image magnitude given in the figure, only when node number is greater than 60, the boundaries between two adjacent patches cannot be observed.

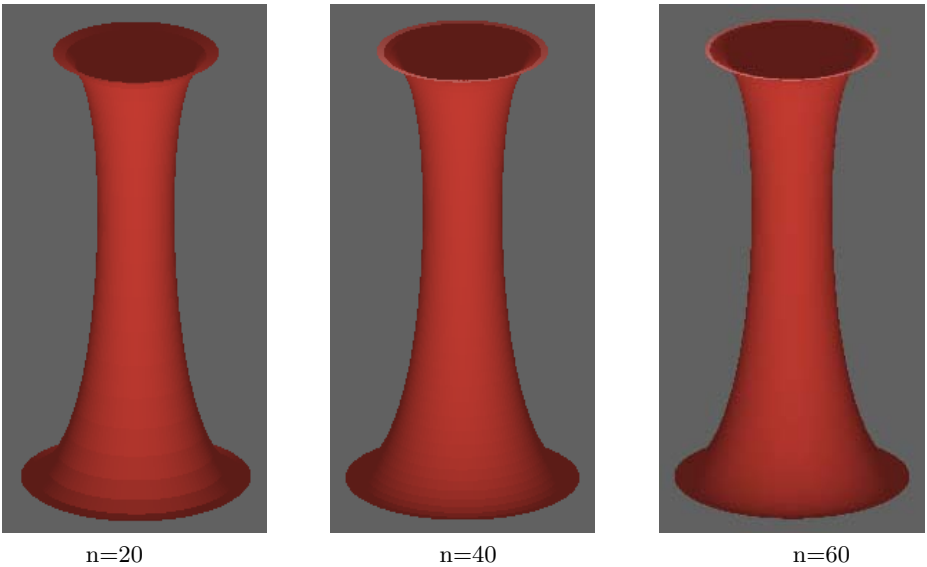


Fig. 3. Surface generated with different node numbers

When the nodes are taken to be 60, only 62 linear algebraic equations are solved with the method given in this paper. However, if Eq. (1) is solved directly with the finite difference method, $62 \times 62 = 3844$ linear algebraic equations must be solved leading to a much more expensive resolution process. Therefore, the method given in this paper is very efficient and especially suitable for the applications requiring real-time performance.

With the same boundary functions (21) but different geometric parameters and vector-valued parameters, more shapes are generated and depicted in Fig. 4.

The second example is to generate a free form surface defined by the following boundary conditions

$$\begin{array}{lll} u=0 & x=0.5\sinh(0.3+1.6v) & \frac{\partial x}{\partial u}=0.5\eta_1\sinh(0.3+1.6v) \\ & y=1.6\cosh(0.6-1.6v) & \frac{\partial y}{\partial u}=1.6\eta_2\cosh(0.6-1.6v) \\ & z=0 & \frac{\partial z}{\partial u}=0 \\ u=1 & x=-1.8\sinh(0.8-1.5v) & \frac{\partial x}{\partial u}=-1.8\eta_3\sinh(0.8-1.5v) \\ & y=1.6\cosh(0.2+1.5v) & \frac{\partial y}{\partial u}=1.6\eta_4\cosh(0.2+1.5v) \\ & z=2 & \frac{\partial z}{\partial u}=0 \end{array} \tag{22}$$

For the above boundary functions, we found that $m = 2$ and $\xi_0 = \xi_1 = 1.6$ and $\xi_2 = \xi_3 = -1.5$ for x component, $\xi_0 = \xi_1 = -1.6$ and $\xi_2 = \xi_3 = 1.5$ for y component, and $\xi_i = 0$ $(i = 0, 1, 2, 3)$ for z component.

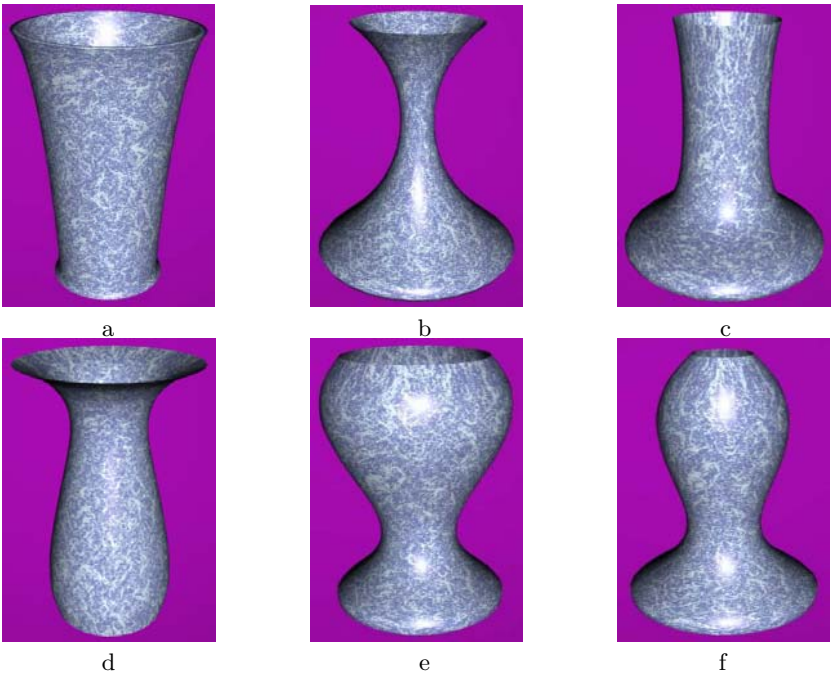
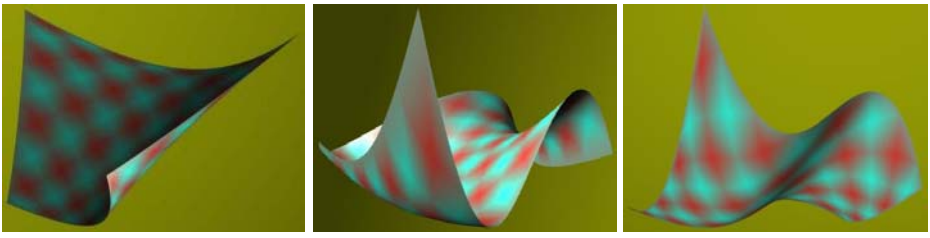


Fig. 4. Generation of surfaces with different shapes

Setting all the vector-valued parameters to 1 and $\eta_i (i = 1, 2, 3, 4)$ to 0, the surface in Fig. 5a is created. Changing η_1 and η_3 to -2 and η_2 and η_4 to 3, the image in Fig. 5b is obtained. Keeping $\eta_i (i = 1, 2, 3, 4)$ the same as those for Fig. 5a and setting $b_x = b_y = b_z = 50$, the surface in Fig. 5c is produced.

The third example is to demonstrate how to create free-form surfaces whose boundary conditions do not satisfy Eq. (16). The positional boundary conditions for this example are



a b c

Fig. 5. Generation of free-form surfaces from accurate boundary conditions

a b

Fig. 6. Generation of free-form surfaces from approximate boundary conditions

$$\begin{aligned} w=0\ x &= -30.01 - 54.39v + 186.71v^2 - 106.09v^3 \\ y &= 5 - 28.62v + 32.12v^2 - 2.5v^3 \\ z &= 3.25 - 26.72v + 17.58v^2 + 5.49v^3 \\ w=1\ x &= -11.1 - 51.07v + 185.38v^2 - 120.05v^3 \\ y &= -3 + 3.06v + 35.77v^2 - 40.83v^3 \\ z &= 14.9 - 28.02v + 68.08v^2 - 45.96v^3 \end{aligned} \tag{23}$$

The functions in Eq. (23) do not satisfy Eq. (16). Here we use a series of cosine functions to transform boundary conditions (23) into those that satisfy Eq. (16). For illustrative purpose, we only take four terms in the cosine series and set tangential boundary conditions to the corresponding positional boundary conditions multiplied by a known coefficient (in this example, it is taken to be 0.1). In doing so, the following boundary conditions are obtained.

$$\begin{aligned} w=0\ \ x &= x_0(v) = -21.2 - 14.4\cos\pi v + 4.3\cos(2\pi v) + 1.3\cos(3\pi v) & \frac{\partial x}{\partial w} &= 0.1x_0(v) \\ y &= y_0(v) = 1.2 - 0.6\cos\pi v + 4.3\cos(2\pi v) + 0.07\cos(3\pi v) & \frac{\partial y}{\partial w} &= 0.1y_0(v) \\ z &= z_0(v) = -2.6 + 1.8\cos\pi v + 4\cos(2\pi v) + 0.06\cos(3\pi v) & \frac{\partial z}{\partial w} &= 0.1z_0(v) \\ w=1\ \ x &= x_1(v) = -4.8 - 9.5\cos\pi v + 0.8\cos(2\pi v) + 2.3\cos(3\pi v) & \frac{\partial x}{\partial w} &= 0.1x_1(v) \\ y &= y_1(v) = -0.01 - 0.06\cos\pi v - 4\cos(2\pi v) + 1.1\cos(3\pi v) & \frac{\partial y}{\partial w} &= 0.1y_1(v) \\ z &= z_1(v) = 12.1 + 1.4\cos\pi v - 0.13\cos(2\pi v) + 1.5\cos(3\pi v) & \frac{\partial z}{\partial w} &= 0.1z_1(v) \end{aligned} \tag{24}$$

It should be pointed out that since only four terms are taken, the positional boundary conditions are a rough approximation of Eq. (23). For a good approximation, more terms should be used.

For each term of Eq. (24), the above method is employed to determine the surface function and the created surface was given in Fig. 6a and 6b where both images are obtained from the same surface but different view angles.

4 Conclusions

Fast surface generation is an important topic for applications requiring real-time performance such as real-time animation, virtual environments and computer games. Analytical PDE based surface generation is only applicable to a small range of problems. Numerical PDE based surface modelling demands large amounts of processing power and memory.

The method presented in this paper transforms a fourth order partial differential equation into a fourth order ordinary differential equation. Then finite difference approximation is applied leading to a small set of one-dimensional finite difference equations. Compared to two-dimensional finite difference equations for solving a fourth order PDE, the one-dimensional finite difference approach involves significantly fewer unknowns and is almost two orders of magnitude faster.

References

1. Farin, G.: Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide. 4th edn. Academic Press (1997)
2. Sederberg, T.W., Parry S.R.: Free-form deformation of solid geometric models. SIGGRAPH'86 **20** (1986) 151-160
3. Coquillart, S.: Extended free-form deformation: A sculpturing tool for 3D geometric modelling. SIGGRAPH'90 **24**(4) (1990) 187-193
4. Bloor, M.I.G., Wilson, M.J.: Using partial differential equations to generate free-form surfaces. Computer-Aided Design **22**(4) (1990) 202-212
5. Bloor, M.I.G., Wilson, M.J.: Representing PDE surfaces in terms of B-splines. Computer-Aided Design **22**(6) (1990) 324-331
6. Zhang, J.J., You, L.H.: PDE based surface representation—Vase design. Computers and Graphics **26** (2002) 89-98
7. You, L.H., Zhang, J.J., Comninos, P.: Blending surface generation using a fast and accurate analytical solution of a fourth order PDE with three shape control parameters. The Visual Computer **20** (2004) 199-214
8. Zhang, J.J., You, L.H.: Fast surface modelling using a 6th order PDE. Computer Graphics Forum **23**(3) (2004) 311-320
9. Monterde, J., Ugail, H.: On harmonic and biharmonic Bzier surfaces. Computer Aided Geometric Design **21**(7) (2004) 697-715

Interactive Visualization for OLAP

Kesaraporn Techapichetvanich and Amitava Datta

School of Computer Science & Software Engineering,
University of Western Australia,
Perth, WA 6009, Australia
{kes, datta}@csse.uwa.edu.au

Abstract. Business data collection is growing exponentially in recent years. A variety of industries and businesses have adopted new technologies of data storages such as data warehouses. On Line Analytical Processing (OLAP) has become an important tool for executives, managers, and analysts to explore, analyze, and extract interesting patterns from enormous amount of data stored in data warehouses and multidimensional databases. However, it is difficult for human analysts to interpret and extract meaningful information from large amount of data if the data is presented in textual form as relational tables. Visualization and interactive tools employ graphical display formats that help analysts to understand and extract useful information fast from huge data sets. This paper presents a new visual interactive exploration technique for an analysis of multidimensional databases. Users can gain both overviews and refine views on any particular region of interest of data cubes through the combination of interactive tools and navigational functions such as *drilling down*, *rolling up*, and *slicing*. Our technique allows users who are not experts in OLAP technology to explore and analyze OLAP data cubes and data warehouses without generating sophisticated queries. Furthermore, the visualization in our technique displays the exploration path enhancing the user's understanding of the exploration.

Keywords: Visualization, Visual Exploration, OLAP, data cube, data warehouse.

1 Introduction

Modern business processes generate an enormous amount of data that needs to be analyzed and understood for better business performance. Executives, managers, and analysts need a tool for making decisions and planning strategies. On-line analytical processing (OLAP) has become an important tool for interactive analysis of multidimensional databases such as data warehouses. This tool helps analysts to explore, analyze, and extract interesting patterns from massive amount of data stored in multidimensional databases. A variety of industries have adopted data warehouses as the preferred mode of data storage in order to manage the explosive growth of their databases [1].

Data in warehouses are historical data which is summarized from a variety of relational databases. In both OLAP tools and data warehouses, a model of data is formed as a multidimensional data cube. The data cube allows viewing aggregated data from different perspectives and is commonly organized into *dimensions* and *measures*. Both

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 206–214, 2005.

dimensions and measures are similar to independent and dependent variables in statistics. For instance, a product is the dimension and the number of unit sales of the product is the measure. The dimensions usually have hierarchies consisting of multiple levels of abstraction from a high level to a low level. For example, the same *product* dimension is composed of *product family*, *product department*, and *product name*. A *time* dimension comprises of year, quarter, and month.

OLAP tools provide functionalities such as *slicing*, *rolling up*, and *drilling down* for an end user to analyze and navigate through dynamic multidimensional data cubes. Though OLAP research has been conducted extensively in the past several years, there is very little available research in interactive visualization for OLAP.

Visualization is a powerful tool supporting visual representation and exploration of massive data sets. The capability of humans to interpret and capture information from graphical formats such as a chart is better than from a list of numbers or from viewing text formats. This paper introduces a novel interactive visual exploration technique for analysis of multidimensional data cubes from data warehouses. To obtain an effective and powerful analysis, the tool incorporates visualization into OLAP service which enables analysts to explore overviews of high levels of data and drill down into levels of detail of each dimension directly. The incorporation of both visualization and OLAP not only helps users to extract interesting patterns but also help them to interpret and analyze the extracted information from OLAP faster. Our technique allow users to view the visualization of all previous selected paths of interest so that users do not need to recognize which levels and dimensions they are looking at.

The rest of the paper is organized as follows. We discuss some related work on OLAP visualization in Section 2. We present our visualization framework in Section 3. We discuss the visual exploration process in our system through the generation of MDX queries in Section 4. We present a case study using our system in Section 5. Finally, we make some concluding remarks in Section 6.

2 Related Work

An extensive amount of research has been done on OLAP including textual presentation of queried results [2, 3, 4] such as a pivot table. We only discuss OLAP services which employ information visualization. ADVIZOR [5] provides visualization tools for exploring database through visual query and analysis. Three techniques, *Single Measure*, *Multiple Measure*, and *Anchored Measure*, are parts of this tool. The Single Measure approach represents a measure by using a 3D bar chart on a centered window called 3D Multiscape. The height of each bar shows a measure value. The Multiple Measure approach applies Scatterplot to visualize two measures along x-y axis rather than 3D Multiscape. Colors can display the third measure. The Anchored Measure approach combines ParaBox, Bubble plots, Parallel coordinates, and Box plots to visualize multidimensional data. Bubble plot axes visualize dimensions and Box plots are used for measures. Both Bubble plot axes and Box plot axes are arranged in the parallel coordinate style [6]. The system allows drilling down into low levels of abstraction only one dimension at a time.

Polaris is an interactive visual exploration tool which employs a table based visualization technique [7]. A graphical display of the table looks similar to a pivot table in

a textual format. The table of Polaris consists of three axes in which the x-y axes represent the numbers of rows and columns while the z-axis is a layout pane for viewing different visual metaphors such as glyphs and bar charts of the same table structure and data sets. The extension of Polaris [8] provides a tool for interactively exploring hierarchical structures in data sets. The system allows users to get overviews of data and drill down into low levels like a pivot table approach. In contrast, our technique allows users to explore independent overviews of data, low levels of detail, and any particular region of interest anytime during navigation.

Maniatis et al. [9] have developed their model for OLAP screens, called Cube Presentation Model (CPM), and applied a visualization technique, Table Lens, into their model. The CPM model consists of two layers which are logical and presentation layers. The logical layer deals with data retrieval while the presentation layer is for data presentation. The model employs cross-joins for retrieving maximum, minimum, and closest average values. The main goal of the system is to determine the window of interest for viewing the data detail of particular areas in large overviews of the cross-join window. The system does not provide drilling down and rolling up features for exploring multiple levels along different hierarchies.

3 Our Visualization Framework for OLAP

Since hierarchical structures have been deployed in most multidimensional databases, we feel that it is difficult for users to explore multidimensional data with a tool providing only overviews of data. It is important for users to interactively drill down through the low levels of detail to refine their views. Furthermore, only interactive textual displays such as PivotTable is not effective for understanding or extracting patterns from multidimensional databases.

Our technique allows users to make a decision whether they want to get overviews of data, to drill down into low levels, to roll up to high levels, or to view any particular region of interest of data anytime. Our system provides a navigation facility which reduces user responsibility of remembering the exploration path of interest.

In the remainder of this section, we give a short review of our method of hierarchical visualization, the design of our tool for visual feedbacks, and the interactive tools for exploring OLAP data cubes.

3.1 Hierarchical Exploration

We briefly introduce the Hierarchical Dynamic Dimensional Visualization (HDDV) [10] technique. In this technique, a horizontal bar, called a *barstick*, represents a dimension in a data set. Each barstick has the same length to optimize the screen space and consists of a sequence of vertical lines representing one or more records. The number of records in each vertical line are equal to the total number of records in the database divided by the length of the barstick. The color of each vertical line shows the density of the number of records satisfying the query range. A darker color represents higher number of satisfying records. The analysis in HDDV is driven by query range. The user can explore a set of attributes or dimensions in a hierarchical fashion by specifying query ranges. Suppose *attribute1* and *attribute2* are two attributes of interest. When the

user specifies a range for attribute1, the vertical lines containing records in this range are rendered in color. Then, the user chooses a range for attribute2. The vertical lines containing the records satisfying both the ranges for attribute1 and attribute2 are rendered.

3.2 Visualizing OLAP Data Cubes

We have modified the idea from HDDV [10] for exploring hierarchical structure of OLAP data cubes. We use a barstick for representing each dimension of the data cube but the details of the display are different. Each barstick is vertically arranged in hierarchical fashion when users select the highest level of dimension of interest. A barstick is divided into small rectangles which represent all existing members of the level in that dimension. A barstick does not show the member with no measure value. In other words, the member which does not have a measure value is hidden from the barstick. The width of each rectangle is calculated based on the proportion of measure value of the member in the selected dimension. For example, if the selected dimension is 'time', it contains four members in the next level called 'quarter'. The total profit is $250,000 and profits for the quarters are $70,000, $55,000, $37,500 and $87,500. The calculated proportions of all profits in each quarter are 28%, 22%, 15%, and 35%, respectively and they are the proportion of the width of all rectangles representing the quarters.

Figure 1 presents an example of visualization for exploring OLAP data cubes. A tree structure in the left panel of Figure 1 represents the hierarchical arrangement of the data cube. The upper right panel displays explored hierarchical barsticks while the lower right panel illustrates one deeper level of the currently selected member. In this example, the variable 'Unit sales' is selected as a measure of interest. The first barstick represents a 'Product Family' level of the 'Product' dimension. There are three members including 'Drinks', 'Foods', and 'Non-consumable' in 'Product Family' and 'Foods' has the highest unit sales displayed by its widest proportion of the rectangle in the barstick. The second barstick illustrates a 'Product Department' level of the 'Product' dimension. In 'Product Family', the members of 'Drink' consists of 'Alcoholic Beverages', 'Beverages', and 'Dairy'. 'Store Type' level of the 'Store Type' dimension is selected in the third barstick and 'Supermarket' has the highest proportion of unit sales of the drink product followed by 'Deluxe Supermarket', 'Gourmet Supermarket' (G), 'Mid-Size Grocery' (M) and 'Small Grocery' (S). The fourth and last barsticks are explored in the 'Year' and 'Quarter' levels of the 'Time' dimension. The data for unit sales of 'Drink' product exist only for 1997 and the fourth quarter has the largest amounts of unit sales. The lower right panel displays the month level consisting of July, August, and September or 7, 8, and 9 when the user places the mouse over the third quarter of the last barstick.

3.3 Interaction

To efficiently support the analysis and exploration processes of the hierarchical structure, our technique provides several navigational functions:

Drill down: Drill down is a function to navigate into deeper hierarchical levels in each dimension. This feature allows users to view a dimension in more details. Our technique

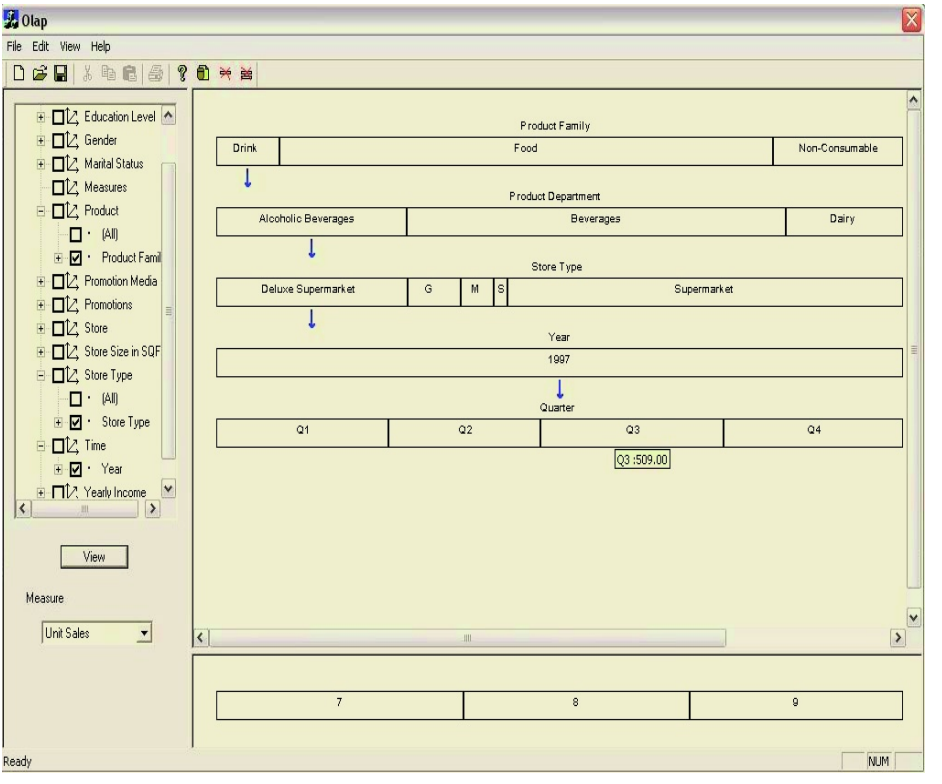


Fig. 1. The left panel shows the tree structure of the OLAP data cube. The upper right panel visualizes the selected dimensions as hierarchical bar sticks including Product Family, Product Department, Store Type, Year, and Quarter levels. The lower right panel displays one deeper level of the data cube in advance depending on the position of the mouse. In this case, the user has positioned the mouse on 'Q3' and this panel shows the 'Month' level of the fourth quarter

provides this function through left mouse double click. Users can drill down into deeper levels any time in any dimension. Figure 1 shows drilling down in the 'Product' dimension from 'Product Family' to 'Product Department' and in the 'Time' dimension from the 'Year' level to the 'Quarter' level of 'Unit sales' sold in the 'Deluxe Supermarket' store type.

Roll up: Roll up is a function to navigate for exploring upper levels of dimensions. Users can roll up any particular dimension by double clicking the right mouse button. Figure 2 illustrates rolling up of the 'Product' dimension from the 'Product Department' level to the 'Product Family' level.

Slice: This function allows users to view a particular sub-cube for any selected dimension of the data cube. Our tool provides this function through left mouse click for viewing a measure value of other members in the same level. Figure 3 shows the change of exploration from Deluxe Supermarket (in Figure 1) to Supermarket.

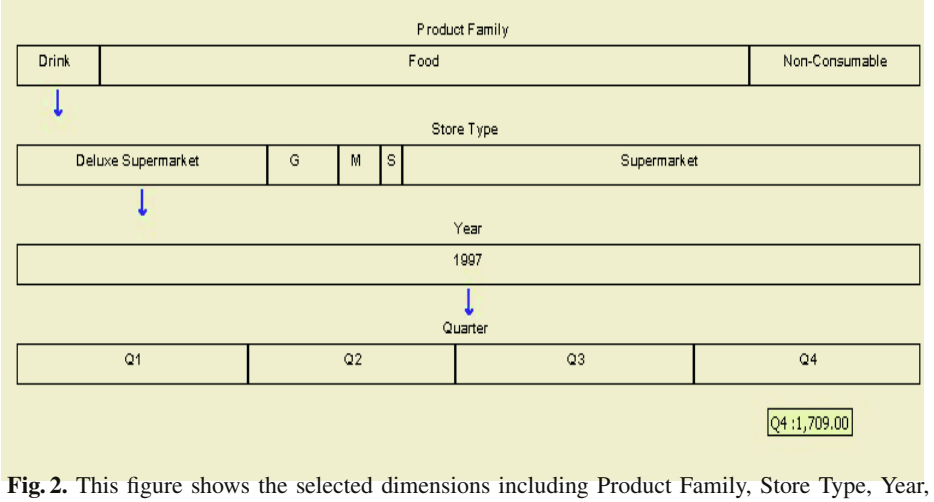


Fig. 2. This figure shows the selected dimensions including Product Family, Store Type, Year, and Quarter. The Product dimension is rolled up from the Product Department level to Product Family level of Figure 1

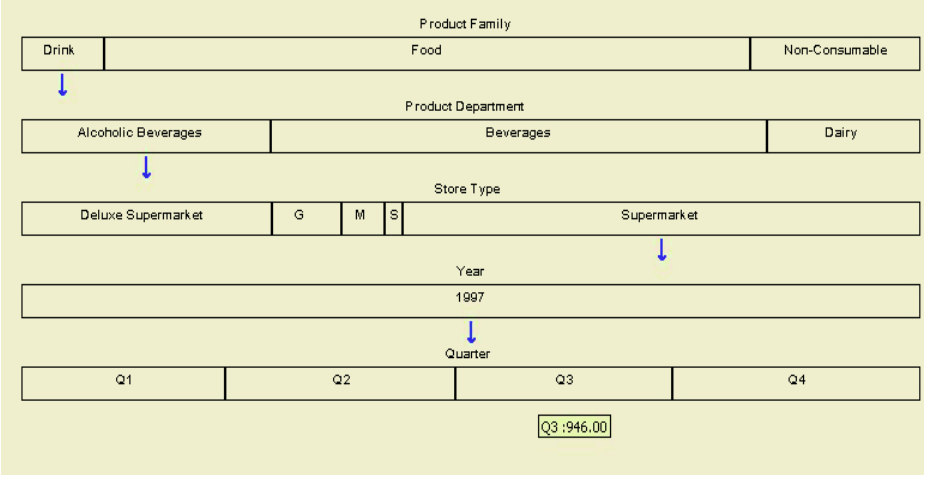


Fig. 3. This figure shows the change of member selection from Supermarket to Deluxe Supermarket in the Store Type level

All navigational functions can be automatically combined when users interact with each barstick so that they can view any particular region of interest in the data cube. For instance, the combination of drill down and slice functions allow users to explore unit sales of Alcoholic Beverages in Product Department for all quarters in 1997. Moreover, users can view each independent dimension when each barstick is first created or by clicking the right mouse button on any barstick. The system supports on demand details. When users move mouse over any rectangle in the barstick, the details of the rectangle

including the name of the specified member and its measure value are shown in a pop-up box as shown in Figure 1.

4 Visual Exploration and MDX Query

Similar to a SQL syntax for querying and manipulating data from relational databases, Multidimensional Expressions (MDX) is a powerful syntax for querying and manipulating multidimensional data from OLAP data cubes [11]. We have implemented the binding of MDX queries with the navigational functions of the interactive tool to enable users who are not OLAP experts to explore OLAP data cubes and data warehouses without generating sophisticated MDX queries.

We describe some examples of the combination of MDX and the interactive tools based on Figure 1. Suppose 'Sales' is a selected data cube, 'Unit Sales' is a selected measure, and 'Product Family' is the selected level of the 'Product dimension' for querying. To view all members of the 'Product Family' level in proportion, a MDX query implying this process can be described as:

```
WITH MEMBER Measures.[sum] AS
 'sum([Product].[ProductFamily].members, Measures.[UnitSales])'
MEMBER Measures.[percent] AS '([Product].CURRENTMEMBER,
 Measures.[UnitSales])/(Measures.[sum]))',
FORMAT_STRING = 'Percent'
SELECT {[Measures].[percent], [Measures].[UnitSales]} on columns,
NON EMPTY [Product].[ProductFamily].members on rows
FROM Sales
```

When the user drills down on the 'Product Department' level through 'Drink' in the 'Product Family' level, an equivalent MDX query as shown below is automatically generated to display the 'Product Department' members on the second bar stick.

```
WITH MEMBER Measures.[sum] AS
 'sum([Product].[AllProducts].[Drink].children, Measures.[UnitSales])'
MEMBER Measures.[percent] AS '([Product].CURRENTMEMBER,
 Measures.[UnitSales])/(Measures.[sum]))',
FORMAT_STRING = 'Percent'
SELECT {[Measures].[percent], [Measures].[UnitSales]} on columns,
NONEMPTY{DESCENDANT[All Products].[Drink],[ProductDepartment])}
 on rows
FROM Sales
```

As shown in Figure 1 and Figure 3, suppose that the user changes the queried members in the bar stick representing the 'Store Type' level from 'Deluxe Supermarket' to 'Supermarket'. The equivalent MDX statement for this process to display the 'Year' members on the fourth barstick of the 'Time' dimension is:

```
WITH MEMBER Measures.[sum] AS
 'sum([Time].[Year].members, Measures.[UnitSales])'
MEMBER Measures.[percent] AS '(([Time].CURRENTMEMBER,
 Measures.[UnitSales])/(Measures.[sum]))',
FORMAT_STRING = 'Percent'
SELECT {[Measures].[percent], [Measures].[UnitSales]} on columns,
NON EMPTY [Time].[Year].members on rows
FROM SalesWHERE ([Product].[All Products].[Drink].[Alcoholic Beverages],
 [Store Type].[All Store Type].[Supermarket])
```

We generate these MDX queries automatically depending on user interaction.

5 Case Study

We use a FoodMart 2000 database [12] in the case study. The database consists of data cubes such as 'Budget', 'HR', 'Sales', and 'Warehouse'. The 'Sales' data cube comprises twelve dimensions excluding the hierarchical levels of each dimension and seven measures. The 'Product', 'Time', 'Store Type', 'Promotion Media', and 'Promotions' dimensions and 'Unit sales' and 'Profit' measures are used for the case study. Suppose the store manager would like to increase the sales of the drink product stocked in the store. Figure 1, 2, 3 shows exploring of the drink product family. The manager can extend the exploration of the 'Promotions' and 'Promotion Media' dimensions to obtain how they affect the sale amounts in each year. For example, it is easy to find the following correlations from exploration of the data cube.

Daily newspaper, radio, and TV tend to be the most effective media to increase the amount of sales in sales days promotion of the supermarkets while bulk mails tend to be the most effective way to advertise the promotions including 'You Save Day', 'Shelf Emptiers', and 'Sales Galore' for gourmet supermarkets. Daily newspaper is the most effective media to advertise the promotions such as 'Big Time Discounts' for mid-size groceries, and 'In-Store Coupon' is the most effective way to increase the sale promotions for small groceries. In addition, the amount of unit sales vary over time. For instance, the fourth quarter has the highest amount of alcoholic beverage and beverage sales in all stores except the small groceries which have the highest amount of alcoholic beverage sales in the third quarter and the gourmet supermarkets which have the highest amount of beverage sales in the second quarter. However, as the analysts, managers, and executives know better market and store situations, they can explore and analyze the data in more efficient ways.

6 Conclusion

We have introduced a novel interactive visual exploration tool for analysis of OLAP data cubes. Our technique provides visual feedbacks while users explore data cubes in graphical formats rather than textual table formats. The incorporation of both visualization and the OLAP service enhances users to deeply understand, gain insight and extract useful information faster from their data. We present hierarchical barsticks for

exploring and visualizing hierarchical structures of the OLAP data cubes. Users can view the trail of the exploration through visualization in order to reduce the recognition load and also view one deeper level in advance before drilling down into deeper levels. In addition, our technique provides users overviews and refined views of interest in the data cubes. Users are allowed to change the exploration views anytime through the combination of navigational functions and interactive tools. We expect our tool to be useful for interactive visual exploration of data cubes.

References

1. Chaudhuri, S., Dayal, U.: An overview of data warehousing and olap technology. SIGMOD Rec. **26** (1997) 65–74
2. O'Donnell, P., Draper, N.: An experimental evaluation of an alternative to the pivot table for ad hoc access to olap dat. In: Proceedings of the 2004 IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World. (2004)
3. : Microsoft excel:user's guide 2, version 4.0. Redmond, WA Microsoft Corperation (1992)
4. Thomsen, E.: OLAP Solutions Building Multidimensional Information Systems. Wiley Computer Publishing (1997)
5. Eick, S.G.: Visualizing multi-dimensional data. SIGGRAPH Computer Graphics **34** (2000) 61–67
6. Inselberg, A., Dimsdale, B.: Parallel coordinates for visualizing multidimensional geometry. Computer Graphics (Proceedings of CG International) (1987) 25–44
7. Chris Stolte, D.T., Hanrahan, P.: Polaris: a system for query, analysis, and visualization of multidimensional relational databases. IEEE Transactions on Visualization and Computer Graphics **8** (2002) 52–65
8. Chris Stolte, D.T., Hanrahan, P.: Query, analysis, and visualization of hierarchically structured data using polaris. In: Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM Press (2002) 112–122
9. Maniatis, A.S., Vassiliadis, P., Skiadopoulos, S., Vassiliou, Y.: Advance visualization for olap. In: Proceedings of the 6th ACM international workshop on data wareshousing and OLAP, ACM Press (2003) 9–16
10. Techapichetvanich, K., Datta, A., Owens, R.: Hddv: Hierarchical dynamic dimensional visualization. In: Proc. IASTED International Conference on Databases and Applications. (2004) 157–162
11. Nolan, C.: Manipulate and query olap data using adomd and multidimensional expression. In: Microsoft Systems Journal. Microsoft (1999)
12. (Microsoft) `http://msdn.microsoft.com`.

Interactive 3D Editing on Tiled Display Wall

Xiuhui Wang, Wei Hua, and Hujun Bao*

State Key Lab of CAD&CG, Zhejiang University,
Hangzhou, Zhejiang, P.R.C 310058
{WangXiuhui, HuaWei, Bao}@cad.zju.edu.cn

Abstract. Recent interest in large displays has led to the rapid development of cluster-based tiled display wall systems, which are comprised of many individual projectors arranged in an array, and driven by a cluster of PCs. In this paper we describe an interaction protocol and the essential synchronization mechanisms to support interactive 3D editing on cluster-based tiled display wall. In order to examine the feasibility of our approach, we have developed a parallel interaction library, named *PIEL* (Parallel Interactive Editing Library). By using the PIEL library, we can conveniently port rendering applications running on stand-along PC to cluster-based tiled display wall systems.

1 Introduction

While the performance of commodity interactive graphics hardware has been greatly improved over the past few decades, not only have the speed improvements kept up with Moore's Law, but also each successive generation of graphics hardware has expanded the feature set. However, because of the enormous economies of scale afforded by commercialization, it is still quite unpractical to make effective use of such systems as general-purpose parallel scientific visualization of large data sets, Distributed Virtual Reality (DVR) and large tiled display wall systems.

The necessary components for large tiled display wall systems on clusters of PCs have matured sufficiently to allow exploration of clusters as a reasonable alternative to multiprocessor servers for high-end visualization. In addition to graphics accelerators, processor power, memory and I/O controllers on commodity PCs have reached a level of sophistication that permits high-speed memory, network, disk, and graphics I/O to all occur simultaneously, the high-speed general purpose networks are now fast enough to real-timely handle the demands of routing commands and data between PCs [11]. By taking advantage of these facilities, we have designed an interaction protocol and several essential synchronization mechanisms for interactive 3D editing on tiled display wall systems. As a typical example, we have developed a parallel scene-editing library, named *PIEL* (Parallel Interactive Editing Library), which is an implementation of our parallel interaction protocol and encapsulates the synchronization mechanisms.

* This work received support from the National 973 Project (No. 2002CB312102) and for Doctor Key Foundation (No. 20030335083), and National Natural Science Foundation (Grant No. 60021201,60203014).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 215–224, 2005.

2 Related Work

2.1 Parallel Interaction Protocol

Crockett introduced the Parallel Graphics Library (PGL), which supports parallel polygon rendering on distributed-memory message-passing (DMMP) architectures [1, 4]. Due to the characteristics of its target architecture, the PGL is intended specifically for distributed memory message passing environments. PixelFlow [5] is another system designed to support multiple simultaneous inputs from a parallel host machine. However, due to the underlying image composition architecture, the PixelFlow also imposes frame semantics. Because of these constraints, the PGL and the PixelFlow do not meet the requirements of many graphics applications. Igehy [3] described a parallel graphics interface, which can be used in conjunction with the existing API as a new paradigm for high-performance graphics applications. Their work has mainly focused on overcoming system bus bottlenecks. Moreover the same technique allows multiple remote hosts to submit streams in parallel to a single rendering server over a high-speed network. In the parallel programs, each stream has a distinct associated OpenGL context, and explicit synchronization primitives are inserted into the command streams if the programmer wishes to express ordering constraints between contexts [2]. Therefore, when the rendering server sees a synchronization primitive, it must quickly switch OpenGL contexts to begin executing commands from another stream until the synchronization primitive can be resolved.

2.2 PC-Cluster Based Large Display Wall

The WireGL [7, 9] is a scalable graphics library to realize PC-cluster based tiled display wall. However, as it is a set of low-level application programming interfaces, the WireGL does not include any architecture to handle interactions. Moreover, it must use some special designed networking hardware to transfer a great amount of data for composing tiled images. Another approach related to our synchronization mechanisms for interactive 3D editing in parallel computing environment was discussed in [8], which repartitions models stored in a scene graph across multiple nodes in a PC cluster. The objective of this approach is to minimize geometry broadcast during rendering. In their experiments, a large static 3D model is preprocessed to create a hierarchical multi-resolution model. Portions of the model are replicated across some, but not all, nodes to reduce potential communication bottlenecks associated with moving graphics data during interactive rendering. However, it does not handle relevant interaction information for 3D scene editing.

3 Objective

Both the interactive 3D editing technique and the cluster-based large display wall have been developed for years separately. A number of algorithms and architectures have been presented for each technique respectively. However, few studies have been conducted on how to integrate them. We are dedicated to designing

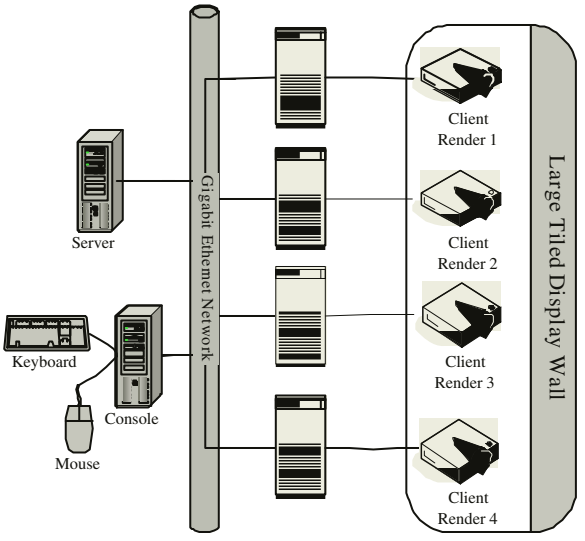


Fig. 1. The architecture of our support environment. Each client render is connected to a projector, and the projectors constitute a large tiled display wall

a support environment for interactive 3D editing on cluster-based tiled display wall, including an interaction protocol and relevant synchronization mechanisms.

The architecture of our support environment is illustrated in figure 1. The operation commands are sent from the console, and the client renders draw 3D scene according to the received commands. The Server is responsible for maintaining frame synchronization and global clock synchronization.

4 Parallel Interaction Protocol

Ongoing conversations between agents often fall into typical patterns. In such cases, certain message sequences are expected, and, at any point in the conversation, other messages are expected to follow. These typical patterns of message exchange are called parallel interaction protocols [14]. Considering the scenario of interactive 3D editing on PC-cluster with a tiled display wall, we focused on the following four factors.

- Network delay. In Distributed Virtual Reality Environment (DVRE), the immersive sensation heavily depends on the real time of interactive manipulation. Therefore, given a set of networking hardware, relieving the influence from network delay is an important factor in designing our interaction protocols.
- Communication bandwidth. Bandwidth is a key factor that affects the interaction efficiency on PC-cluster. A great deal of rendering primitives is often transferred from some computation nodes to others. Thus, poorly designed interaction protocols can greatly debase the interaction performance.

- Reliability. The reliability problem of interaction protocols roots in a variety of synchronization mechanisms. We must deal with a great variety of synchronization problems such as frame synchronization, viewpoint synchronization, object status synchronization and global clock synchronization, etc. Improper mechanisms for dealing with the synchronizations may result in some severe problems as deadlock or starvation.
- Scalability. The choices of parallel mechanisms will determine the scalability. Whether peer-to-peer or client/server parallel is dependent on applications.

In our scenario as the interactive editing on PC-cluster with tiled display wall, many types of messages must be handled. Our interaction protocol defines five types of messages. All messages are of the same format of data package, as shown in Figure 2. Because all of them are predefined, the types of their senders and the length of their parameters are determined implicitly. A data package consists of three parts: Priority Level, Message Type and Message Parameters. The Priority Level is used for priority control, the detail of which will be discussed in the next section. The Message Type occupies a single byte, and the Message Parameters' length is variable according to the Message type. The syntax and semantics of each message types are given as follows.

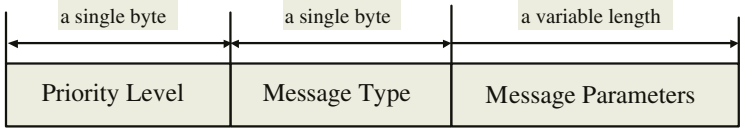


Fig. 2. The format of data package. Each message type corresponds to a set of message parameters. For a given kind of message, the length of entire data package is fixed

All messages defined in our interaction protocol are described in an augmented Backus-Naur Form (BNF) as follows [14].

```
OCTET       = <any 8-bit sequence of data>
CHAR        = <any US-ASCII character (octets 0 - 127)>
BOOLEAN     = <any US-ASCII digit "0".."1">>
DIGIT       = <any US-ASCII digit "0".."9">
CTL         = <any US-ASCII control character
           (OCTET 0 - 31) and DEL (127)>
TEXT        = <any OCTET except CTLs >
HEX         = "A" | "B" | "C" | "D" | "E" | "F"
         | "a" | "b" | "c" | "d" | "e" | "f" | DIGIT
```

4.1 Background-Render-Complete Message and Begin-Swap-Buffer Message

These two messages are used for the frame synchronization, which is crucial for various applications such as walkthrough and interactive 3D editing. The

background-render-complete message notifies the server that some client render has completed the back-buffer rendering. And the begin-swap-buffer message informs the entire client renders that the time to swap buffer has arrived.

```
message     = msg_type  * BOOLEAN
msg_type    = OCTET
```

4.2 Camera-Update Message

This message is for changing viewing parameters, wrapped as a camera object. The camera-update message informs the entire clients renders to update their local cameras. The camera updating includes camera translating, panning, rotating, etc.

```
message     = msg_type  *msg_para
msg_type    = OCTET
msg_para    = *TEXT | *HEX
```

4.3 Object-Status-Update Message

This message is used for modifying objects in a scene. The object-status-update message informs the entire clients renders to update a given object. The related operations mainly include object selecting, adding, deleting, transforming, etc.

```
message         = msg_type  *msg_para
msg_type        = OCTET
msg_para        = *(target_object  [operation_data])
target_object   = *TEXT
operation_data  = *HEX
```

4.4 Global-Clock-Signal Message

Since continuous interactive objects may move by themselves even without any operation issued on them, the states can continuously update not only according to the operations issued on it, but also because of the passage of time as well [12]. Thus, in addition to the object-status-update message, a global-clock signal is necessary for all client renders.

```
message     = msg_type *DIGIT
msg_type    = OCTET
```

5 Synchronization Mechanism

Maintaining the consistency of models in a scene graph cross nodes in a PC-cluster is one of the most significant challenges. Accounting for our scenario, we deal with four kinds of synchronization, such as frame synchronization, camera synchronization, object status synchronization and global clock synchronization, as shown by figure 3.

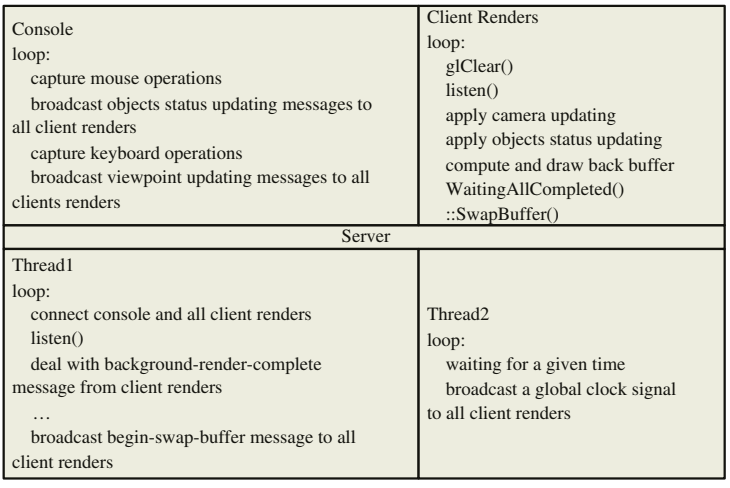


Fig. 3. A simple interactive process. The client renders compute and write back frame buffer concurrently, and the console collects operation commands and dispatch them into the client renders. The server coordinates per-frame operations and offers global clock signals

5.1 Frame Synchronization and Camera Synchronization

Frame synchronization is to keep every client render in the cluster displaying images in the same rate. It is crucial for interactive 3D editing visually on the tiled display wall. In our synchronization architecture, once a client render completes the rendering in its back buffer for a frame, it sends a background-render-complete message to the server. After the server has received all of background-render-complete message sending by each client render for a given frame, it broadcasts a begin-swap-buffer message to notify each client render to swap frame buffer.

The camera synchronization is to keep all viewing frustums of client renders consistent when viewing parameters are adjusted. It is realized by broadcasting a camera-update message from the console to all client renders whenever some interactive operation messages arrive.

5.2 Object Status Synchronization and Global Clock Synchronization

Object status synchronization is necessary for the scene edit. After users edit some objects in a scene, the status of the object is modified. Then, the console must obtain the objects' ID and related parameters, and send them to all client renders. On the other hand, if there are some dynamic objects driven by time signal, a global clock signal across all nodes in the cluster should be provided for this kind of objects instead of local clock signal of each node [15, 13]. In our implementation, a high frequency clock signal is broadcast from the server to all client renders. Nevertheless, broadcasting high frequency clock signal is a

bandwidth-consuming operation. As an alternative, we let the server send time adjustment signals to all client renders in low frequency. This kind of signal is used to adjust the local clock of each client render.

6 Implementation

In order to test the feasibility and reliability of our interaction protocol and synchronization mechanism presented in this paper, we have developed a library named PIEL. The PIEL, which implements the interaction protocol and encapsulates the synchronization mechanism, is capable of handling multiple client renders for interactive 3D scene editing on tiled display wall.

6.1 PIEL

The PIEL is intended both as a test bed for research on parallel interaction protocols, as well as a semi-production interactive 3D editing library for parallel scientific and engineering applications. Figure 4 shows the flow of messages through the PIEL pipeline. The pipeline contains several threads, shown as gray boxes, which communicate through some connection-oriented sockets. In this example, the server act as a bridge between the console and the client renders. The operation commands from the console are firstly sent to the server, and then dispatched to all client renders. The server and the client renders communicate with each other to achieve frame synchronization. Meanwhile, the server broadcasts the global-timer-signal for timer synchronization between the console and all client renders.

One key design issue that comes up in developing the PIEL is how to keep the consistent of the status of 3D scene objects, which are distributed in several client renders working on different PCs. When the console sends commands that affect some objects' status in the scene, the console must update all client renders and itself at the same time. But whether or not we should update the console immediately is a contradiction. If we update the console immediately, then the console's output is uninterrupted while the console may be inconsistent with the client renders. On the other hand, if we update the console till the entire client renders re-

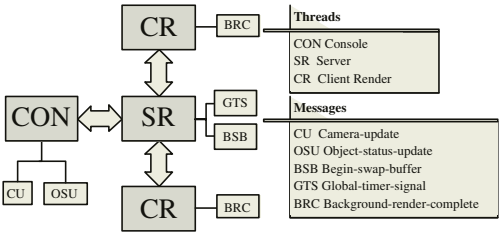


Fig. 4. The Pipeline of the PIEL. Two client renders work synchronistically with a console through a server

Fig. 5. A typical view from the Model of the West Lake, Hang Zhou

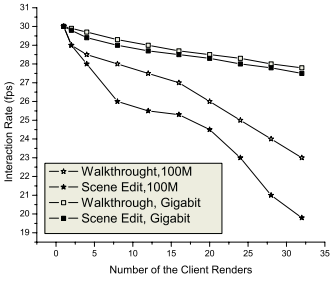


Fig. 6. Performance Graphs. In order to compare with the stand-alone applications, all the cures for performance start from a single client render. To show the impact of the network bandwidth on the interaction rate, we have tested the performance under the conditions of 100M network and gigabit network respectively

sponses arrive, then the user may feel as if the keyboard failed to function. As a tradeoff, we implement it by maintaining a command queue and lazily updating. If a client receives new updating commands while waiting for frame synchronization, it just pushes them into commands queue. After the frame synchronization has completed, the pushed commands are executed immediately.

Another key implementation design issue that comes up in implementing the PIEL is the global clock synchronization. There may be some dynamic objects driven by one clock, so a global clock signal in the PC-cluster is necessary. However, sending high frequency global clock signal from server to the entire client renders is bandwidth-consuming and unreliable. We deal with this problem by discontinuously sending correcting signals, instead of sending high frequency clock signals. The correcting signals are used to adjust the local clock where the client renders run.

6.2 Performance

Because poor performance often hides architectural bottlenecks, the performance is in the first consideration of designing the PIEL. To test its performance, we applied the PIEL to two stand-alone applications: a VRML render engine and a terrain render engine.

By using the PIEL, the VRML render engine[10] can concurrently render VRML scenes on multi-PCs to display images in a tiled display wall with the aid of multi-projectors. Moreover, we can edit the scene interactively through a virtual mouse on the tiled display wall. The terrain render engine enable our to

work though on the tiled display wall by a virtual mouse or keyboard. Figure 5 shows a view of the tiled display wall.

Figure 6 shows the performance of the PIEL. According to the figure, the walkthrough application is always faster than the scene editing application under the same conditions. The reason is that the network traffic during the walkthrough is less. We can also find that the loss of the interaction rate, especially for the editing, can get considerable decrease in the case of gigabit networks. At the same time, the figure shows that the decrease of the interaction rate is acceptable, even under conditions of the 100M networks, when the number of client renders is bellow sixteen.

7 Conclusion

We present an interaction protocol for interactive 3D editing on tiled display wall, and address the related synchronization issues. The PIEL is one typical implementation of these ideas. However, the architecture embodied by the PIEL is not the only possible choice, and it is instructive to examine the design considerations of alternative implementations.

To gain the maximum performance, client renders should have roughly the same amount of work to do. On the other hand, as most sort-first [6] parallel rendering systems, applications using the PIEL go without dynamic load balancing schemes. If primitives are uniformly distributed among client renders, the rendering process is almost perfectly parallel, although overheads are introduced due to primitive splitting. Unfortunately, most 3D scenes are not so well-balanced, thus integrating dynamic load balancing with this work would be an interesting research direction.

Acknowledgements

We thank Mr. Zhong Ren and Mr. Huaisheng Zhang for providing the VRML Render Engine and Terrain Render Engine.

References

1. T. Crockett. Design Considerations for Parallel Graphics Libraries. Proceedings of the Intel Supercomputer Users Group (1994)
2. G. Chiola and A. Ferscha. Performance comparison of distributed petri net simulations. In SCS, editor, Summer Simulation Conference, Ottawa, Canada, July 24-26 1995
3. H. Igehy, G. Stoll, and P. Hanrahan. The design of a parallel graphics interface. In SIGGRAPH Computer Graphics (1998)
4. T. W. Crockett. Parallel rendering. In A. Kent and J. G. Williams, editors, Encyclopedia of Computer Science and Technology, volume 34, Supp. 19, A., pages 335–371 (1995)

5. Molnar S., J. Eyles, and J. Poulton. PixelFlow: HighSpeed Rendering Using Image Composition. SIGGRAPH 92, pp. 231-240
6. Steven Molnar, Michael Cox, David Ellsworth, and Henry Fuchs. A sorting classification of parallel rendering. IEEE Computer Graphics and Applications, 14(4) pages 23–32, July 1994
7. Greg Humphreys, Matthew Eldridge, Ian Buck, Gordon Stoll, Matthew Everett, and Pat Hanrahan. WireGL: A scalable graphics system for clusters. Proceedings of SIGGRAPH 2001, pages 129-140, August 2001
8. Rudrajit Samanta, Thomas Funkhouser and Kai Li. Parallel rendering with K-Way replication. IEEE 2001 Symposium on Parallel Graphics, October 2001
9. G. Humphreys, M. Houston, R. Ng, R. Frank, S. Ahern, P. Kirchner, and J. T. Klosowski. Chromium: A Stream Processing Framework for Interactive Rendering on Clusters. In SIGGRAPH 2002
10. Xiuhui Wang, Wei Hua, Hujun Bao. Cluster Based Parallel Rendering And Interactive Technology. In Chinagraph 2004
11. Qin Xiao. Delayed Consistency Model for Distributed Interactive Systems with Real-Time Continous Medis. Journal of Software, 2002,13(6): 1029-1039
12. Mills, D.L. Internet time Synchronization: the network time protocol .IEEE Transactions on Communications, 1991,39(10): 1482-1493
13. E.Wes Bethel, Greg Humphreys, Brian Paul, and J.Dean Brederson. Sort-First, Distributed Memory Parallel Visualization and Rendering. Proceedings of the 2003 IEEE Symposium on Parallel and Large Data Visualization and Graphics (PVG 2003)
14. FOUNDATION FOR INTELLIGENT PHYSICAL AGENTS (FIPA). FIPA Interactive Protocol Library Specification. http://www.fipa.org.
15. Ian Buck, Greg Humphreys, and Pat Hanrahan. Tracking graphics state for networked rendering. Proceedings of SIGGRAPH/Eurographics Workshop on Graphics Hardware, August 2000

A Toolkit for Automatically Modeling and Simulating 3D Multi-articulation Entity in Distributed Virtual Environment*

Liang Xiaohui, Wang Chuanpeng, Che Yinghui, Yu Jiangying, and Qu Na

Key Lab. of VR Technology of BeiHang University, Ministry of Education,
No. 6863 BeiHang University, No.37 Xueyuan Road, Beijing, 100083, P.R.China
lxh@vrlab.buaa.edu.cn

Abstract. This paper describes the Entity Modeling and Simulating Platform (EMS), a toolkit which can automatically model and simulate multi-articulation entities (actor or avatar) in a distributed virtual environment. EMS focuses particularly on modeling the physical and behavioral attributes of an entity and defining the relationship among different attributes. Using EMS, a user can interactively define the attributes of an entity, run and control the entity in a distributed virtual environment without programming. In this paper, the main modules and key technologies of EMS are introduced.

1 Introduction

Distributed Virtual Environment (DVE) creates a shared and consistent environment for the users located in different areas. In the DVE, every real object in the world can be abstracted as an entity (actor or avatar), such as a terrain, a building, a car or even a virtual human etc. The users can then immerse into the virtual environment, to control these entities and to collaborate with others.

As Distributed Interactive Simulation (DIS) is an important application field of the DVE, two important IEEE standards [1] [2] have been setup to regulate the developing process of DIS, mainly focused on network protocols [1] and the interface specification between application layer and communication layer [2]. In [1], the entity in a DVE is defined as follows: *Entity is an element of the synthetic environment that is created and controlled by a simulation application and effected by the exchange of DIS protocol data units.* Creating an entity includes two important processes: the modeling process and the simulating process. The modeling process is in charge of describing three kinds of important properties of an entity, the geometrical properties (the geometry shape of an entity, such as point, line, surface and geometry topology), the physical properties (such as

* This research work is supported by China High Technology Development Project (863 project), Number: 2004AA115130 and China National Key Basic Research and Development Program (973 Program), Number: 2002CB312105.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 225–234, 2005.

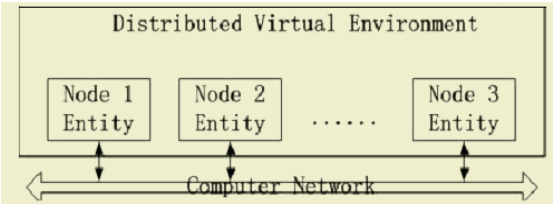


Fig. 1. Distributed Virtual Environment and the Entities

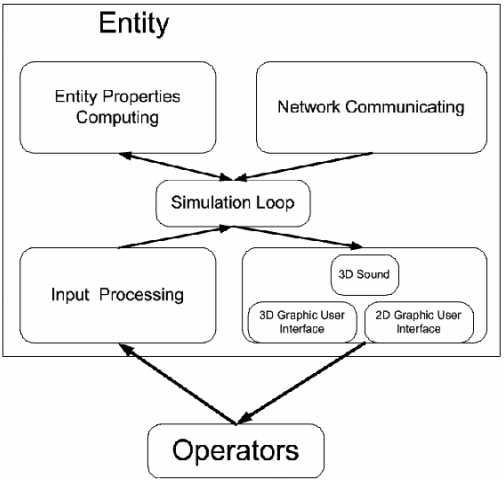


Fig. 2. Software Architecture of an entity

weight, material) and behavioral properties (such as the kinematics or dynamics of an entity), the simulating process then makes use of these modeling information obtained from the modeling process to make the entity run. In addition, to get a 3D entity in a DVE, both a 3D graphics rendering module and a network communication module also need to be added in the simulation process.

There are three main characteristics of an entity. The first is **Diversity**.There are many kinds of entities in a DVE. Different properties can be used to define a particular kind of entity. Those entities with the same properties can be regarded as one kind of entity. For example, there might be three kinds of entities such as aerial, oceanic and terrestrial; each must have different properties to differentiate itself from the other.The second is **Multi-Articulation**.Typically, an entity in a DVE has multi articulations. The number of articulations of each entity might vary in a wide range depending on how complex the target entity is. For example, there should be more articulations for a complex entity like a virtual human comprising a head, all kinds of organs, arms, hands an so on than a relatively simple entity such as a tank, which is made up with only 3 parts as a body, a turret and a gun barrel. Usually, the user need to control each articulation of an entity in a DVE.The third is **Multi-Granularity**. The entity also has multi

granularity. For example, the "car" entity has a different view from different granularity. In the highest granularity, it can be viewed as the instruments of the driver and the seat of the passenger, the user can operates these instruments; in a middle granularity, it can be only viewed as the shape of the car and the user can only make it turn left, turn right, go ahead or stop.

The more features added to an entity, the more complex the modeling and simulating process will be. To get the DVE developing process easier and faster, keeping researching the automatic entity developing method becomes very important.

The structure of this paper is as following: Section2 serves as an introduction to the background and some related work. Section3 focuses on the core modules of the EMS built in this work as well as the runtime results of it. The last Section 4 summarizes the work have been done and proposes the potential research work in the future.

2 Related Works

A lot of research work has been done on investigating a variety of methods to create the entities in a DVE. All the different methods being developed so far can be characterized into three main categories.

2.1 Program Based Method

Program Based Method might be the most widely used method to create entity. The users first use software such as Multigen creator, 3DMAX to make geometry model, and then programming. The program usually includes physical and behavioral modules, 3D rendering, network communication and interactively controlling. To reuse entity program, some general frameworks (such as dll, lib) [3][12][13][15] are developed.

To reuse entity program, some general frameworks (such as dll, lib) [3] are developed. However, in most cases, different entities need different codes. As a result, the maintenance work of software becomes very hard.

2.2 Script Based Method

In this method, an Entity Definition Script (EDS) is given to describe the geometry, physical and behavior properties of an entity and a compiler is also given to compile the script in the runtime. This method can be divided into two types [4]: geometrical property based and behavioral property based. The first type includes VRML [5], OML [6] and Starship [7] etc; the second type includes SCPL [4] and ABL[14].

As this method is focused on the EDS, the users/ developers have to be familiar with it and to be able to integrate their own codes to the runtime software (such as 3D rendering and network communication).

Table 1. Differences of Entity Developing Methods

	Program Based	Script Based	Toolkit Based
Entity Description	No	Yes	No
3D Effect Module	User writing	User writing	No
Network Module	User writing	User writing	Yes
Multi Articulation	User writing	No	No
Convenience	Bad	Medium	Good
Maintainability	Bad	Medium	Good
Expansibility	Bad	Medium	Good
Main User	Programmer	Programmer	Professional

2.3 Toolkit Based Method

This method uses toolkit to help the users develop an entity. A toolkit usually has a library containing many basic entities. The user can either choose one of or combine some of these basic entities via user interface to get an entity which is close to their target, and then add the specific behaviors as demands. This method is widely used in the DIS application development, the tools includes VRForce [8], STK [9] etc. These tools usually have a network communication component and some of them even follow DIS or HLA standard.

There are some limits to these tools as follows:These tools only provide 2D but 3D UI. Although 2D UI does help simplify development procedure, in some particular cases when a user need 3D user interface, he has to use the library of a toolkit and write program by himself (Like Programming Based Method). These tools also do not provide multi articulation support, so the users can not control the components of the entities. The differences of the three methods are shown in table 1.

In this work we developed the Entity Modeling and Simulating Platform (EMS) as an entity development toolkit for distributed virtual environment. The EMS overcame the limits of toolkit based method by providing more functionalities, including 3D GUI and multi-articulation support. Now users get the control over both the entities and the articulations of them via 3D GUI.

3 System Overview and Main Modules

3.1 Architecture

The EMS comprises two main modules: the modeling module and the simulating module, which is designed mainly for the individual entity developing based on MDA (Model Driven Architecture) principle. To be able to interactively communicate with each articulation of the entity and to create multi articulation entity automatically forms the basic requirements of developing the EMS. Nowadays the EMS has achieved its communication through the whole network by adopting the IEEE1516 (High Level Architecture) standard. The entities generated from the EMS can run on different nodes in the network and communicate with each other. Its architecture is as follows:

3.2 Modeling Module

The key functionality provided by modeling module is to facilitate the users to define both the properties of an entity and the relationship among these properties. These definitions then will be stored in a XML format file. Another functionality provided by modeling module is environment information modeling etc. Main characteristics of this module are as follows:

-Supporting multi articulation entity modeling.
-Providing Entity Template to describe an entity.
-Providing GUI Interface.
The key technologies of this module are as follows:

-Skeleton of Entity for Multi-articulation
The most important and complicated part of the multi-articulation entity developing is how to make as many as possible articulations of the entity controllable by the user as well as to keep the structure of the entity remaining clear For this purpose, we make the following basic definitions:

Definition 1: Component. Component is a user controllable articulation part of entity.

Definition 2: Action. Action is a kind of property of a component, by which the user can control the entity.

Definition 3: Attribute. Attribute is a variable belongs to a component. Each component can be controlled by the user.

The component can be described as follows:
Component = {GeometryName, ActionSet, AttributeSet}.

GeometryName is the name of a component in the geometry model, ActionSet contains all the actions by which user can control a component (To make EMS more reusable, each action is a dll module. The user can rewrite the action dll.), AttributeSet contains other variables belonging to a component, two important variables of them are the position and rotation of the entity.

To help describe the relationship between different components, we also define the skeleton of the entity. The skeleton consists of many joint points which are used to store the DOF information of adjacent points. The user can either define by themselves or select a propagation module that can be used between adjacent points.

-Entity Template
The core of this module is the Entity template (ET) that abstracts the main properties of individual entity, including geometrical, physical and behavioral properties. From a developer's side, the ET can be regarded as a Geometry Based Entity Description Script (EDS) as well as the modeling module can be regarded as a way to visualize these script by which the users can set the parameters of the script.

The runtime result of this module is shown in Figure 4.

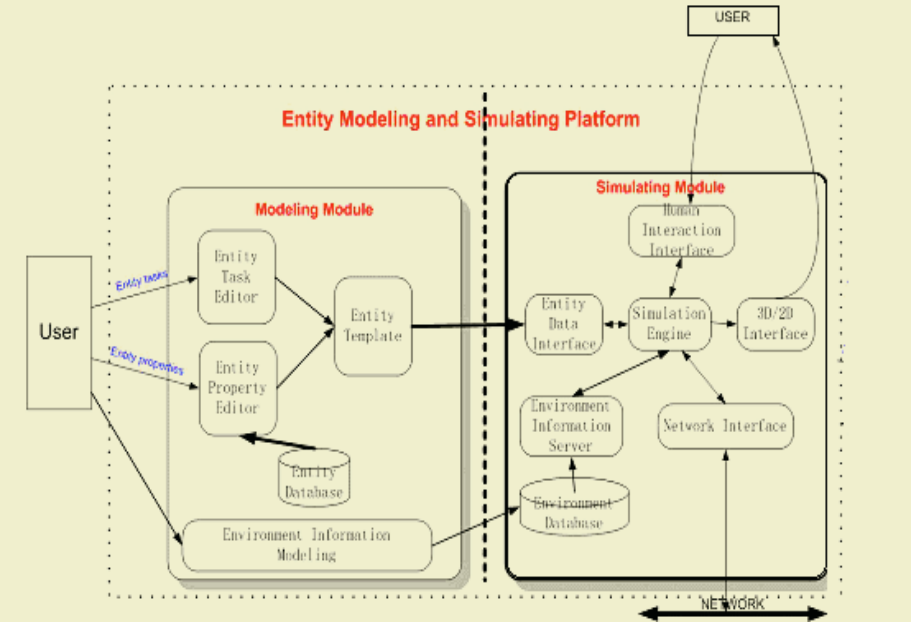


Fig. 3. The EMS Architecture

3.3 Simulating Module

This module is the runtime part of the EMS. Before running this module, the description script of the entities must have been generated by the modeling module. This module then interprets the entity description script and makes the entities to run. There are four windows available from different aspects to display the runtime information of the distributed virtual environment: the 3D window is in charge of displaying the DVE by using 3D graphic, the 2D window

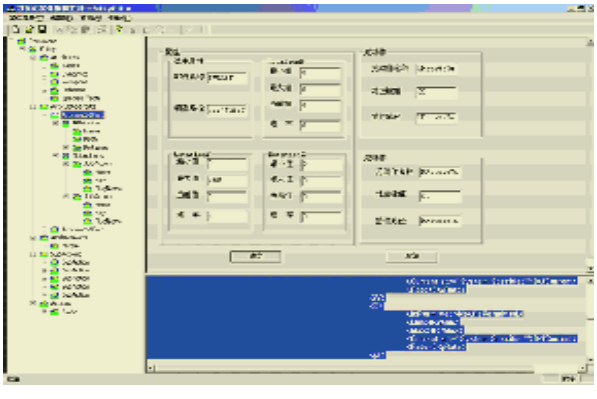


Fig. 4. The Runtime Result of Modeling Module

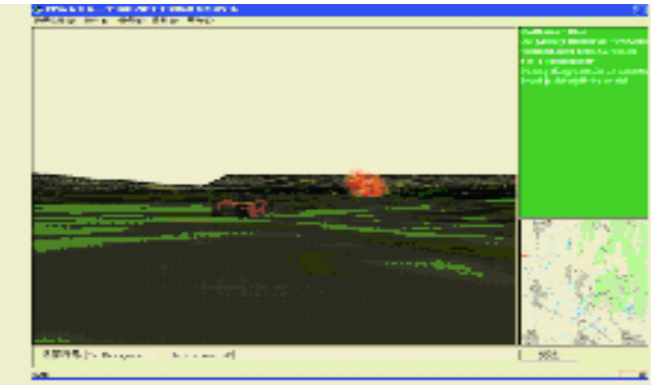


Fig. 5. Runtime Result of the Simulation Module

works as the plan view of the DVE, the EVENT widow is for displaying the event information of the DVE (such as collision, fire and explosion) and the COMMunication window allows the users to chat through the DVE. The runtime result of the SM is shown in figure 5.

The characteristics of this module are as follows:

-Providing 3D user interface by which the users can control each component of the entity through 3D interface.

-Providing IEEE 1516 (HLA/RTI) network interface.

This module comprises four main sub modules:

-Simulation Engine

Simulation Engine is the core of the SM. Its main tasks include loading and interpreting the ET file, controlling the progress of the simulation.

-Environment Information Server(EIS)

To make the DVE usable, it should not only provide the visual information (geometry and image) of the DVE, but also provide the attributes information of it. In the EMS, the EIS sub-module provides the environment information such as elevation of terrain, features of terrain and atmosphere etc. The EIS does not actually generate the geometry model of target entity like terrain and sky, it gets the information from the model files such as OpenFlight and 3ds and sets up a communication between the users and the DVE to allow the users interactively change the environment attributes.

To support different applications, EIS adopts multi-layers and multi-resolution regular grid to store the attributes, as shown in Figure 6.

The different layers of the grid are used to store different information of the terrain seperately, each layer corresponding to an attribute of the terrain such as buildings, roads, rivers, soil type etc. The different resolution of the grid is used according to the different types of the entities (such as plane, car or virtual human).

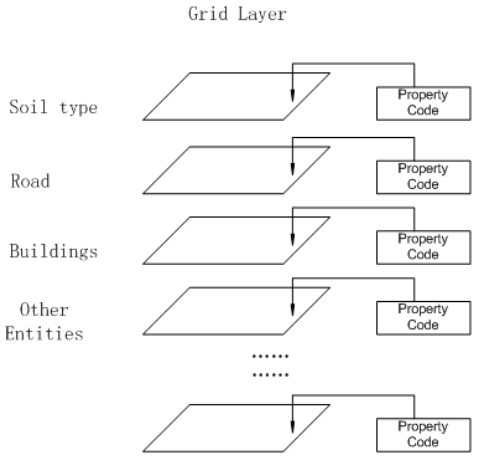


Fig. 6. Multi-layer Grid of the Environment Information

Now we use the component technology to implement the EIS, the runtime result of it is as follows:

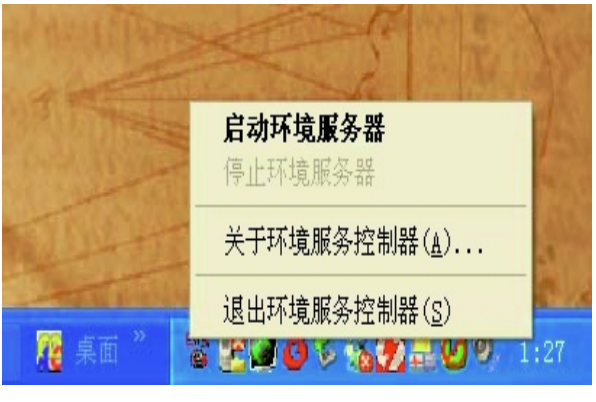


Fig. 7. The EIS Runtime Result

-3D Interface

This Interface provides a 3D graphics user interface by which the users can not only view the DVE but also control the entity in a 3-dimentional mode. This interface is based on a 3D graphic engine called VEE (Virtual Environment Engine) also developed by ourselves, providing an object level control in the virtual environment. The application programming interface (API) of the VEE includes loading model, setting position or orientation of an object, testing interaction, setting global light etc. The core of the VEE is a scene graph which can be regarded as data structure. The rendering process of the VEE is shown in fig. 8.

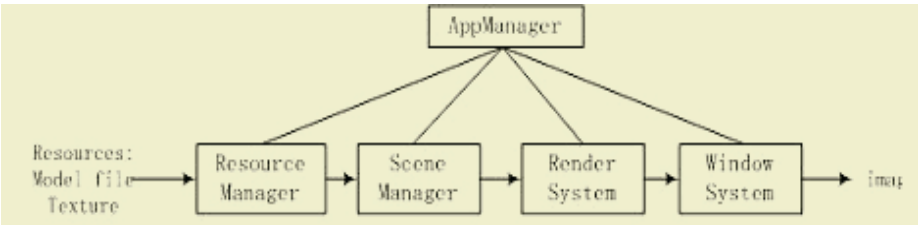


Fig. 8. Structure of the VEE

The key technology adopted in the 3D interface of the EMS is the primary attributes set., aiming to save users/developers from repeating coding work each time when a new entity need to be added. By defining a lot of primary attributes, the 3D interface can be independent of a particular entity type. So the 3D codes don't need to be changed whenever a new entity is added into the DVE.

-Network Interface (NI)

The network interface is in charge of providing network communication service, which is an essential sub-module to develop in creating a DVE. To make the network interface more reusable, IEEE launched two important standards: IEEE 1278 in 1995 and IEEE 1516 in 2000. Comparing the two IEEE standards, we found that IEEE 1278 concentrates on the protocol itself of a DIS application, on the other hand, 1516 is focused on the interface specification (relating to the management) of a DIS application and the RTI. Table 1 in [10] illustrates the main differences between IEEE 1278 and IEEE 1516 as we have pointed out.

In this work, the network interface of the EMS supports IEEE 1516 interface (High Level Architecture and Runtime Infrastructure) and both the Federation Object Model (FOM) and Simulation Object Model (SOM) are similar to the entity template (ET).

4 Conclusion and Future Works

In this paper, we introduced the Entity Modeling and Simulating Platform for the Distributed Virtual Environment. Assuming that the end users don't have a particular knowledge of computer programming, the EMS needs to be transparent for the users. All the users need to do is to set the properties of entity and the relationship among these properties. The rest work then will be done by the EMS based on the information provided by the users, including automatically generating the entity program integrated with interactive 3D user interface and maintaining communication among entities in the network., The main technologies involved in the development of the EMS include entity template, skeleton of entity, simulation engine, environment information service etc. Using the EMS can accelerate the process of the development of the distributed virtual environment. Now the EMS has also been used in the development of the DVENET. The future research work can be carried on in improving the overall performance of the EMS from the following aspects: adding more realistic 3D effort of the

3D user interface, achieving more standardized interface of the action DLL and optimizing the action set.

References

1. IEEE Standard for Information Technology, Protocols for Distributed Interactive Simulation Application, Entity Information and Interaction, Order No. SH16154, IEEE Service Center, 445 Hoes Ln, Piscataway, N.J, 08855-1331, 1993.
2. IEEE Standard 1516.1-2000, IEEE Standard for Modeling and Simulation , High Level Architecture.
3. Wookho son, kyunghwan Kim, Nancy M. Amato, Jeffrey C. Trinkle. A Gneralized Framework for Interactive Dynamic Simulation for MultiRigid Bodies, IEEE trans on system, man and cybernetics, Vol.34, No.2, 2004(4)
4. Zheng Yuan, Li Sikun, Hu Chengjun, Peng Yuxing, Virtual Artifact Modeling Language SCPL, Chinese Journal of Computers, Vol.22, No.3, 1999(3). (In Chinese)
5. The Virtual Reality Modeling Language, Version 2. 0. ISO?IEC CD 14772, 1996. 6.
6. Green M. Object Modeling Language (OML) Version 1. 1, Programmer's Manual. Alberta, Canada: University of Alberta.
7. Gurminder Singh et al. BrickNet: Sharing Object Behaviors on the Net. In: Proc Virtual Reality Annual International Symposium, North Carolina, 1995.
8. VRForce: http://www.mak.com.
9. STK: http:// www.stk.com.
10. LIANG Xiaohui, LU Weiwei, LIU Zhiyu, ZHAO Qinping, SOADIS: An New Architecture for Distributed Interactive Simulation, 8th International conference of CADCG, Marco, 2003.11.
11. Philip Willis, Virtual Physics for Virtual Reality, IEEE Proceedings of the Theory and Practice of Computer Graphics 2004
12. D. C. Ruspini and O. Khatib, A Framework for Multicontact Multibody Dynamic and Simulation Haptic Display, in Proc. IEEE Int. Conf. Intel.Robotic Syst., Oct. 2000,
13. J. Cremer., Issac, http://www.cs.uiowa.edu/
14. Michael Mateas, Andrew Stern, A Behavior Language for Story-Based Believable Agents, IEEE Intelligent Systems,2002(4)
15. B. Mirtich, Multibody dynamic package, http://www.merl.com/projects/rigidBodySim

Footprint Analysis and Motion Synthesis*

Qinping Zhao and Xiaoyan Hu

Virtual Reality & Visualization Institute,
Beijing University of Aeronautics and Astronautics,
Beijing, 100083, P.R. China
{zhaoqp, huxy}@vrlab.buaa.edu.cn

Abstract. In this paper, we describe a novel method to analyze captured human motion data. As described later, we analyze the sequences of footprint of the moving figure to extract different motion stages, as well as the switch frame between two adjacent stages. The displacement between left and right footprints is a time varying function. The change of motion (in direction, posture, speed etc.) can be reflected by the displacement between left and right footprints. Each motion stage is a sequence of motion data. Two adjacent motion stages can be different just in movement direction despite they are in the same movement style. In the second half of this paper, we apply footprint analysis on a short human motion clip of walking forward, and synthesis new motions of walking (running) along a path, of which the direction can be changed interactively in running time.

1 Introduction

Realistic human motion synthesis is a rising research area in computer graphics and computer animation in recent years. Many advances have been made in both physics/biomechanics based method and motion capture based method. Here, we roughly divide these methods into two categories. The first class is called physics/biomechanics based methods. The methods of this class generate motion mainly based on physics and biomechanics knowledge such as direct and inverse kinematics, direct and inverse dynamics, dynamics and controllers.

The methods of the second class are based on the data captured by motion capture equipment, are called motion capture based methods. With advances in motion capture techniques, people can record human motion data from real actors. Applying the captured data to computer generated characters can achieve fine realistic motion. The pattern and structure of the motion data are unknown. It is hard to describe the data by several parameters, and it is even more difficult to edit the data. In order to reuse the collected motion data, researchers have built several excellent models for human motion data e.g. motion graph [14],

* This research work is supported by China High Technology Development Project (863 project), Number: 2004AA115130 and China National Key Basic Research and Development Program (973 Program), Number: 2002CB312105.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 235–244, 2005.

motion texture[25]. The basic mechanism of most of these models is separating one motion clip into a series of segments of frame sequences, and building transition matrix or paths or other relationships for those segments. Some researchers follow another novel way by introducing signal processing techniques into motion analysis and synthesis. They use temporal domain and frequency domain analysis techniques on human motion to blend or change motion. MRA(Multiresolution analysis) [11] makes it easy to modify motion at different resolution levels and provides more control on synthesized human motion.

Currently, most work focuses on analyzing motion data not only building a pure mathematics model but also introducing physics constraints onto the model. More physics realistic motion then can be synthesized in this way.

As more and more techniques on human motion synthesis emerging, the need for evaluation on synthesized result seems rising. Paul et al. has done some interesting work on it [18].

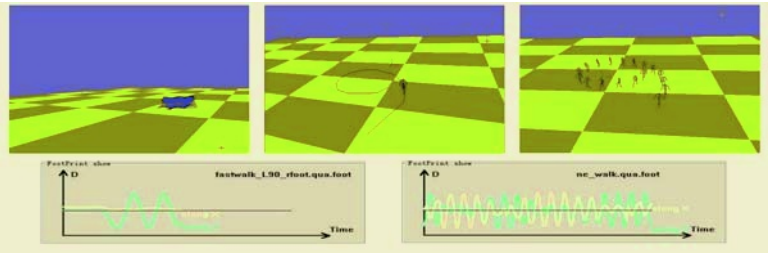


Fig. 1. (*upper left*) A short motion stage of walking forward, (*upper middle*) and (*upper right*) synthesized motion of walking along a specified path, (*lower left*) the displacement of footprints of the motion stage in (*upper left*), (*lower right*) the displacement of footprint of the synthesized motion

This paper focuses on analysis of footprints of moving figure rather than the full skeleton motion data. Footprint Analysis has some advantages. First, position of footprint in 3D space is naturally three dimensional function of time. Secondly, the sequences of footprint show the path of the motion. Lastly, the capability of generalization: Walk with head nodding and walk with hand shaking may be the same type of motion called "Walking" or so in footprints analysis despite having different upper-body motion. In fact, it can be seen later that a clip motion of walk and run are looked as the same type and dealt with by the same technique.

The remainder of this paper is organized as follows. In section 2, we discuss related work. In section 3 we introduce our method of footprint analysis and build a set of parameters to describe one footstep cycles and motion stages. Section 4 focuses on our motion synthesis approach. In this section, foot skate is cleaned up in a simple way and form a short motion stage a long motion which follows a user specified path is synthesized. Section 5 discusses the result. Lastly, we summarize our contributions and outline possible future work.

2 Related Work

Realistic human motion synthesis is an active research area. Since our work is based on capture motion data, the following brief review of related work is mainly focused on motion capture based methods.

There are many studies on footprint and gait of human walking in biomechanics [17][16]. R. Boulic et al. proposed a human walking model which produces the values of the spatial, temporal and joint parameters for each frame. They treated walking motion as walking cycles according to the footprint on the ground [6].

Foot as constraint is used in motion control [5][19]. In [19], the foot placement of support leg served as a kinematics constraint, then footprint planning was generated in high-level locomotion control and all body joint angles were adjusted in low-level control. Kovar et al. [15] described a simple method to remove foot skate using IK algorithm. Anthony et al. [2] used a set of objective functions and constraints that lead to linear time analytical first derivatives including constraints on physical validity, such as ground contact constraints. Torkos et al. [22] proposed a footprint model and scattered the footprints along a user specified path, and, reconstructed the motion using inverse kinematics.

To analysis raw motion data directly is difficult. So researchers built many mathematical models for captured motion data. Since human motion is recorded as motion sequences which are both nonlinear and time-varying, effective models of human dynamics can be learned from motion capture data using statistics and stochastic processes theory. These models include hidden Markov model (HMM) [8][1][20][21], directed graph [14][3], and statistical model [25][7].

Techniques from the image and signal processing domain have been successfully applied to designing, modifying, and adapting animated motion [9]. Unuma et al.[23] used Fourier expansions of experimental human motions to interpolate or extrapolate the human locomotion. Witkin and Popovic [24] modified a reference trajectory by interactively turning the position of selected key-frames and by scaling and shifting the angles. Jehee Lee et al. [13] [11] edit the motion data with a hierarchy of displacement maps. In Jehee's work, Multi-Resolution analysis is applied on motion data, and the motion data is represented as a Multi-Resolution hierarchy and be reconstructed at different resolution levels.

The system described in [12] allows users to control the avatar in virtual world by making desired motions in front of a single video camera interactively. Arikan et al. [3] proposed an approach which can generate motion sequences that satisfy a variety of constraints automatically in real time, and then users can author complex motions interactively. Kovar et al.'s work [14] also provides users the ability to create natural walk motions along arbitrary paths under user's control.

In this paper we analyze the displacement of footprints to separate one motion into different motion stages and extract some info about each motion stage such as direction and speed. We also develop a simple foot skate removal algorithm when generating motions of walking (running) along a path specified by the user interactively.

3 Footprint Analysis

In this paper we focus on the motions that move body from current location to a particular destination. And we call these motions Move Motivated Motions (MMM or M3). In M3 the footprint of moving figure is important because that the change of body location is achieved by moving legs alternately and is reflected by the footprint on the ground. Fig. (2) shows the footprint of a clip motion of walking forward and then turning left.

Obviously, the footprint can start at any position, and we would apply some preprocess on it to extract relative position.

3.1 Footprint Displacement

To extract relative position from footprint of a motion, we subtract the position of start footprint from all the sequence footprints. Thus, the first footprint becomes origin and the entire footprint is relative to this point. We go further to extract the displacement of left and right feet projection on the ground at each frame. We define displacement of footprint as follow:

$$\begin{aligned} FP'_i &= FP_i - FP_0 \\ D_i &= FP'_{i\,left} - FP'_{i\,right},\ (i = 0, 1, 2 \ldots, n) \end{aligned} \tag{1}$$

In Eq. (1), D_i is the displacement between left and right feet; FP_i is the original footprint of the moving figure at i frame; FP_i' is the counterpart of FP_i after preprocess, and n+1 is the length of the motion clip.

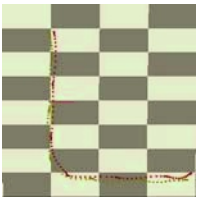


Fig. 2. The footprint of a walking motion: walk forward, then turn left and walk forward again

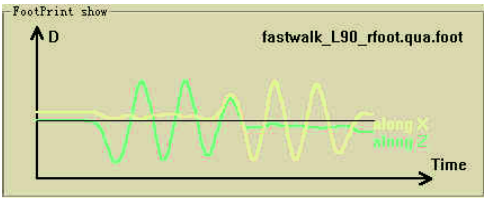


Fig. 3. The displacement of the footprints of a walking motion: walk forward, turn left and forward again. (yellow: displacement along X axis; green: displacement along Z axis)

Fig. (3) shows the displacement of left/right feet along X and Z axes. As can be seen in Fig. (3), we ignore the Y axis because the projection of feet on the ground is more essential for analyzing the Move Motivated Motions, and this makes it much easier to analysis.

3.2 Motion Stages and Parameters

In order to describe the motion data more precisely we make the following definitions.

Footstep: Here we define a motion cycle as the motion sequence between two adjacent foot takeoff and foot strike. As described later, we replace adjacent takeoff and foot strike with two adjacent wave crest and trough in Fig. (3). This means that one footstep can be regarded as a half period of the displacement curve shown in Fig. (3).

Path: the path of the motion clip is represented by the footprint itself to represent. The reason for not using the root joint (centre of mass) is that, to match two trajectory of root joint in 3D space means more strict constraints and requires these two motions almost exactly the same. For most M^3 the path often has a clear destination, but for non-M^3it is not the case.

Direction of footstep cycle: Direction is defined for each footstep cycle in this paper. Here we define the direction cycle to be a vector that starts from the takeoff and ends at the next foot strike.

$$DV_j(DirectionVector) = FP_{j(footstrike)} - FP_{j(takeoff)} \tag{2}$$

DV_j is the direction vector of the j-th footstep cycle. According to the above definition, we can tell the footstep and direction from the direction vector of one footstep cycle.

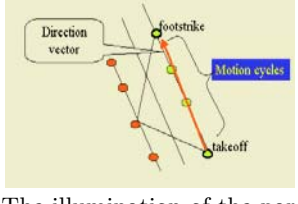


Fig. 4. The illumination of the parameters

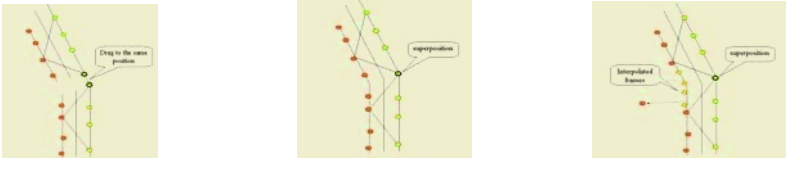


Fig. 5. footskate removal: (*left*) Two rearranged motion stages, (*middle*) drag the second motion stage to superpose the two feet, (*right*) remove the inter frames and interpolate transition frames

Motion Stage: As many other models or methods in motion editing and motion synthesis, we would break a given motion sequence into several clips for further analysis and reuse. We call each motion clip Motion stage. A motion stage is an instance of a particular pattern of motion. So two motion stages may belong to two different patterns, or the same pattern but with different parameter values. Motion stage is more like "object" in Object Oriented Programming Language rather than "class".

To extract motion stages from the original motion sequence, a two pass process is used. In the first pass, we break the original motion into segments of M^3 and non-M^3. To achieve this, we compute the angle between two adjacent direction vectors by the following formula.

$$\alpha_j = \frac{DV_j \cdot DV_{j+1}}{|DV_j| \cdot |DV_{j+1}|} \quad j = 0, 1 \ldots k \tag{3}$$

Here, k is the number of total footstep cycles in the given motion. Then a low-pass filter is applied on all the angles (α_j). The passed segment which has 5 or more footstep cycles would be a motion stage. Footsteps in the same motion stage have approximate directions.

Switch frame: When the motion breaks into motion stages, the problem of how to synthesize new motions with these motion stages rises. The original motion seems fine at first, but turns to be unnatural when we exchange the order of its motion stages. We should find the key point to integrate them smoothly – switch frame. The first and last frame is naturally the switch frames when we arrange the motion stages as they are in the original motion. In other cases, constrain on feet should be considered, and we present a simple way to satisfy it in the next section.

Direction of motion stage: the direction of a motion stage is a vector that starts from beginning footprint of the motion stage and points to the last footprint of the motion stage. As the direction of footstep cycles, this direction vector indicates the net displacement and change in direction of the motion stage.

4 Motion Synthesis with Motion Stages

The basic idea of our motion synthesis method is very simple: to rearrange the motion stages directly. There are more things to do to make the simple idea work. First, we should allow users to predefine a path even change the path in real time. Secondly, the motion stages should be arranged along user specified path. And lastly, we would find out the switch frame of each motion stage and smooth the connection between them.

4.1 Path Design and Arrange the Motion Stages

Users can define a path as well as the motion stages along the path. Each motion stage has its own net displacement and direction, these two properties may do not match the specified path very well. For some kind of motion stages such as walking forward, it is not a problem because we can extend it to satisfy the path constrains. For those non-M3 motion stages, we loop the motion stage along the specified path and fill the remainder up with walking motion. The remainder is deleted if it is small enough.

Rotation of Motion Stage: To fit the predefined path, the direction of the motion stage should be changed to be consistent with the path. This is easy to achieve

by rotating the entire motion stage to new direction. When applying a rotation to the motion stage, the original of coordinate system is an invariant and should be chosen carefully.

4.2 Smooth the Synthesized Motion

Smoothing the synthesized motion is divided into two processes: Find the switch frames between the adjacent motion stages; Then interpolate the transition frames between the adjacent motion stages with Cardinal Curve.

Finding Switch Frames: Two adjacent motion stages (e.g. walk forward in two different directions), should have approximate motion state at the conjunct frames so that we can smooth them to produce a natural transition. An important constrain on feet is to cleanup foot skate. We solve this constrain by drag the successive motion stage to superpose the feet of the conjunct frame of both predecessor and successor motion stages.

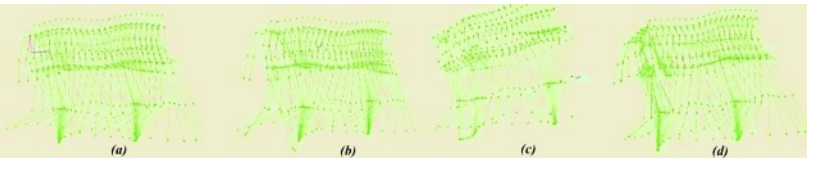


Fig. 6. Interpolating between two adjacent motion stages: (a) Motion stage of walk forward, (b) connect two motion stages directly, (c) interpolated from (b), need foot skate removal, and (d) interpolated after drag two motion stages to superpose their feet

According to the "Characteristic phases of a walking motion" [Multon et al. 1999], we select the support feet to superpose. So, we should find out the corresponding frame in the successor motion stage which has the same support foot (e.g. left foot) with the end frame of the predecessor. As shown in Fig. (4), one footstep cycle is depicted by a half period curve from one crest to the neighboring trough. To find the best fit frame we define the error function between two frames as follow:

$$error_{ij} = \sum_{k=2}^{k=n} \|joint_{ik} - joint_{jk}\| \tag{4}$$

Where error $_{ij}$ is the error between i-th frame and j-th frame, and n is the total number of joints. This error function is a full skeleton compare between two frames. In Equation (4), k starts from 2 because the root joint is discarded in the error function. This makes the error be calculated in local-coordinate system.

The definition of the distance function used in Equation (4) depends on the representation of joint rotation angles. In our work, we use exponential maps [Jehee Lee et al. 1999] and we define the distance as follow:

$$\|j - j'\| = \sqrt{((j_x - j'_x))^2 + ((j_y - j'_y))^2 + ((j_z - j'_z))^2} \tag{5}$$

Here j and j' are the joint rotation angles, and x,y,z are the three components of j and j'. This distance equation is similar to Euclidean distance in the form.

Suppose that the index of the last frame of predecessor motion stage is k, and the first footstep cycle of the successor motion stage is ranged from s to t, we compute the errors between them – $error_{ki}$, for i =s, s+1...t. and select the frame with the lowest error. We call this error Error(K), and record this frame to be the corresponding switch frame in the successor motion stage. Further more, we can compare Error(K-2) and Error(K-1) with Error(K) to select a more appropriate switch frame in the predecessor motion stage.

Interpolate: The switch frames do not always match each other very well, so we should interpolate some transition frames between this two motion stages. Suppose that the switch frame is k-th frame of the successor motion stage, we would replace the frames range from 0 to k-1 with the interpolated frames. This makes the transition between two motion stages more smooth.

There are many interpolate techniques for motion synthesis, here, we take the Cardinal curve to accomplish the job. Cardinal method has some advantages as follows: First, Cardianl spline is easy to implement. The algorithm is simple and efficient. Secondly, Cardianl spline passes the control point. In our application, the endpoints are switch frames of adjacent motion stages. Lastly, Cardinal spline provides user the ability to control the tightness of the interpolate curves.

5 Result

We choose walk and run motion to analyze in our work because they are two typical human motions and are move motivated motions. It took about 4 ms to extract 3 motion stages from a clip of 187 frames walk motion on Inter Pentium IV 1.7G computer with 256M memory. To deal with a more complex dance motion of 1470 frames, it took approximately 10ms and extracted about 23 motion stages on the same computer. The algorithm to extract motion stages from a perticular motion has a linear complexity -O(n), where n is the number of footstep cycles.

To synthesize a new walk/run motion along a specified path is not a time consuming job. The implementation of our interpolating algorithm based on Cardinal splines involves $4 \times n \times n$ additions where n is the number of generated frames and k is the number of total DOFs of the skeleton. In our experiments, n is 8 and k is 57.

6 Conclusion and Future Work

We present a novel way to analyze human motion - footprint analysis. Footprint analysis considers the motion as a series of motion stages which consist of footstep cycles. And both motion stage and footstep cycle are described by a set of intuitionistic parameters. This representation of motion is more perceptually and people can grasp the characters of the original motion without replaying it.

Different from most methods which divide the motion into several motion patterns, motion stage is more precise for distinguishing two instances of the same motion pattern.

The footprint is mostly related with the lower-body of the skeleton. So the same motion stages may have different upper-body motions. We call this the capability of generalization. With the capability of generalization, we can apply the same technique on a set of same motion stages with different upper-body motions.

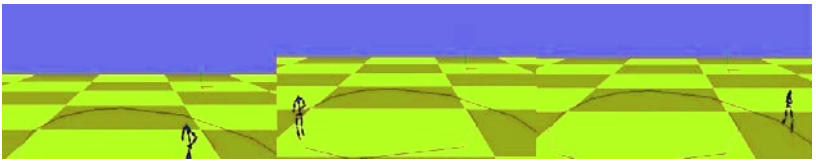


Fig. 7. Synthesized motion of running along a specified path: several screenshots of the synthesized running motion

References

1. A. GALATA, N. JOHNSON, AND D. HOGG.: Learning Variable Length Markov Models of Behaviour. Computer Vision and Image Understanding. (2001) 81(3) 398- 413.
2. ANTHONY C. FANG, NANCY S. POLLARD.: Efficient Synthesis of Physically Valid Human Motion. , SIGGRAPH 2003 Proceedings. ACM Transactions on Graphics. (2003) 22(3) 417-426.
3. O. ARIKAN, D. A. FORSYTH.: Interactive Motion Generation from Examples. In Proceedings of SIGGRAPH 2002.
4. ARIKAN, O., FORSYTH, D. A., O'BRIEN, J. F.: Motion Synthesis from Annotations. In Proceedings of ACM SIGGRAPH 2003.
5. N. I. BADLER, C. B. PHILLIPS, B. L. WEBBER.: Simulating Humans: Computer Graphics Animation and Control. Oxford University Press, New York, Oxford. (1993)
6. R. BOULIC, N. MAGNENAT-THALMANN, D. THALMANN.: A Global Human Walking Model with real time Kinematic Personification. Visual Computer, Vol.6, No6, (1990) 344-358.
7. BOWDEN, R.: Learning Statistical Models of Human Motion. In IEEE Workshop on Human Modelling. Analysis and Synthesis, CVPR2000. (2000)
8. M. BRAND AND A. HERTZMANN.: Style Machines. In Proceedings of ACM SIGGRAPH 2000. (2000) pages 183- 192.
9. A. BRUDERLIN AND L. WILLIAMS.: Motion signal processing. In Proceedings of SIGGRAPH '95. (1995) pages 97-104.
10. F. FEBASTIAN GRASSIA.: Practical Parameterization of Rotations Using the Exponential Map. Journal of Graphics Tools. (1998)
11. JEHEE LEE AND SUNG YONG SHIN.: Multiresolution Motion Analysis with Applications, The international workshop on Human Modeling and Animation, Seoul. (2000) 131-143.

12. JEHEE LEE, J. CHAI, P. S. A. REITSMA, J. K. HODGINS, N. S. POLLARD 2002. Interactive Control of Avatars Animated with Human Motion Data, ACM Transactions on Graphics, SIGGRAPH 2002 Proceedings. (2002) 21(3), PP. 491-500.
13. JEHEE LEE AND SUNG YONG SHIN.: A Hierarchical Approach to Interactive Motion Editing for Human-like Figures. Computer Graphics (Proc.SIGGRAPH '99) 33 (AUG.), (1999) 395-408.
14. KOVAR L., GLEICHER. M. AND PIGHIN. F.: Motion graphs. In Proceedings of SIGGRAPH 2002, (2002) 473-482.
15. KOVAR L., GLEICHER. M. AND SCHREINER. J.:Footstake cleanup for motion capture editing. In ACM SIGGRAPH Symposium on Computer Animation (2002) 97-104.
16. LAMOREUX. LW.: Kinematic Measurements in the Study of Human Walking, Bulletin of Prosthetics Research, (1971) Vol.BPR 10, No15.
17. J. NILSSON, A. THORSTENSSON AND J. HALBERTSAM.: Changes In Leg Movements and Muscle Activity with Speed of Locomotion and Mode of Progression in Humans. Acta Physiol Scand. (1985) 457-475.
18. PAUL S. A. REITSMA NANCY S. POLLARD.: Perceptual Metrics for Character Animation:Sensitivity to Errors in Ballistic Motion. In Proceedings of ACM SIGGRAPH 2003.
19. SHIH-KAI CHUNG AND JAMES K. HAHN.:. Animation of Human talking in Virtual Environments, Computer Animation 1999 (CA'99), 26-28 May, 1999, Geneva, Switzerland, Proceedings. IEEE Computer Society, 1999
20. L. M. TANCO AND A. HILTON 2000.: Realistic synthesis of novel human movements from a database of motion capture examples. In IEEE Workshop on Human Motion.
21. TIANSHU WANG, NAN-NING ZHENG, YAN LI, YING-QING XU, HEUNG-YEUNG SHUM. Learning Kernel-Based HMMs for Dynamic Sequence Synthesi. Pacific Conference on Computer Graphics and Applications (2002) 87-95.
22. TORKOS, N., AND VAN DE PANNE, M.Footprint-based Quadruped motion synthesis. In Graphics Interface '98. (1998)151-160. ISBN 0-9695338-6-1.
23. M.UNUMA, K. ANJYO, AND R.TAKEUCHI.: Fourier principles for emotion-based human figure animation. In Proceedings of ACM SIGGRAPH.(1995) pages 91-96, Los Angeles, California, August 1995. Addison Wesley.
24. A. WITKIN AND Z. POPOVIC. Motion wraping.: In Proceedings of ACM SIGGRAPH, pages 105-108, Los Angeles, California, August 1995. Addison Wesley.
25. YAN LI, TIAN-SHU WANG, HEUNG-YEUNG SHUM.: 2002. Motion texture: a two-level statistical model for character motion synthesis. ACM Trans. Graph. (2002) 21(3): 465-472.

An Adaptive and Efficient Algorithm for Polygonization of Implicit Surfaces

Mingyong Pang[1], Zhigeng Pan[2], Mingmin Zhang[2], and Fuyan Zhang[1]

[1] Department of Computer Science and Technology, Nanjing University, Nanjing 21009, P. R. China
panion@netease.com
[2] State Key Lab of CAD&CG, Zhejiang University, HangZhou 310027, P. R. China
zgpan@cad.zju.edu.cn

Abstract. This paper describes an adaptive and efficient algorithm for polygonization of implicit surfaces, which consists of two steps: initial polygonization and adaptive refinement. The algorithm first generates an initial coarse triangular mesh from implicit surface using a variation of the traditional Marching Cubes (MC) Algorithm. And then the triangles in the coarse mesh are iteratively subdivided by employing a sampling rate that varies spatially according to local complexity of the surface. The new created vertices in refined mesh are projected onto the implicit surface by gradient descent method. Consequently, the algorithm produces the minimum number of polygons required to approximate the surface with a desired precision and the final mesh is simplicial complex. Our algorithm can be used in the real-time environment of visualization of implicit surfaces.

1 Introduction

Implicit surfaces are represented by mathematical functions of the form $f(x, y, z) = 0$, where for arbitrary functions $f(x, y, z)$, it is not possible to explicitly express x,y and z in terms of parameters generally. The functions classify points (x, y, z) in $\mathbf{R}^3$ space as inside the shape, outside the shape or on the shape's surface according to sign of $f(x, y, z)$. The advantage of implicit technique over that of parametric counterpart is that very intricate geometric and topological (smooth) shapes can be described extremely quickly using a skeletal model[1]. Additionally, typical modelling operators such as Constructive Solid Geometry (CSG) are closed about implicit surfaces[2]. Since the early 1980's, implicit surfaces have become very attractive and were widely used, especially for complex object modelling, computer animation, medical data visualization and entertainment industry[3]. However, they still suffer from the lack of a good algorithm which could draw them in real-time with a high quality[4]. The main reason is that they have no 2D parameter domain to decide coordinates for a given point on surface, just like their parametric cousins, so visualization techniques of them involve inevitably a notoriously heavy computing consuming.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 245–255, 2005.

Roughly speaking, there are two types of methods to render implicit surfaces today, based on Ray-tracing[5] and polygonization techniques[4],[6] respectively. The ray-tracing based methods directly render realistic implicit surfaces from their implicit equations, but they are so slow that only can be used in post-processing rendering environment, because that they have to calculate multitudinous intersections of rays and surfaces and the implicit functions very often involve complex operations. Polygonization based methods realize rapidly visualization of the surfaces by transforming the surfaces to approximated polygonal meshes and by using widely existed graphic hardware of even nowadays cheap personal computer systems. Furthermore, sometimes a conversion to another representation is necessary for implicit surfaces because of the simplicity of polygonal meshes. So the latter becomes practically mainstream of rendering of implicit surface in interactive environment.

The algorithms of polygonization of implicit surfaces must solve two interdependent problems[7]: a) How to optimally sample vertices on the surface, and b) How to connect correctly the samples in order to create mesh topology for the surface.

Since finding roots of implicit function during sampling is inevitably involved, the most of the existing algorithms use spatial dissection technique to accelerate process of sampling. Allgower and co-workers[8] induce tetrahedrons to dissect the space; moreover, Wyvill *et al.*[9] use more simpler uniform cube units. Bloomenthal[10] brings the famous Marching Cubes (MC) method that is often used in science data visualization into polygonization, in which the cubes, which are transverse with the surface, propagate across those faces of a *seed* cube. The method effectively removes needless computation in virtue of the consistency of the surface in three-dimension space. In [11], Wyvill gives a subdivision table data structure used in Bloomenthal's algorithm, and [6] and [10] give some implementation skills of the table.

In the algorithms mentioned above, all samples are obtained by using uniform subdivision strategy without considering local characters of the surface, for example, local distribution of curvature on the surface. In this case, it is obviously that sampling is not sufficiency in the curved parts and over sampling in relatively flat regions, on the surface. For this reason, the adaptive methods are presented. [6] gives an adaptive method employing an octree data structure of spatial subdivision. But in order to avoid "crack"[6] occurred at adjacent positions between cubes that belong to different levels of subdivision, the algorithm restricts the difference that is only one between two adjacent cubes in octree. This strategy makes the algorithm complex and declines the efficiency of the algorithm. For the sake of this, so far there are not general octree-based adaptive polygonization methods of implicit surfaces. In [12], Yu *et al.* present an adaptive algorithm based on projection idea and interval arithmetic, which is appropriate to visualize a trimmed implicit surface. But it doesn't resolve adaptive problem in deed for that the algorithm is strongly relative to the direction of projection. Turk and co-workers[13] bring the potential energy model of spring system into sampling process and present a physically based adaptive algorithm,

and Figueiredo *et al.*[14] also present a discrete physically-based methods simulating particle systems to visualize implicit surface. Although the two methods resolve adaptive sampling problem finely, they involve so much excess computation that cannot be used in interactive applicant. In [7],[15], Velho presents a simple adaptive method to construct adapted polygonal meshes of implicit surfaces. The method starts with a coarse uniform polygonal approximation of the surface and subdivides each polygon recursively according to local curvature. The shortage of it is that there are many redundant triangles and "T"-edge in resulting meshes (please refer to Fig. 6, the first-left column). Ning[16] considers an inverse process: divide the space into tiny cells firstly, and then implement adapted sampling by uniting interrelated tiny cells. This method is unable to solve efficiency problem, too.

In this paper, we present an adapted polygonization of implicit surface, which is similar to Velho's method. The main differences between Velho's method and ours are: a) our method employs quick cubes-propagating algorithm to generate a coarse meshes instead of Velho's complete and uniform spatial dissection; b) our method employs new adaptive subdivision rules of triangles in meshes to subdivide mesh facets and the "T"-edge problem occurred in Velho's method is avoided.

2 Algorithm Description

The main difficult of adapted polygonization of implicit surfaces comes from two aspects: a) how to change sampling rate according to the local curvature distribution on the surface, i.e. to select appropriate number and geometric positions of sample points on surface, and b) how to keep local and global consistency of topology of mesh, i.e. to correctly connect the points to form a correct topology for mesh approximating linearly the surface. Thus, we have to design a mechanism to synchronize the two aspects. If they are dealt with incorrectly, sample points would be mistakenly connected and "hole" or "crack"[4] would appear in the resulting mesh.

In this paper, just like Velho's method, we divide the polygonization task into two stages, which called construction and adapted-subdivision respectively. The former firstly constructs a coarse mesh. In this process, polygonization is executed in a relative lower resolution and is expected to correctly grasp the topology of implicit surface. The coarse mesh provides a basic method for controlling adaptive rate of sampling in next stage, and a basic scheme to maintain consistency of topology/geometry of resulting mesh with the surface. The latter adaptively and iteratively subdivides the coarse mesh according to the curvature of the surface at a triangle facet until pre-defined tolerance is satisfied. The curvature change of the surface along with an edge of a given triangle is characterized by the inter-product of two normal vectors at ends of the edge, which can be directly calculated from implicit equation by different operator.

2.1 Computing Initial Mesh

The first step of our algorithm is generation of a coarse linearly approximated mesh for the implicit surface. The output of this process is a triangular mesh that served as the basis for adaptive refinement. Velho does this work using uniform space decomposition in [15]. But in this paper, we do it by employing a more efficient algorithm presented in [6]:

At first, an initial *seed* cube, called *cell*, which intersects surface is established. New cells propagate along established edges that intersect the surface, of the seed cell. The cells may be stored as an adjacency graph, with care requires that no cell be added redundantly to the graph. Then new cubes are expanded or propagated according to the cases that the implicit surface crosses the six faces of seed cell. All the new cells intersect with the surface in deed. Taking new expanded cells as new seeds and iteratively execute aforementioned research process, we can find total cells bestriding the surface (see Fig. 1a), which are registered in the graph. In this procedure, no computation for redundant cells or cubes is involved.

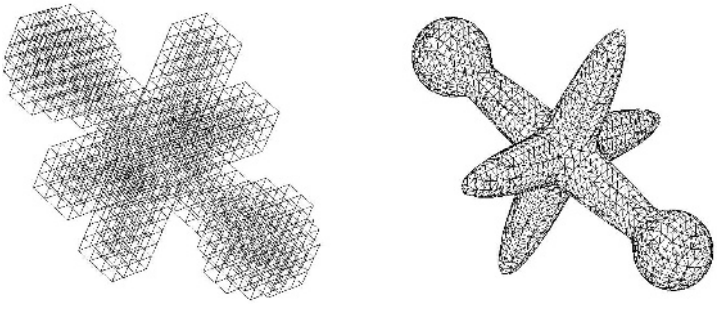


Fig. 1. Constructing initial mesh by using cell propagating method: Catched total cells bestriding the surface(left), and Mesh constructed from bestriding cubes(right)

Very different from traditional method decomposing cubes into tetrahedrons and linearizing surface to triangles in tetrahedrons, we polygonize surface parts in cubes as Fig. 2[6]. The three dimensional ordering of intersects, called *surface vertex*, of a cube can performed by a simple procedure as: begin with any surface vertex on the cube and proceeds towards the positive corner and then clockwise about the face to the right until another surface vertex is reached. This process iterates until a repeat surface vertex is found. Connect the surface vertex orderly, a closed space polygon is obtained.

For the purpose of adapted refinement, all the polygons should be triangulated by the following rules: if a polygon is a quadrilateral, connect the nearer diagonal vertex-pair and two triangles are created; otherwise, calculate the barycenter coordinate, denoted by $\boldsymbol{v}$, of the vertices of the polygon and then connect $\boldsymbol{v}$ with the vertices respectively, n (vertex number belonging to the polygon)

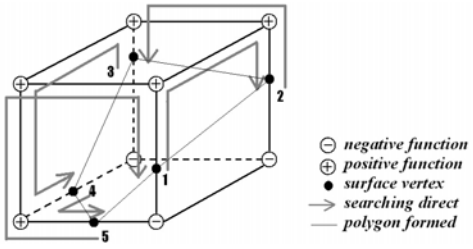


Fig. 2. Generating a 3D linearly approximate polygon from a bestriding cube[6]

triangles can be obtained. Collecting all the triangles, a coarse mesh is built naturally (Fig. 1b).

As we known, our algorithm needs to compute coordinates and its function value of vertices of each cube or cell in the course of establishing the coarse mesh. Besides, normal or gradient direction, $\nabla f = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$, at each vertex of cell should be calculated Since implicit function usually is complex, we compute centric difference instead of complex derivatives:

$$\nabla f \approx \frac{f(\boldsymbol{v} + \Delta) - f(\boldsymbol{v} - \Delta)}{2\Delta}$$

2.2 Adaptive Refinement

Whether to subdivide a triangle in coarse mesh or not depends on surface curvature change along the three edges of it. If all the edges of a triangle are plat enough (or non-curved), i.e., tolerance of difference between normal vectors at two ends of edges of the triangle is satisfied, we regard the facet approximating the surface so good that it is unnecessary to be refined further; otherwise, the triangle should be subdivided into smaller ones.

In [7],[15], as long as there is one of edges of a triangle is curved, the triangle should be subdivided into four sub-triangles (see Fig. 3). In this paper, we use a very different strategy:

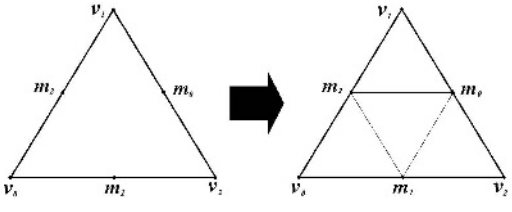


Fig. 3. 1-4 subdivision of a triangle, where $\boldsymbol{m}_i$ is midpoint of corresponding edge

– If all the edges of a triangle, $\triangle \boldsymbol{v}_i \boldsymbol{v}_j \boldsymbol{v}_k$, are curved, the triangle is subdivided into four sub-triangles and the midpoints of corresponding edges of are moved onto surface. This case is illustrated by $\triangle \boldsymbol{v}_2 \boldsymbol{v}_3 \boldsymbol{v}_4$ in Fig. 4.

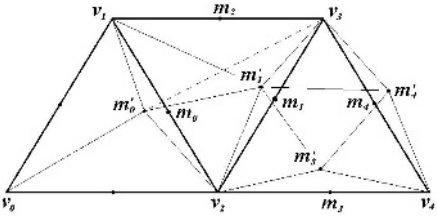


Fig. 4. Adaptive subdivision rules used in our algorithm

- If exactly two edges of a triangle $\triangle \boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{v}_k$ are curved, a 1-3 subdivision would be done. Without loss of generality, let us suppose that the two curved edges are $\boldsymbol{v}_i\boldsymbol{v}_j$ and $\boldsymbol{v}_i\boldsymbol{v}_k$ respectively, and corresponding final positions that their midpoints are moved to are $\boldsymbol{m}'_k$ and $\boldsymbol{m}'_j$ respectively, connecting $\boldsymbol{m}'_k$ and $\boldsymbol{m}'_j$ generates the first new sub-triangle, $\triangle \boldsymbol{v}_i\boldsymbol{m}'_k\boldsymbol{v}'_j$. Next work is contrasting the distances of $\boldsymbol{m}'_k \sim \boldsymbol{v}_k$ to that of $\boldsymbol{m}'_j \sim \boldsymbol{v}_j$, if $||\boldsymbol{m}'_k - \boldsymbol{v}_k||^2 > ||\boldsymbol{m}'_j - \boldsymbol{v}_j||^2$, connect $\boldsymbol{m}'_j$ and $\boldsymbol{v}_j$ to generate new triangles $\triangle \boldsymbol{v}_j\boldsymbol{v}_k\boldsymbol{m}'_j$ and $\triangle \boldsymbol{v}_j\boldsymbol{m}'_j\boldsymbol{m}'_k$; otherwise, connect $\boldsymbol{v}_k$ and $\boldsymbol{v}'_k$, $\triangle \boldsymbol{v}_j\boldsymbol{v}_k\boldsymbol{m}'_k$ and $\triangle \boldsymbol{v}_k\boldsymbol{m}'_j\boldsymbol{m}'_k$ would be created. This case illustrated by triangle $\triangle \boldsymbol{v}_1\boldsymbol{v}_2\boldsymbol{v}_3$ in Fig. 4.
- If exactly that only one edge, e.g. $\boldsymbol{v}_j\boldsymbol{v}_k$, of triangle $\triangle \boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{v}_k$ is curved, and suppose its midpoint moved onto the surface is vertex $\boldsymbol{m}'_i$, we can obtain a 1-2 subdivision by connecting $\boldsymbol{v}_i$ and $\boldsymbol{m}'_i$, and two triangles $\triangle \boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{m}'_i$ and $\triangle \boldsymbol{v}_i\boldsymbol{m}'_i\boldsymbol{v}_k$ should be resulted. See $\triangle \boldsymbol{v}_0\boldsymbol{v}_1\boldsymbol{v}_2$ in Fig. 4.
- If all the edges of triangle $\triangle \boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{v}_k$ are flat enough, the triangle isn't necessary to be subdivided anymore and the triangle is output.

For the new-generated sub-triangles, aforesaid rules are repeatedly used and subdivision is carried out to deeper levels until there is no triangle can be subdivided. Since subdivision rules are adapted and the rules disenable the generation of "T"-edge, these insure the simplexity and topological consistency of resulting mesh.

2.3 Controlling Curvature and Topological Consistency

Change of surface curvature along with a triangle edge can be measured by an angle, α, between two normals, $\boldsymbol{n}_i$ and $\boldsymbol{n}_j$, at endpoints of the edge (Fig. 5). This is because that geodesic curvature radius may be approximated by formula[17]:

$$\frac{d}{2\sin(\alpha/2)}$$

where d is length of the edge. When α is larger than threshold, α_0, the edge can be considered as curved. In interactive case, α_0 may be customized by user in order to attain the more visual effect of polygonization.

Once a triangle is subdivided, a key step, moving new-generated midpoints of edge onto the surface, must be done in order to enable mesh to approximate

the surface more accurately. This process is simple but important because it provides a mechanism to insure topological consistency between two adjacent facets, which share a common edge in mesh. For example, edge $\boldsymbol{v}_2\boldsymbol{v}_3$ of triangle $\triangle \boldsymbol{v}_1\boldsymbol{v}_2\boldsymbol{v}_3$ in Fig. 4 is curved, then its midpoint $\boldsymbol{m}_1$ is moved to $\boldsymbol{m}_1'$ on surface; obviously that $\boldsymbol{v}_2\boldsymbol{v}_3$ is also curved in $\triangle \boldsymbol{v}_2\boldsymbol{v}_3\boldsymbol{v}_4$ in the nature of things for same threshold, its midpoint is moved to $\boldsymbol{m}_1'$ in same way. Above operating is unable to leave "hole" or "crack" in resulting mesh.

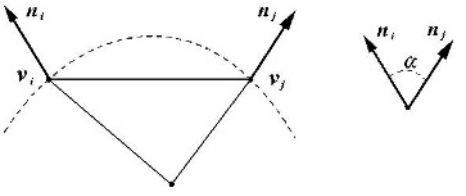


Fig. 5. Measure of surface curvature along an edge of a triangle

2.4 Moving Midpoint onto Surface

During the period of subdividing the coarse mesh, if an edge $\boldsymbol{v}_i\boldsymbol{v}_j$ is curved, it would be split into two pieces by its midpoint $\boldsymbol{m}$ and the point $\boldsymbol{m}$ would be moved onto the surface so that the subdivided mesh is a finer linear approximation of the surface. Therefore, we wish find a sample point $\boldsymbol{m}'$ on the surface such that $f(\boldsymbol{m}') = 0$ and it minimizes the distance, $||\boldsymbol{m} - \boldsymbol{m}'||$, between $\boldsymbol{m}$ and $\boldsymbol{m}'$. The solution of such an optimization problems can be calculated by various numerical methods. In this paper, gradient descent is employed, which is simple to implement and provides good accuracy control and takes fewer time.

The basic idea is as following: considering implicit function f of the surface as a potential function, which defines a potential field in $\mathbf{R}^3$ and has the lowest potential on the surface, furthermore, considering midpoint $\boldsymbol{m}$ a unit mass particle. When the particle is placed in the field, it will move to the lower energy places forced by the field. The gradient of $|f|$ can be used to generate such a force that drives particle to the implicit surfaces[14]. This physically-based method is modelled by following differential equation:

$$\frac{dx}{dt} + sign(f)\nabla f = 0$$

The above equation can be solved iteratively using an explicit Euler's time integration method. The position of a particle $\boldsymbol{m}_{t+\Delta t}$ at the next time-step $t + \Delta t$ can be calculated from its current position $\boldsymbol{m}$, at the current time t by :

$$\boldsymbol{m}_{t+\Delta t} = \boldsymbol{m}_t + \Delta t * sign(f(\boldsymbol{m}_t)) * \nabla f(\boldsymbol{m}_t)$$

The iteration computing is repeated until the particle is close enough to the surface, i.e. $f(\boldsymbol{m}_t)$ less than a predefined minimum ε.

If the time step is too big, the particle would oscillate from one side of the surface to the other, rather than converge onto the surface. In order to prevent such a case, the time step is reduced by half the particle crosses the surface every time.

In addition, owing to $f = 0$ is the minimal level set of f^2, the "force" defined as following formula is also "push" particle onto the surface f[18]:

$$F(f) = -C\nabla f^2 = -2Cf\nabla f$$

where C is a constant.

3 Experimental Results

Our algorithm is implemented by using C^{++} language on P4(1.2GHz, 256M) PC. The sample implicit functions used in the paper are listed in the Appendix at the end of the paper. Fig. 6 illustrates the adaptive comparing effect of our polygonization of implicit surface and Velho's. Table 1 is comparison of the time consumptions between the two algorithms.

In Fig. 6, the top and middle rows are curved parts of *Jack* surface and *Peanut* surface (see Appendix) respectively. The bottom row is surface, called *Blobby*,

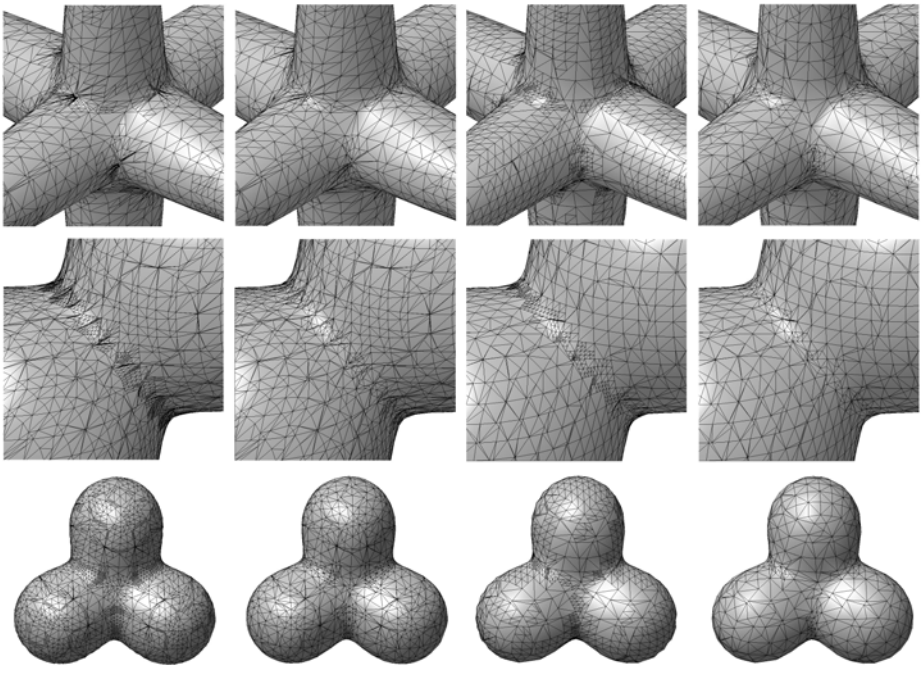


Fig. 6. Graphic effects of polygonizations using different initial meshes and different adaptive subdivision rules

Table 1. Time comparison for polygonizations of implicit surfaces between our algorithm and Velho's (time: ms)

Surface	Unit sphere		Peanut		Blobby		Jack		Topology	
	Triangles	Time	Triangles	Time	Triangles	Time	Triangles	Time	Triangles	Time
Velho's	3561	330	5416	237	11718	931	12323	911	68574	1452
Ours	1500	50	2460	58	6134	200	8338	291	58501	1132

blended from three spherical surfaces, on which the blended local has high curvature. The left column is generated by Velho's algorithm, there are many trivial triangles in meshes; the second left column is generated by our adaptive rules but refined based on Velho's uniform sampled coarse mesh; the second right column is generated by Velho's adapted method but based on coarse mesh built by Bloomethal's quick algorithm. The right column is result created by our method at same precision just used by its left columns. It has fewer triangles and eliminates "T"-edge automatically. In Table 1, both Velho's algorithm and ours use the same subdivision precision of 0.9 and the equal-size cubes during generating basic coarse mesh. Subdivision space used in Velho's algorithm is near-tight axis-parallel bounded-box that envelops corresponding surfaces. Time consuming in the table is confirmed by averaging time of algorithm running 1000 times continuously, in the process OpenGL is employed.

From Fig. 6 and Table 1 we can take a conclusion that our algorithm prevents redundant triangles without loss of approximation precision and eliminating the unexpected"T"-edges. In fact, the more the areas of curved parts on surface and the deeper the subdivision levels, the more efficient in reducing redundant triangles and economizing the time consumed in polygonization. Furthermore, executing time of the algorithm is affected greatly by the geometric character of implicit surface and the complexity of implicit function involved in the implicit function.

4 Conclusions

By dividing the polygonization of implicit surface into two stages of construction and adapted-subdivision, we present a quick adapted algorithm that can be used to render implicit surface in interactive environment. The algorithm not only eliminates completely unexpected "T"-edges in resulting mesh by using new defined subdivision rules, but also reduces efficiently the number of triangles in polygonization without reducing approximation precision of polygonal mesh. Moreover, it also optimizes the configuration of triangles in mesh.

References

1. Blinn, J.F.: A generalization of algebraic surface drawing. ACM Transactions on Graphics, Vol.1, No.3, (1982) 235–256
2. Buff, T.: Interval arithmetic and recursive subdivision for implicit surfaces and constructive solid geometry. In: ACM SIGGRAPH'92, (1992) 131–138
3. Pasako, A., Adzhiev, V., Sourin, A., *et al.*: Function representation in geometric modeling: concepts, implementation and applications. The Visual Computer, Vol.2, No.8, (1995) 429–446
4. Bloomenthal, J.: Polygonization of implicit surfaces. Computer Aided Geometric Design, Vol.5, No.4, (1988) 341–355
5. Roth, S.: Ray casting as a method for solid modeling. Computer Graphics and Image Processing, Vol.18, No.2, (1982) 109–144
6. Bloomenthal, J.: An implicit surface polygonizer. In: Heckbert P, ed. Graphics Gems IV, New York: Academic Press, (1994) 324–349
7. Velho, L.: Simple and efficient polygonization of implicit surfaces. Journal of Graphical Tools, Vol.1, No.2, (1996) 5–24
8. Allgower, E., Schmidt, P.: An algorithm for piecewise-linear approximation of an implicitly defined manifold. Journal of Numerical Analysis, Vol.22, No.2, (1985) 322–346
9. Wyvill, B., McPheeters, C., Wyvill, G.: Animating soft objects. The Visual Computer, Vol.2, No.4, (1986) 235–242
10. Bloomenthal, J.: Chapter 4: Surface tiling. In: Bloomenthal J, ed. Introduction to implicit surfaces. San Francisco: Morgan Kaufmann Publisher, Inc., (1997)
11. Wyvill, G., McPheeters, C., Wyvill, B.: Data structure for soft objects. The Visual Computer, Vol.2, No.4, (1986) 227–234
12. Yu, Z.S., Wu, Q.D., Li, Q.Y., *et al.*: Projection based polygonization of implicit surfaces. Journal of Computer-Aided Design and Computer Graphics, Vol.33, No.4, (2001), 333–337(in Chinese)
13. Turk, G.: Re-tiling polygonal surface. Computer Graphics, Vol.26, No.2, (1992) 55–64
14. Figueiredo, L.H., Gomes, J., Terzopoulos, D., *et al.* Physically-based methods for polygonization of implicit surfaces. In: Proceedings of Graphics Interfaces, Vancouver, (1992) 250–257
15. Velho, L.: Adaptive polygonization made simple. In: Brazilian Symposium on Computer Graphics and Image Processing , San Carlos, (1995) 111–118
16. Ning, P., Hesselink, L.: Adaptive isosurface generation in a distortion-rate framework, Society of photo-optical instrumentation engineers, San Jose, (1991) 11–21
17. Manfredo P. do Carmo.: Differential geometry of curves and surfaces. Pearson Education, Inc., (1976)
18. Ohtake, Y.: Mesh optimization and feature extraction. Ph.D Thesis, University of Aizu, (2002)
19. Frdric T., Philippe, M., *et al.*: Fast polygonization of implicit surfaces. In: Winter School of Computer Graphics, Plzen(Czech Republic), (2001) 283–290
20. Hart, J.C.: Ray tracing implicit surfaces. Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH'93, Anaheim, California. Course Notes No.25,"Modelling, Visualizing and Animating Implicit Surfaces", (1993) 1–15

Appendix: The Used Implicit Surfaces in the Paper

Unit-Sphere: $f(x,y,z) = 1 - x^2 - y^2 - z^2$

Peanet:[6] : $f(x,y,z) = e^{-((x-0.78)^2+(y-0.78)^2+3.25(z-0.78)^2)} + e^{-((x-0.23)^2+(y-0.23)^2+3.25(z-0.23)^2)} - 0.9$

Blobby: $f(x,y,z) = e^{-((x-1.0)^2+y^2+z^2)} + e^{-(x^2+(y-1.0)^2+z^2)} + e^{-(x^2+y^2+(z-1.0)^2)} - 0.9$

Jack:[10] : $f(x,y,z) = \left(\left(\frac{1}{\left((4x/3-4)^2+16y^2+16z^2/9\right)^4} + \frac{1}{\left((4x/3+4)^2+16y^2+16z^2/9\right)^4} + \frac{1}{\left(x^2/9+4y^2+4z^2\right)}\right)^4 + \left(\frac{1}{\left(y^2/9+4x^2+4z^2\right)}\right)^4 + \left(\frac{1}{\left(z^2/9+4x^2+4y^2\right)}\right)^4 + \frac{1}{\left((4y/3-4)^2+16x^2/9+16z^2/9\right)^4} + \frac{1}{\left((4y/3+4)^2+16x^2/9+16z^2/9\right)^4}\right)^{-\frac{1}{4}} - 1$

Topology: $f(x,y,z) = \left(1 - \left(\frac{x}{6}\right)^2 - \left(\frac{y}{3.5}\right)^2\right)\left((x-3.9)^2 + y^2 - 1.44\right) \times \left((x+3.9)^2 + y^2 - 1.44\right) - z^2$

A Framework of Web GIS Based Unified Public Health Information Visualization Platform

Xiaolin Lu

School of Information Technology,
Zhejiang University of Finance & Economics, Hangzhou 310012, China
luxiaolin@mail.hz.zj.cn

Abstract. The GIS plays a vital role in public health information visualization for public health information management, broadcasting, data management, statistical analysis, and decision supporting. This paper described the elementary requirement and the essential technology for public health information visualization and proposed a framework of the unified public health information visualization platform based on the Web GIS and visualization technology. The system framework adopted multi-tier system infrastructure that consist the sever tier and the front tier. In the server tier, the J2EE based architecture was adopted to construct a distrusted system infrastructure. In the front tier, the GIS map java applet is used to show the public health information with spatial graphical map, and the web based graphics figures such as curves, bars, maps and multi-dimensional visualization technology are used to visualize the public health information. The public health information contained the geo-referenced data, such as specific location, area code, latitude and longitude, street address, and geopolitical boundaries can be visualized with GIS distribution maps. The system infrastructure, functions, system integration, and some key technology were discussed in this paper. It would have the important practical value for constructing the visible public health information system.

1 Introduction

The GIS plays an important role in public health information visualization. In the public health crisis, such as the burst of infectious disease, the geo-reference information visualization will be essential to protect people from the nature disaster. Web GIS based visualization information can be used for warning, monitoring, processing and controlling the emergent event with its visible spatial information. With the development of Internet, the Web GIS becomes a desirable technology for building the public heath information system to prevent against the infectious disease and public health crisis. It is an essential tool for sharing the infectious diseases information, rescuing the infectious patients, isolating virus source area, and sending out the alarm to the public in short time [1-3].

GIS has been used for public heath area for long time. As early as in 1854, Dr. John Snow had used a map to track the original area where the cholera disease erupted in London in first time. From that time, map had been used in the infectious disease controlling and preventing, and it becomes an essential tool in protecting the public health environment [4]. The research subjects of applying the GIS to the public health have attracted

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 256–265, 2005.

much interest in recent years, especial after the eruption of SARS in 2003 [5]. The World Health Organization has used the GIS technology to forecast and analyze the spreading tendency of SARS with a GIS. ESRI Corporation has developed a WEB-GIS based information system to issuing the SARS distribution map in China and Hong Kong area. The SARS infected regions were shown with the visual technology in the map to report the SARS cases and distributed information to public. SarsNet, WHO and the Freach INSERM 444 research institute have developed an infectious disease supervisory system, which can be used to search the newest SARS surveillance statistical information from different nations and different time periods. The system has also provided the functions of searching the special disease information according to the geographical position and expressing the results with the visual technology. SarsNet also proposed SARS dissemination models, which can dynamically produce a new SARS distribution map to express developing information of SARS disease in next time [5,6,7,8].

In China, the Remote Sensing Institute of Chinese Academy of Science has conducted a knowledge innovation-engineering project named "Research on the SARS infectious situation information visible decision-making support environment" [9]. This project has applied the GIS technology to show spatial chart of the SARS infectious situation in China and Beijing area. The project has preliminarily investigated the relationship between the urban population transportation passenger flow and the SARS dissemination. They also have simulated the future SARS infectious situation. Their researches have provided the scientific foundation for understanding the SARS dissemination mechanism and provided a good tool for the SARS infectious situation controlling and the decision-making. In addition, the Science Geography Science and Resources Research Institute of Chinese Academy, the Beijing Supper Map GIS Company, and the China Infectious Disease Control Center also have conducted the research on Web-GIS based the public health application.

We started a research project on the SARS information management system for preventing and controlling the SARS disease from 2003. We have developed a Web GIS based interactive SARS information system, which can enable the public to participate in the prevention activities and reported the SARS information to the Health Cure Center. With that help of the system, the public and the medical worker could work together to explore and report the cases of illness, survey infectious disease area situation, isolate the infectious area in the shortest time.

In this paper we will report our research work on the elementary requirement and the essential technology in building the Web GIS based public health information visualization platform. The our research aims to apply the Web GIS based visualization technology for public health information broadcasting, data management, statistical analysis, and decision supporting. The research subjects of system infrastructure, functions, Web-GIS, web based visualization technology, system integration, and some key technological will be discussed in this paper.

2 WEB-GIS Based Public Health Visualization

Web-GIS is a technique that can set up geographical information system on Web. User can get the geographical information mutually by Web-GIS application through Inter-

net. It makes GIS function of system expand to web site by combining web with GIS technology. Various kinds of geographical space data, attribute data, picture can be obtained through web. With the rapid development of Web-GIS technology, GIS based applications can be developed with low costs, little maintain work. Web-GIS based systems have been popularizing in large scale [12,13].

While Web-GIS based system used the CGI based technology in early time, presently GIS java applet is mainstream technology as the front end in the browser to show the graphical map. In the CGI based Web-GIS system, web server transfers outside graphical user interface of GIS system. It expanded the network function of web server. The CGI is a bridge between graphical map interfaced and GIS application program in web server. Web server responds the request from web browser and transmits the GIS map and data information as picture, and sends the information back to the web browser.

Another technique to realize Web-GIS is to utilize ActiveX. The controlling parts and COM model technology are used to set up Web-GIS system. Microsoft Company provides the COM technology that can develop powerful Web-GIS system both in client/server and browser/server system structure. It also supports many kinds of development environments such as VB, VC, Delphi and PowerBuilder at the same time. With such technology, the GIS systems have the good flexibility and ability of expanding. To use the COM and ActiveX in the browser/server system structure, the COM

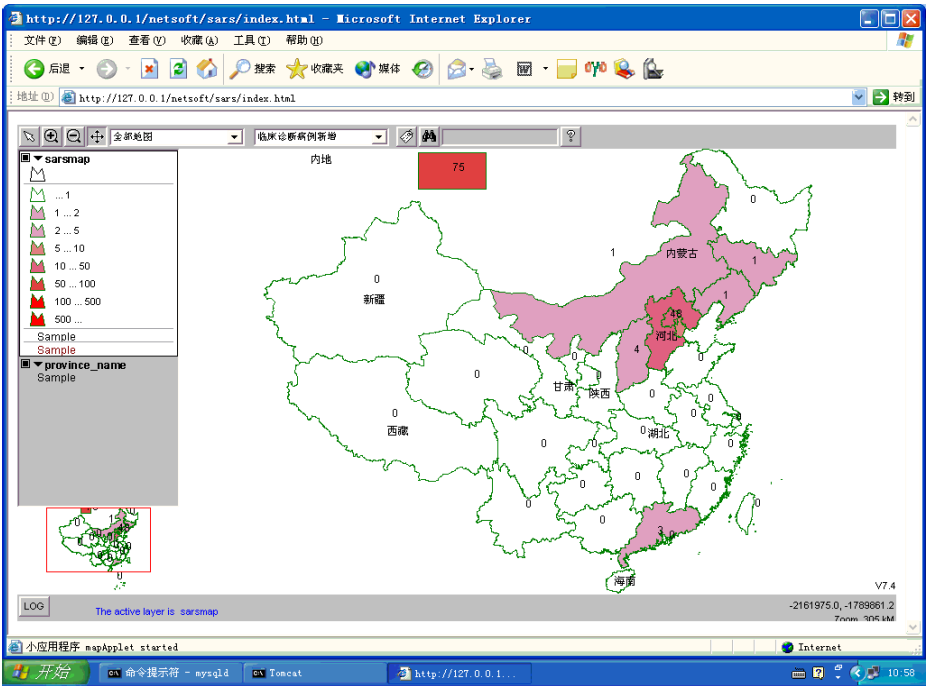


Fig. 1. Web-GIS based public health information visualization: distribution map with grade color to express quantities of infectious situation

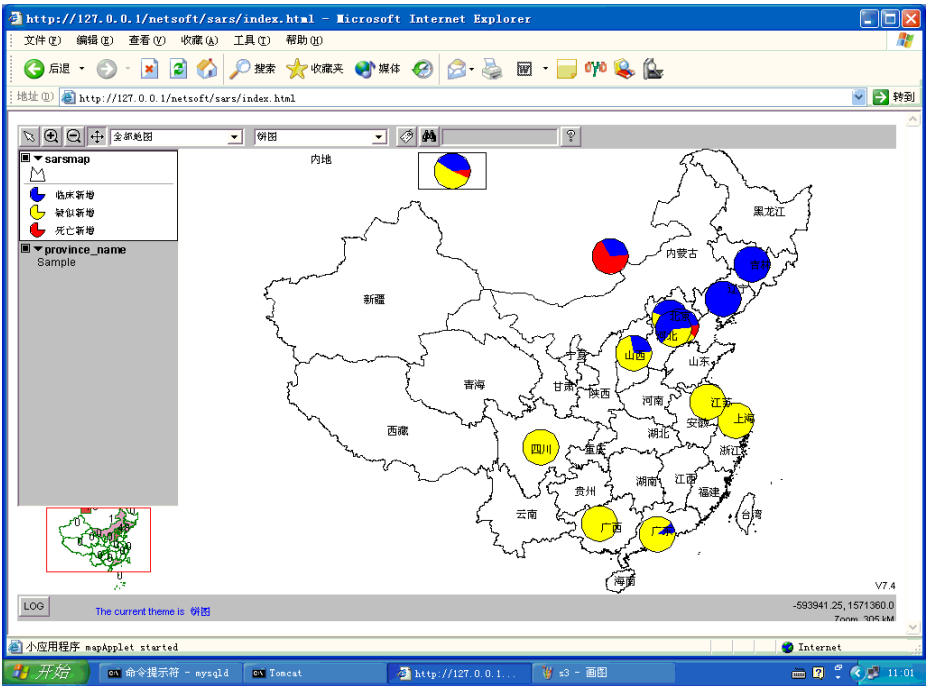


Fig. 2. Web-GIS based public health information visualization: Distributions map with the pie chart to express the public health information

and ActiveX serve as the plug-in in the browser. It is a good technique to set up Web-GIS system. There many Web-GIS application systems use this kind of technology. The plug-in gets the GIS data in server and display the graphical map in the browser. User can directly view and operate the graphical user interface by plug-in. It reduced consumedly the data transmitted in the network compared to the CGI based GIS system and solves the deliver bottleneck of graphical data in network. At the same time, plug-in also provides database operation function to view attribute data, search information, and operate the map by a graphical user interface.

The COM and ActiveX can only run in browser of IE produced by Microsoft Company and run on the Windows platform. It limits their application in other browser such as Netscape and other platform such as Unix operation system. Because java is platform independent, java GIS applet becomes the mainstream technique to establish the system of Web-GIS. The Java language is a language for network and object oriented. It seals object, dynamical inherit. With characteristic of object, overwhelming majority of data type appears in object form. Because Java adopts Virtual Machine technique, the target code has nothing to do with system platform. At the same time, it supports the distribute network computing. Because of these characteristics, java GIS applet becomes an ideal technique to realize the Web-GIS application. Presently GIS java applet is mainstream technology as the front end in the browser to show the graphical map.

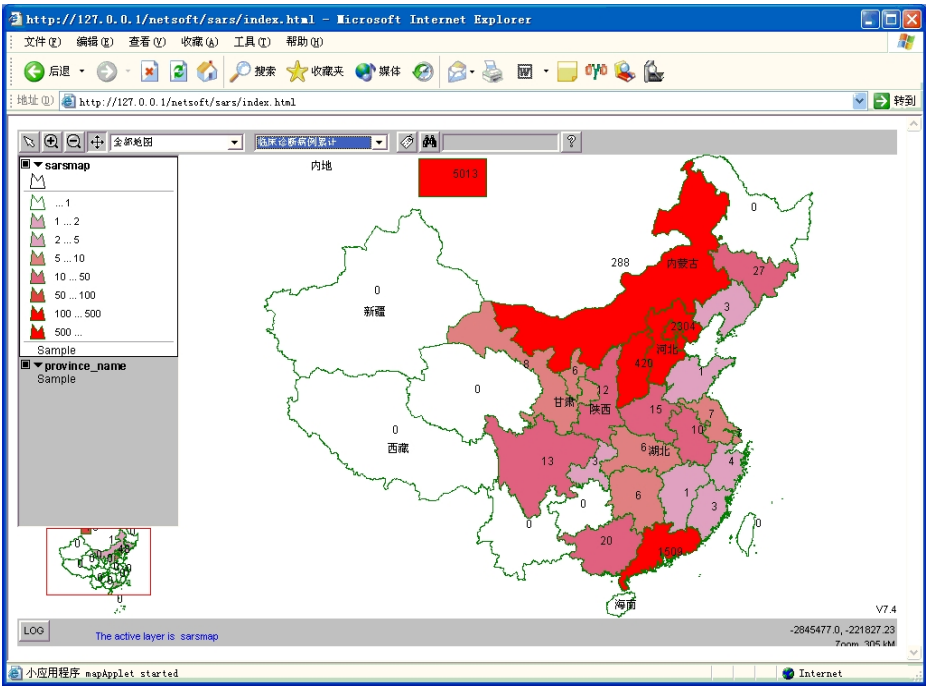


Fig. 3. Web-GIS based public health information visualization: a distribution map of accumulated infectious disease cases

2.1 The GIS Data and Public Health Information Visualization

By combine of the data of public health information with spatial data of GIS, distributing information of infectious disease situation can be visually displayed on electronic map of GIS. The infectious disease situation data such as the patient's number, the virus source areas can be shown at corresponding geographical position. The electronic map with GIS can realize the flowing functions [14,15]:

1) Necessary theme layers and marks;
2) Zoom in, zoom out, pan and selection functions on map;
3) Search functions such as the graphic objects search, the key words search, etc.;
4) Data inputting functions on map, such as the functions to input GIS relevant information in the corresponding area on the map.

Basic GIS and infectious disease situation data needed to be collected for setting up the distribution maps of newly increased case in real time. By connect with the GIS database, we can browse and search the basic public health information, infectious situation area, hospitals distribution in the area.

2.2 The Visualization of Infectious Disease Situation with Distribution Maps

There are several kinds of distribution maps: the distribution map of newly increased case and the distribution map of the accumulative total cases. Using the grade color

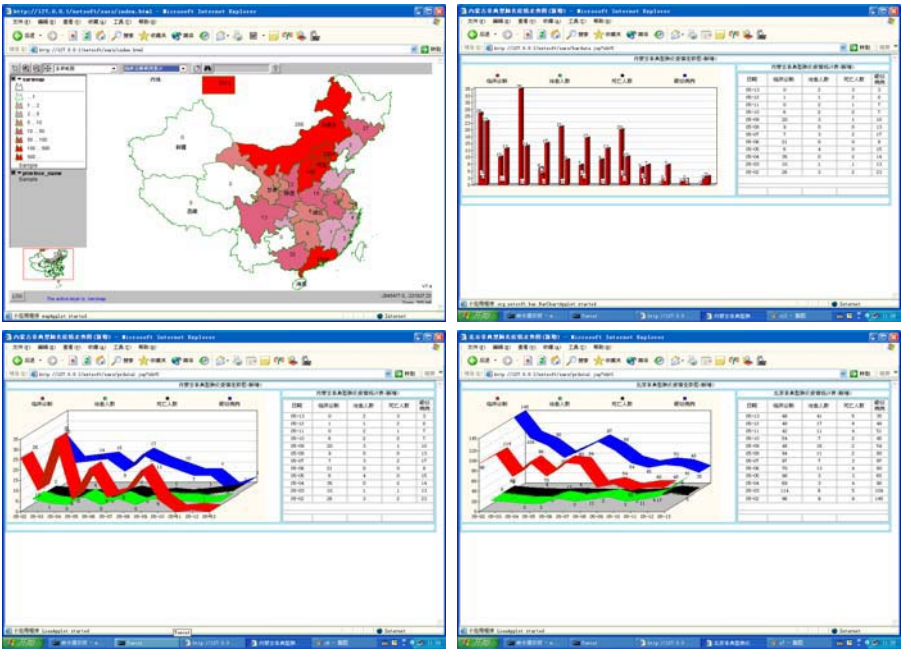


Fig. 4. The pictures show the visualization the relationship of the public health historical data and virus-developing tendency

can show the distribution situation information. The bar and pie pictures can be used to express the dimensional quantities of infectious situation.

The figure 1 to 3 shows the distribution map for public health information visualization with geo-reference data. The figure 1 shows a distribution map of newly increased infectious disease cases. The different color express different numbers of the cases, the legend figures are shown in the left. The figure 2 shows pie chart distribution map of percentage in newly increased cases, suspect cases, and death cases. The figure 3 shows a distribution map of accumulated infectious disease cases.

2.3 The Visualization of the Historical Data and Infectious Spreading Tendency

We can click on the map to inquire about the infectious situation of this area about the information of historical data, developing tendency, which can display visually three-dimensional line chart, bar to express the data of infectious situation.

The figure 4 shows the visualization of the historical data, virus-developing tendency.

1. The left-top picture is a distribution map.
2. The right top picture is a bar chart to show the historical data.
3. The bottom pictures are three-dimensional line charts to express the tendencies of infectious situation.

3 A Framework for the WEB GIS Based Public Health Information Visualization Platform

In order to setup up a visual work environment for public and hygiene department to research, statistical analyze and broadcast the public health information between the spatial graphical data and the infectious disease data, we proposed a framework for the Web GIS based public health information visualization platform.

Web-GIS applet can be used to get and show the geographical information for public in web site. Also, public can use the GIS java applet in browser to reporting the most recent public health information. After the message received from the public, there will be a serious of immediate actions, such as confirming the information, notifying relevant department, rescuing the patient, isolating the infectious disease area and broadcasting the information. The all processes will get many departments involved, such as hospital rescuer center, police station and relevant government department. The unified visualization information platform is a powerful assistant for achieving of communication, interactive operation, effective data processing, and GIS data visualization.

Because there is many computers with different kinds of operating system will work in coordination in the system, the distributed, platform independent system architecture needed. To enable application system be accessed by lots of different computers in network, we choice the J2EE and EJB technology that offer framework of service system to construct enterprise system structure.

The middleware of CSCW, GIS, and public health information EJB offers distributed transaction processing. Many host computers can join to offer the many services. Compared with other distributed technology, such as CORBA technology, the system structure of EJB has hidden the lower detail, such as distributed application, the events management, the target management, multi-thread management and connoting pool management etc. In addition, J2EE technology offers many kinds of different middleware to be applied to business logic. The data are stored and managed with EJB. Distributed computing enable users operate in any time, any place, and obtain business logic and data processing in remote server. The distributed systems enable the databases and services in the same or different computers. The databases consist of CSCW database, public health information database, and interoperable GIS information system database. The database uses JBBC to communicate with EJB (back-end server). The system front-end adopts JSP/Servlet. JSP and Servlet use the back-end work to provide service. The front end adopts a java GIS applet to communicate with JSP/Servlet and back end EJB in Browser.

System architecture is platform independent, multi-layers and distributed structure. It can combine GIS database, CSCW database, and many host computers together and share the system resource and data.

Web-GIS based public health information visualization platform should be considered and realized the following basic functions.

Information Sharing: The databases of GIS, public health data are the information center of the system. All the cooperating work and information sharing is dependent on

them. For example, when the public reports the infectious situation through system, the information will be stored in the database and be shared by others.

Human Computer Interaction Based on WEB-GIS: The Web-GIS is the basis graphical user interface of man-machine interface in the system. Based on WEBGIS, interoperable information visualization platform obtains the geographical position at the same time. The friendly interface and interactive system is convenient for information management.

4 Discussions and Further Work

We developed a framework of the unified visible public health information visualization platform based on the Web GIS technology. There are two components for the platform: the platform server end and the client end. In the server end, we will develop GIS server based on the J2EE framework. In the client end, we will investigate a visible harmonious cooperation user interface to realize the public health information visualization and cooperation.

Although we completed a framework the public health information visualization based on the WEB-GIS technology, there are still many technical problems to wait for the research:

(1) The different GIS spatial data frequently existed in different format document. The special GIS data format transform software is needs to translate the GIS data format (for example ARC/INFO, ArcView, MapInfo, ERDAS, and Atlas and so on). The system lacks the exchange standing for information and data and the resources cannot be shared by other system.
(2) The spatial data is usually huge (several hundred MB to several GB). The user must have enough storage and the bandwidth to transmission data. While in very many situations, user possibly only needs a small area of the whole map or the partial GIS data, not the complete spatial data. As a result of the network and the bandwidth limit, the system performance becomes very slow.

These problems are the essential technical problems, which need to be solved in building a Web-GIS, based the public health information visualization platform. The further research work will focus specifically on the following aspects:

(1) The essential technology that realizes the interactive and convenience man-machine user interface will be investigated to break the limitation of the WEB-GIS system.
(2) The visualization technology for the public health information data, such as 3D and the virtual reality technology (VR) would be considered to applying the simulation of the real public health environment. It will enhances the cognition effect in the interactive observation and the analysis public health environment
(3) The essential technology of data compression and data optimization are the bottleneck problems on the Internet transmission when the spatial data transferred from the far-end server to the browser in the Web GIS system.
(4) The system data specification of the public health information based on OGIS (OGIS, Open Geo-data Interoperable Specification) and SVG (Scalable Vector Graph-

ics) should be investigated to enable the geography data and the public health attribute data exchangeable between the different geographic information system software and system.

5 Conclusions

The public health information management is a socialized work that the public would be very important to participate in the activity for preventing and controlling the acute bursting infectious disease. In the process of preventing and controlling the acute bursting infectious disease, an Web-GIS based visible and interactive public health information operation environment will be helpful for the public to participate in prevention disease on their own initiative. The public and the hygiene department could be cooperated together to prevent and control the acute bursting infectious disease, rescue the patient, and isolate the infectious disease source area.

This paper proposed a framework of the unified public heath information visualization platform based on Web-GIS and virtualization technology. The Java GIS applet technology combined with web based visualization technology has been applied to the public heath information visualization of the geographical distribution and the historical developing tendency. It can be applied to the visualization of public health information management, such as information broadcasting, data management, statistical analysis, and decision supporting.

The Web-GIS based visible and interactive public health information visualization environment will be helpful to track infectious disease-spreading tendency, to build up the immunity isolation mechanism for the infectious disease, establish the best transportation line for the personnel and the equipment supply in the infectious region, dynamically issue the medical service health device information on the Internet. It may provide the tools for the visible infectious disease trend analysis, the visible medical service demand analysis, the hospital and the outpatient clinic location, the visible region analysis, and the resources and equipment management. It will have the important practical value and to protect the life and health, play a vital role in the visible public health information system construction.

References

1. Kamel Boulos, M.N., Roudsari, A.V., Carson, E.R.: Health Geomantic: An Enabling Suite of Technologies in Health and Healthcare (Mythological Review). Biomed Inform, Vol. 34(3) (2001) 195-219
2. Mennecke, B.E., Crossland, M.D.: Geographic Information Systems: Applications and Research Opportunities for Information Systems Researchers. In: Proceedings of the 29th Hawaii International Conference on System Sciences (HICSS), Maui Hawaii. Vol. 3 (1996) 537-546
3. Johnson, C.P., Johnson, J.: GIS: A Tool for Monitoring and Management of Infectious. In Proceedings of Map India 2001, the 4th Annual International Conference and Exhibition, 7–9 February, New Delhi, India (2001)

4. Kistemann, T.: GIS for communicable disease control: perspectives and pitfalls. In Proceedings of the First European Conference for Geographic Information Sciences in Public Health, 19–20 September, Sheffield, UK (2001)
5. World Health Organisation: Communicable Disease Surveillance and Response: Severe acute respiratory syndrome (SARS): Status of the outbreak and lessons for the immediate future (2003)
6. Keola, S., Tokunaga, M., Tripathi, N.K., Wisa, W.: Spatial Surveillance of Epidemi-ological Disease: A case study in Ayutthaya Province. Thailand. GIS @ devel-opment (2002)
7. Kamel Boulos, M.N.: Location-based health information services: a new paradigm in personalized information delivery. Int J. Health Geogr. Vol. 2:2 (2003).
8. Midtb T.: Visualization of the temporal dimension in multimedia presentations of spatial phenomena. In Bjrke, J.T., Tveite, H. (eds.): Proceedings of ScanGIS'2001 – The 8th Scandinavian Research Conference on Geographical Information Science, 25–27 June, Norway (2001) 213-224
9. Science Remote Sensing Application Research Institute of Chinese Academy: Research on SARS epidemic situation information visible decision-making support environment, Research Bulletin of the SARS epidemic situation analyzes and forecasts, No. 1, May 2 (2003)
10. Isaac Brewer, Alan, M., MacEachren, Hadi Abdo, Jack Gundrum, George Otto: Collaborative Geographic Visualization: Enabling Shared Understanding of Environmental Processes. IEEE Symposium on Information Visualization, Salt Lake City Utah (2000)137-144
11. Shanzhen, Y., Lizhu, Z.., Chunxiao, X., Qilun, L., Yong, Z.: Semantic and Interoperable WebGIS. In: Proceeding of the International Conference on Web Information Systems Engineering (WISE'01), Vol.2, Kyoto Japan (2001) 42-48
12. Jiang, J., Jun, C.: A GIS—based computer supported collaborative work CSCW system for urban planning and land management. Phonograms metric Engi-neering & Remote Sensing, Vol. 68(4) (2002)353-359
13. Yingwei, L., Xiaolin, W., Zhuoqun, X.: Design of a Framework for Multi-User/Application Oriented WebGIS Services. In Proceeding of the 2001 International Conference on Computer Networks and Mobile Computing (ICCNMC'01), Beijing China (2001)151-157
14. Shashi Shekhar, Sanjay Chawla, Siva Ravada, Andrew Fetterer, Xuan Liu, Chang-tien Lu: Spatial Databases-Accomplishments and Research Needs. IEEE Transactions on Knowledge and Data Engineering, Vol. 11(1) (1999)45-55
15. Candan, K.S., Rangan, P.V., Subrahmanian, V.S.: Collaborative multimedia systems: synthesis of media objects. IEEE Transactions on Knowledge and Data Engineering, Vol. 10(3) (1998) 433-457

An Improved Colored-Marker Based Registration Method for AR Applications

Xiaowei Li, Yue Liu*, Yongtian Wang,
Dayuan Yan, Dongdong Weng, and Tao Yang

School of Information Science and Technology,
Beijing Institute of Technology, Beijing, 100081,
Tel :+8610-6891 2565 ext.11,12,13 or 17
{nemesis, liuyue, wyt, ydyuan, crgj, yorktim}@bit.edu.cn

Abstract. Registration is crucial in an Augmented Reality (AR) system for it determines the performance of alignment between virtual objects and real scene. Colored-makers with known world coordinates are usually put in the target scene beforehand to help get a real-time, precise registration because they can provide explicit 3D/2D correspondences and four such correspondences can produce adequate and accurate equations of the pose matrix if the camera's intrinsic matrix has already been calibrated, and then registration can be achieved by solving these equations. However, usually only limited number of (e.g. two or three) markers out of four can be captured and this will make the colored-marker based method fail. In order to overcome such shortcomings an improved colored-marker based registration method is proposed in this paper which works when the target scene is a plane. The proposed method integrates both 3D/2D and 2D/2D information by updating the cost function used in the optimization step of RANSAC, and thus combines the virtues of homography based method. Experimental result shows that the proposed method can provide acceptable pose estimation and its potential to be applied in actual AR systems.

1 Introduction

AR is a new technology in which computer-generated virtual objects can be merged into the image sequences of real scene. Rather than immersing the user into an absolutely virtual world as a Virtual Reality (VR) system does, an AR system allows its user to view the real environment with the virtual objects imposed on. An AR system can be widely applied in many fields such as medical care, entertainment, manufacturing, outdoor navigation and cartoon film etc. [1], [2].

One of the key issues when building an AR system is accurate registration, whose solution is used for the alignment of virtual objects and real scene. Compared with the sensor-based registration approach, vision-based method is much

* Corresponding Author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 266–273, 2005.

cheaper and more reliable, though its computation cost is greater, which causes lower frame rate. In a vision-based method, colored markers are usually placed in the target scene beforehand to help get accurate registration, which can provide stable, explicit 3D/2D correspondences, and four coplanar correspondences could provide adequate equations of the camera's pose matrix after the camera's intrinsic matrix has already been obtained during the process of calibration, then precise registration result can be obtained by solving these equations as described in [3], [4].

Although four coplanar marker based method is fast, accurate and widely adopted in many real-time AR applications [5], [6], it does have some flaws. When the number of captured markers is below four, registration will fail because for every frame, all four markers have to be captured as an input to the registration process. Moreover, applications with only four markers usually have very narrow view fields and are sensitive to the occlusions on the markers, which sometimes results in high-frequency jitter. Using more markers seems a feasible solution but in many situations it's not an easy thing, because defining each marker's world coordinate is also a hard work

When the target scene is a plane, relationships between two images shot from different angles and between world plane and final image we see are both known as homographies, which are much easier than those in general scenes. With these homographies, there has already been registration approach to recover camera matrix for each frame. [7], [8]. By tracking the world plane, this method can get a larger view field. Because the estimation of homography with RANSAC is robust and reliable, it can accommodate situations with occlusions. But it will cause accumulative errors which leads to drifting, so it can not be applied for long sequences.

To compensate for the shortcomings of the above-mentioned methods, a hybrid registration method based on both homographies and partial 3D/2D information that comes from less than four colored markers placed on a world plane is proposed in this paper. Although the conventional colored marker based method will fail in such situation, the proposed registration method can be performed and experiments show that it can still achieve reliable result, which shows the robustness of the proposed method and its potential for future AR applications.

2 Coplanar Colored-Marker Based Registration Method

If the camera used to shoot the image sequence has already been calibrated, the registration can be achieved with four coplanar markers whose world coordinates have been predefined. A brief description of this method will be given.

2.1 3D/2D Homography

A 3D/2D correspondence (m, M) includes a 3D point M, which is always represented in homogenous form $(X, Y, Z, 1)^T$, and a 2D pixel point m, which is

also represented in homogenous form $(x, y, 1)^T$. (m, M) is related by the 3×4 projective matrix P_i [4]:

$$m = \lambda_i P_i M, \quad P_i = K\,[R_i \mid t_i] \tag{1}$$

where R_i is a 3×3 rotation matrix, t_i is the translation vector of the camera and λ_i is the homogenous scale factor which is dependent of $P_i M$. The camera's 3×3 intrinsic matrix K can be written as:

$$K = \begin{bmatrix} f & s & u \\ 0 & af & v \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where f is focal length, (u, v) is principle point, a is aspect ratio and s is skew.

When the target scene is a plane and the world coordinate frame X-Y plane is defined on the world plane, 3D point M on this plane will be in the form $(X, Y, 0, 1)^T$, then equation (1) can be written as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = m = \lambda_i P_i M = \lambda_i K \begin{bmatrix} r_{i1}, r_{i2}, r_{i3}, t_i \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \lambda_i K \begin{bmatrix} r_{i1}, r_{i2}, t_i \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{3}$$

where r_{ij} is the $\boldsymbol{j}$th column of R_i [7], [8]. Thus, (m, M) is related by a 3×3 matrix H_w^i [3]:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda_i H_w^i \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad H_w^i = K \begin{bmatrix} r_{i1}, r_{i2}, t_i \end{bmatrix} \tag{4}$$

H_w^i is called 3D/2D homography, which is a homogeneous matrix with 8 independent elements, and each 3D/2D correspondence (m, M) can provide two independent equations of H_w^i. Thus, four coplanar markers can be used to solve an accurate H_w^i.

2.2 Recovering Projective Matrix from 3D/2D Homography with Known K

For the ith image, if H_w^i has been precisely obtained by the four coplanar markers and the camera's intrinsic matrix K has already been known and fixed, the projective matrix P_i can be easily recovered from equations (1) and (4) because R is a unit orthogonal matrix. [7].

$$K^{-1} H_w^i = \begin{bmatrix} r_{i1}, r_{i2}, t_i \end{bmatrix}, \quad P_i = K \begin{bmatrix} r_{i1}, r_{i2}, r_{i1} \times r_{i2}, t_i, \end{bmatrix} \tag{5}$$

When all four markers have been captured, an accurate registration result can be achieved. Although actually markers can be tracked and extracted by using optical flow and color information etc. [5], it's not an easy job to locate all four color markers which are crucial for precise registration.

3 Homograhpy Based Registration Method

3.1 2D/2D Homography

The relationship between consecutive images i and $i+1$ of the same target scene is usually described as fundamental matrix F. With F, a point in image i is mapped to a straight line (called epipolar line) in image $i+1$, and vice versa. But when the target scene is a world plane, the relationship is a 2D/2D homography H_i^{i+1}. With H_i^{i+1}, a 2D/2D correspondence (m, m') is related by:

$$m' = H_i^{i+1} m, \quad m = {H_i^{i+1}}^{-1} m' \tag{6}$$

where H_i^{i+1} is introduced by the world plane which is paid attention to in [4].

In real images of a scene plane, all H_i^{i+1} are obtained using a robust and fast estimation method, e.g. RANSAC method [4].

3.2 Homography Based Method

A 2D/2D correspondence (m, m') is the image of the same 3D world point M in two images, thus equation (7) can be written as by combining equations (4) and (6):

$$H_w^{i+1} = H_i^{i+1} H_w^i \tag{7}$$

If the target scene is a plane and the first image 3D/2D homographyH_w^0has been accurately gained in advance, equation (7) could be written as [7],[8]:

$$H_w^{i+1} = H_i^{i+1} H_{i-1}^i ... H_0^1 H_w^0 \tag{8}$$

Thus, H_w^{i+1} of every image in such sequence could be computed sequentially. After getting H_w^{i+1}, registration can be accomplished with equation (5) after the camera has already been calibrated just as the colored-marker based method does.

However, it's obvious that there is an accumulative error in H_w^{i+1} if H_i^{i+1} is not robustly and reliably estimated, and this will result in the drifting of the latter images of relatively longer sequences. In this homography based method, $H_i^{i+1}$is accurately obtained with RANSAC algorithm.

3.3 Robust Estimation of H_i^{i+1} Between Two Images

RANSAC stands for "**RAN**dom **SA**mple **C**onsensus", which is a robust estimation method. The process to estimate H_i^{i+1} with it can be written as follows [4]:

(1) Compute interest points in each image i and $i+1$.
(2) Make putative correspondences.
(3) RANSAC estimation:
- (a) Select a random sample of four correspondences and compute an H_i^{i+1}.
- (b) Calculate the reprojection errors of each putative correspondence $\{d_j\}$.
- (c) Compute the number of inliers consistent with H_i^{i+1} by comparing the errors with certain threshold t by $d_j < t$.

Choose H_i^{i+1} with the largest number of inliers.

(4) Optimization:
Re-estimate H_i^{i+1} from all correspondences classified as inliers by minimizing certain cost functions.

(5) (Optional) Guided Matching:
Further interest point correspondences could be found using the optimized H_i^{i+1} and these newly found correspondences can be determined as inliers directly.

Steps (4) and (5) can be iterated until the number of correspondences is stable.

4 The Improved Marker Based Method

As mentioned before, the typical sub-sequence that causes abrupt jitter in normal colored-marker based method is that, in the first image, all the four markers are captured, and in the subsequent images usually only two or three markers can be tracked As a result, the conventional four coplanar marker based method will fail in the subsequent images.

To compensate for such shortcomings, an improved colored-marker based registration scheme is proposed, in which under normal situations conventional registration approach is performed, and under unsuccessful situations with markers lost, homography based registration approach is performed. Besides, the remaining accurate 3D/2D information from the remaining markers has also been integrated in the process of H_i^{i+1} estimation, and plays a more important part in the makeup of residuals for RANSAC method.

4.1 Updated Cost Function in 2D/2D Homography Estimation with RANSAC

When applying RANSAC algorithm to compute 2D/2D homography between consecutive images, certain cost function of residuals is used for the final iterative refinement on the inlier matches. Usually, these residuals are 2D/2D reprojection error caused by the correspondences$\left(m_j, m'_j\right)$, and this cost function's normalized form is:

$$c_1 = \frac{1}{2k}\sum_{j=1}^{k} d(m_j, H_i^{i+1^{-1}} m'_j)^2 + d(m'_j, H_i^{i+1} m_j)^2 \tag{9}$$

where k refers to total number of $\left(m_j, m'_j\right)$ used to compute H_i^{i+1}, and $d(X_1, X_2)$ is:

$$d(\begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix})^2 = (\frac{x_1}{w_1} - \frac{x_2}{w_2})^2 + (\frac{y_1}{w_1} - \frac{y_2}{w_2})^2 \tag{10}$$

In the typical sub-sequence, there are still remained markers in the images and with these markers, extra 3D/2D information can be exploited. If H_w^i estimated is assumed reliable, when less than four markers are captured in image $i+1$,

considering (4) and (7), a normalized cost function from 3D/2D reprojection error is:

$$c_2 = \frac{1}{s}\sum_{j=1}^{s} d(n_j, H_w^{i+1}N_j)^2, \quad H_w^{i+1} = H_i^{i+1}H_w^i \tag{11}$$

where (n_j, N_j) is the accurate 3D/2D correspondences from the remained markers, and s refers to the total number of remained markers, $s < 4$, usually 2 or 3.

Generally the cost function C used to optimize H_w^i in the step (4) of RANSAC contains only c_1, which mean that it only uses the 2D/2D information.[4],[7]. But for the above-mentioned situation, C should of course integrate both c_1 and c_2, and further, c_2 should play a much more important role, because (n_j, N_j) which are obtained from the remained markers are very reliable. Thus the final improved cost function is:

$$C = c_1 + \lambda c_2, \quad \lambda = 3 \tag{12}$$

4.2 The Improved Registration Scheme

For a calibrated image sequence, registration for image i will be performed as follows:

(1) Track and record markers in image i, with a KLT feature tracker.
(2) Count the number of tracked markers,denote it as N.
(3) If $N = 4$, perform normal marker based method. Firstly, compute current image's 3D/2D homography H_w^i. Secondly, deduce registration result from H_w^i.
(Record this H_w^i for possible further use in image $i + 1$)
(4) If $N < 4$, firstly robustly estimate H_{i-1}^i using RANSAC algorithm, with the remained markers information integrated to update optimization process for higher accuracy. Secondly, get current image's 3D/2D homography H_w^i by $H_w^i = H_{i-1}^i H_w^{i-1}$. Thirdly, deduce registration result from H_w^i.
(Record this H_w^i for possible further use in image $i + 1$)
(5) Proceed to image $i + 1$.

5 Experimental Result

A camera is used to shoot a 320×240 sequence of a plane scene and the camera has been calibrated with a chessboard using Zhang's method [3]. Two consecutive images of the sequence are shown below, i.e. Fig. 1(a) image 0 and Fig. 1(b) image1, one with all four markers captured and the other only three.

Surely colored-marker based registration would succeed on image 0. The registration for image 1 can be achieved by homography based method and our improved marker based method separately. Harris corner detection is used to extract interested points. In order to make putative correspondences, Harris points in image 0 and 1 are matched by maximizing the cross-correlation in a 7×7 window, within a 30 pixel disparity. Totally there are 75 putative correspondences as input into RANSAC algorithm.

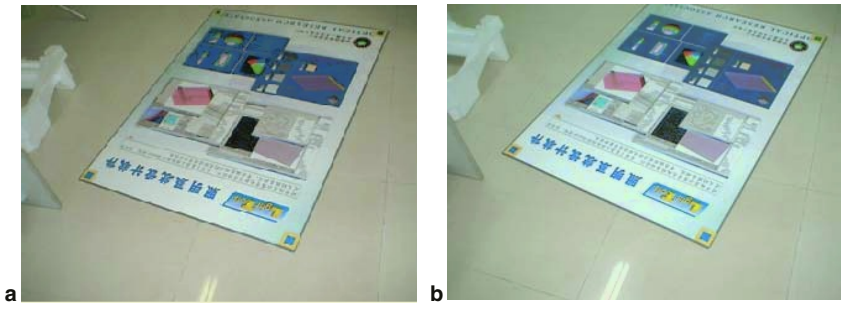


Fig. 1. Two consecutive images of a sequence: (**a**) Image 0 with all four markers captured. (**b**)Image 1 with only three markers captured

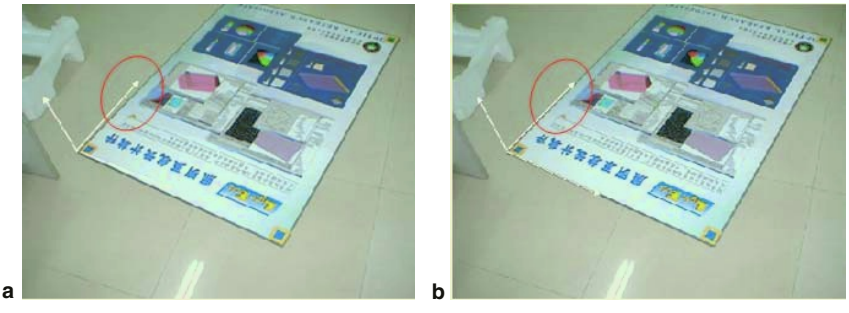


Fig. 2. (**a**) Registration Using Homography Based Method. (**b**)Registration Using the Improved Colored Marker Based Method

In the optimization process, the minimization of equation (12) is achieved using the Levenberg-Marquard iterative algorithm.

Fig. 2 shows the experimental result. It can be easily found that our method can give a relatively better result. In Fig. 2(a), the optimization of H_0^1 is not very reliable, because the number of final inliers we could achieve is small which is not sufficient for the calculation of accurate estimation. However, in Fig. 2(b), by integrating with 3D/2D information, our improved method can give a much more reliable result, because the error are accumulative, our improvement does make sense to reduce drifting.

6 Conclusion

To conclude, we have presented a registration scheme which combines the virtues of both the colored-maker based method and the homography based method, by developing a hybrid cost function to update the optimization step of RANSAC, in which 3D/2D information is also integrated for higher accuracy. Experiments show that the final registration result can be much improved.

The advantages of the proposed method are: Firstly, when the number of captured markers is less than four, registration can still be robustly and reliably

estimated. Secondly, although the homography based method can achieve registration without any markers, it does have an accumulative error causing drifting problem, but when integrated with partial 3D/2D information, its registration result can be greatly improved, and this will reduce drifting.

However, the proposed method still has the following shortcomings: Firstly, the environments the proposed method applied in must have a dominant plane, which can be easily satisfied in many applications, though sometimes it's indeed a hard constraint. Secondly, as the putative correspondences input into RANSAC are from the whole images and thus, the computation is too heavy to perform in real-time on an modest PC. Although our proposed solution can compensate for some shortcomings of colored-marker based method, in fact it loses its important real time quality. On the other hand, since no batch adjusting is used and the proposed method is actually sequential-processing, it may be implemented in real-time together with the rapid development of computer technology.

Acknowledgements

This project is supported by the National Basic Research Program of China (National 973 Project, Grant No. 2002CB312104) and National Natural Science Foundation of China (Grant No. 60025513).

References

1. R.T.Azuma et al.: Recent Advances in Augmented Reality. Computers Graphics and Graphics, Vol. 21, No. 6. Nov./Dec. (2001)
2. Princeton Video Image: Product showcase. `http://www.pvi-inc.com/showcase` (2000)
3. Zhang Z.: Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. Proceedings of 7th ICCV. (1999) 666–673.
4. Richard Hartley and Andrew Zisserman: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press. (2003)
5. Jing Chen: Research in the Spatial Registration Technology of A Video AR System. Ph.D Thesis. Beijing Institute of Technology (2002)
6. `www.hitl.washington.edu/research/shared_space/download/`
7. G. Simon, A.W. Fitzgibbon, and A. Zisserman: Markerless Tracking using Planar Structures in the Scene. Proceedings of ISAR2000, 5-6 Oct. (2000) 120–128
8. Simon J.D. Prince, Ke Xu, and Adrian David Cheok: Augmented Reality Camera Tracking with Homographies. Computer Graphics and Applications, Vol. 22, No. 6, Nov./Dec. (2002) 39–45.

Non-photorealistic Tour into Panorama

Yang Zhao, Ya-Ping Zhang, and Dan Xu

Department of Computer Science and Engineering,
Yunnan University, Kunming 650091, China
danxu@vip.sina.com
bootcool@163.com

Abstract. In this paper, we describe a system NP-TIP (Non-Photorealistic Tour Into Panorama). It provides a simple non-photorealistic scene model in which users can freely walk through and obtain enjoyable and real-time artistic experience. In NP-TIP, firstly we design a new algorithm for fast converting a photo or image to a synthesized painting following the painting style of an example image. By treating painting styles as sample textures, we reduce the problem of learning an example painting to that of texture synthesis, which improves the speed of non-photorealistic rendering. Secondly, we propose a new modeling scheme for TIP based on the cubic panorama, which not only overcomes the disadvantage of fixed viewpoint in browsing panorama, but also models easily and computes simply. According to users' selection from example paintings of different artistic style, NP-TIP can provide stylized interactive and real-time panoramic walkthrough of scenes.

1 Introduction

In recent years image-based modeling and rendering (IBMR) has been increasingly interested by people. Unlike traditional 3D computer graphics in which 3D geometry of the scene is known, image-based rendering techniques render novel views directly from input images. Panoramic image is synthesized by using several images took from the same viewpoint with certain overlapping areas. Then the viewable portion of the panoramic image is re-projected to a view plane to realize navigation of a real scene. Compared to traditional geometry based modeling method, it can be not only realized easily, but also get better effects for natural scenery which is very difficult to be represented by geometry model.However, panoramic representations can only simulate camera rotation under fixed viewpoint and provide simple zooming effects. It is not indeed walkthrough, which is very important visual effect for navigation. In 1997,Y. Horry [2] and H. Kang [3] proposed a navigation method of single image respectively, called TIP (Tour in to the Picture). According to the comprehension and imagination to reference image, this method reconstructs a simple 3D model with respect to reference image and generates realistic walk-through images by navigating the model. TIP not only overcomes the problem of fixed viewpoint, but also has the advantages of modeling easily and computing simply. It can be applied to different kinds of real-time rendering of complex virtual scenes.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 274–283, 2005.

While current research in virtual environments has traditionally striven for photorealism, for many applications there are advantages to non-photorealistic rendering (NPR). Firstly, artistic expression can often convey a specific mood difficult to imbue in a photorealistic scene. Secondly, through abstraction and careful elision of detail, NPR imagery can focus the viewer's attention on important information while downplaying extraneous or unimportant features. Finally, a NPR look is often more engaging than the traditional, photorealism computer graphics rendering [1].

With the development of NPR, it not only processes 2D images, but also extends to represent 3D scenes and objects. More and more researchers endeavor to depict 3D model by NPR means. Robert D. Kalnins etc. have introduced a very classic interactive NPR system for 3D model [4]. An outstanding feature of this system is that the designer can choose varieties of brush styles and paper textures to directly paint on the 3D model as his pleasure. The shortage of the system is that the painting speed is limited by the amounts of polygon surfaces on the 3D model. The similar systems can be referred to references [5] [6].

Allison W. Klein etc. have also described a system for non-photorealistic rendering (NPR) of virtual environments [1]. By adopting the image-based rendering methods at the off-line preprocessing step, the system achieves the good interactive frame rates at run-time as well as frame-to-frame coherence. But there are three disadvantages: (1) It has to build a coarse 3D model of the realistic 3D scene, so it is not fit to the fast rendering of the complex outdoor scene. (2) Even to the simple indoor scene, the preprocessing steps of their method also require several hours in all. (3) It has to design many different NPR filters, those NPR filters may achieve different representations in their programming complexity and computational complexity.

In this paper, by using a hybrid NPR/TIP approach we present a new method for none-photorealistic tour into panorama (NP-TIP). First, we design a new algorithm for fast converting a photo or image to a synthesized painting following the painting style of an example image. Second, we proposed a new modeling scheme for TIP based on the cubic panorama technology, making users experience the feeling of "walking-into" the panorama. The amalgamations of the NPR and TIP can not only make the users wander in the non-photorealistic 3D virtual environment, looking on the original real environment in any view of artist, immersing himself into the artistic kingdom, both the vision and spirit, but also avoid some critical problems in the photorealism systems mentioned above.

2 Basic Architecture of NP-TIP System

In this section we will introduce the basic architecture of our NP-TIP system. At a high level, our system proceeds in two steps as shown in Fig. 1. First, during preprocessing, the system has to do three important works (1) Stitches three fisheye images captured at a fixed viewpoint to synthesis the cubic or spherical panorama. The detail of the stitch method is described in papers [13] [14].

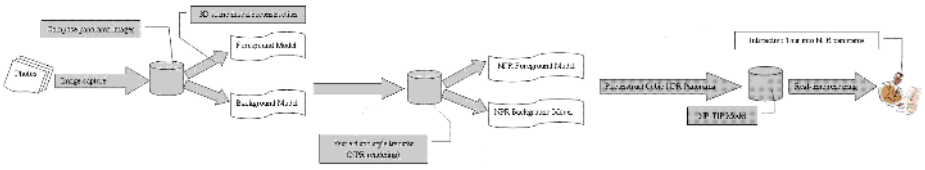


Fig. 1. Basic architecture of NP-TIP system

(2) Reconstructs the foreground model and background model. (3) Fast transfers the artistic style of an example image to the panorama image. Second, during real-time rendering, the system has to do two important works (1) Re-renders the NPR foreground model and background model (2) Interactively tour into the NPR cubic panorama based on the 3D model we created. By abstracting NPR preprocessing into a versatile rendering operation, our architecture supports a number of NPR styles within a common framework. This feature gives users flexibility, as the same panorama model can be used to produce real-time and interactive walkthroughs in different NPR styles.

3 Fast Artistic Style Transfer Algorithm

For the purpose of providing a versatile NPR rendering interface, we need to design a versatile algorithm. Hertzmann presented a universal technique to learn non-photorealistic transformations based on pairs of unpainted and painted example images [7], but, the image analogies method is rather slow since it is pixel-based-it normally takes a few hours to generate a synthesized painting on a PC. Michael Ashikhmin has developed a fast texture transfer technique based on the texture synthesis algorithm, and then put it into NPR rendering [8]. Though this method has quickened the speed of artistic style transformation, the experiment results are not very perfect. Bin Wang has put forward an efficient example-based painting and synthesis of 2D directional texture algorithm [9], but the artistic style of the output images still looked very crudity and stiffness.

In order to make the target image has the similar artistic style with the example painting, a simple method is to make the texture feature of the example painting transfer into the target image according to given rules. Because of human's vision is most sensitive to the change of luminance of image, it is critical to transfer the luminance feature of the example painting into target image, and that the color features of the target image should be consistent with the source image. All of above is the basic thinking of our algorithm. Based on Alexei A.Efros's image texture synthesis algorithm [10], we have designed another patch-based fast artistic style transfer algorithm. In the following sub section, we will introduce the details of our method.

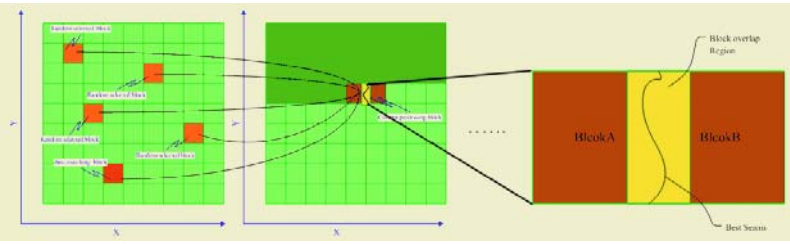


Fig. 2. Fast artistic style transfer

3.1 The Color Space Transform and Luminance Transfer

The texture synthesis algorithm [10] makes the color of output synthesized texture image be in full consistent with the referenced sample image, but we only need to do the process in the luminance channel when transferring the texture features in practice. In order to decorrelate color space, we should process in $l\alpha\beta$ color space [11]. Our NPR rendering algorithm will be discussed in $l\alpha\beta$ color space. In addition, for the purpose of achieving more perfect and accurate matching result when finding the local best block, our approach is to apply a linear map that matches the means and variances of the luminance distributions.

Let $L(t')$ be the luminance of a pixel t' in T, let $L(e')$ be the luminance of a pixel e' in E. Then the luminance of T can be remapped as:

$$L(t') \leftarrow k(L(e') - \mu_l E) + \mu_l T \tag{1}$$

Where $\mu_l E$, $\mu_l T$ are the mean luminance, $\partial_l E$, $\partial_l T$ are the standard deviations of the luminance, both taken with respect to luminance distributions in E and T, respectively. Set k to adjust the luminance variation in target image. When $k = \frac{\partial_l T}{\partial_l E}$, the luminance distribution of target image equals to the distribution of example image. Our means is similar to Hertzmann [7]'s Luminance remapping.

3.2 Fast Artistic Style Transfer

Definition 1. *Let N_1, N_2 be the two isometric blocks, then* the L_2 distance between N_1,N_2 *can be defined as:*

$$Erorr(N_1, N_2) = \sum_{p \in N_1 p' \in N_2} [l_1(p) - l_2(p')]^2 \tag{2}$$

Where l_1,l_2 are the luminance channels in N_1,N_2 respectively; p, p′ are the corresponding pixels in N_1, N_2 respectively.

Definition 2. *Let N_t be an arbitrary block in the target image, SN_e be a set of blocks that are randomly sampled in the example image. Select one block N_e from SN_e which makes the value of $Error(N_e, N_t)$ be minimum. Then N_e is the* best matching block *in SN_e with N_t.*

Fig. 3. Fast artistic style transfer by example image (Watercolor style)

Fig. 4. Fast artistic style transfer by example image (Van Gogh's Style)

Definition 3. *A* best seam *in this paper has the same conception with the* Minimum Error Boundary Cut*[10], but we add the boundary restrictive condition to make it work more properly. (The* **blockOverlap** *in the equation 3 represents the aperture ratio of the overlap region between neighborhood blocks.) The definition is as follows:*

$$\begin{cases} E_{i,j} = e_{i,j} + \min(E_{i-1,j}, E_{i-1,j+1}) \ \textit{if } j = 0; \\ E_{i,j} = e_{i,j} + \min(E_{i-1,j-1}, E_{i-1,j}, E_{i-1,j+1}) \ \textit{if } 0 < j < \textit{blockOverlap}; \\ E_{i,j} = e_{i,j} + \min(E_{i-1,j-2}, E_{i-1,j-1}) \ \textit{if } j = \textit{blockOverlap} \end{cases} \quad (3)$$

When transferring the artistic style between the example image and the target image, we first scans each block in target image in scan line order, at the same time randomly choosing some sampling blocks (usually 400-500 blocks in our experiment) from the example image, which are isometric with the blocks in target image. Second, we computes the L_2 distance between the sampling blocks and the current scanning block in target image, then we choose a sampling block that has the minimum L_2 distance with the current scanning block in the target image. Third, we copy this sampling block to target image. In this way, we can fast transfer the luminance features of the example image to the target. But this method will yield distinctive boundary between blocks. In order to eliminate the boundary between blocks, we have introduced a method as same as A.Efros's image texture synthesis algorithm [10]. It is just allowing some overlap between blocks when doing the synthesis process. Then we can find the best seam in the overlap region by using the dynamic programming. Finally, we sew the overlap region of blocks along the best seam of image by using alpha blend. Fig. 2 shows

two overlapping texture blocks selected for the corresponding source blocks and the seam line between them.

3.3 Color Transfer

After being processed by this way, the target image has taken on the similar texture features with the example image. In order to keep the color characteristics of the target image consistent with the source image, we only need to copy the color information of the channels from the source image to the target image. Fig. 3 and Fig. 4 shows the NPR rendering results; the example images are from reference [7].

4 Non-photorealistic Panorama Walkthrough

4.1 3D Scene Model Reconstruction Based on a Vanishing Line

According to the imagination and understanding to the input image, user can easily distinguish foreground from background. Thus the whole 3D scene model reconstructed mainly consists of background model and foreground model.

Backgroud Model. The input image is divided into two disjoint regions by the vanishing line. The region below the vanishing line in the image corresponds to ground plane, and that above the vanishing line corresponds to back plane. Since vanishing line contains all vanishing points formed by all parallel lines on a plane, back plane can be thought of as a plane of infinite distance and the points on it can be called the ideal points. For easy computation, we assume that the camera is positioned at the origin, the view direction is towards $+z$, the view-up vector is towards $+y$ and the focal length of the camera is d. Thus, the coordinates of each vertex in the background model can be obtained [3].

Foreground Model. A foreground object specified in the image is modeled as a 3D polygon, called foreground model [3]. Suppose that the polygons stand perpendicular to the ground plane. The coordinates of its vertices are then computed by finding the intersection points between the ground plane and the ray

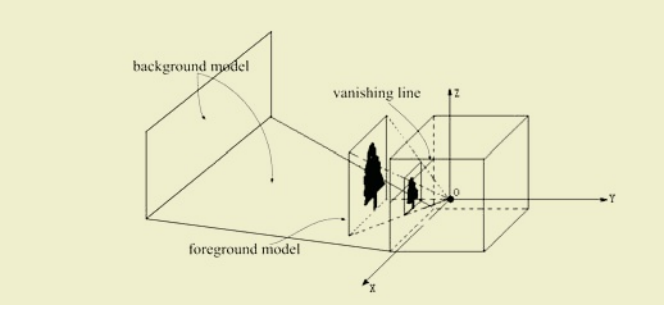


Fig. 5. the local model of one side of cube

starting from the viewpoint and passing through the corresponding vertices in the image plane.

As proposed in [3], a foreground object can have a hierarchical structure in a more complex environment. That is, another foreground object can be attached on a foreground object standing on the ground plane to form a hierarchical structure. Together with the scene model, the images called background image and foreground mask are generated by segmenting the foreground objects from the input image. The background image is used as a texture, which is to be mapped onto the background model. Foreground mask is used to distinguish the exact portion of the foreground object from the background. After performing the above all steps, we can render the scene model by changing the parameters of camera.

4.2 Non-photorealistic Tour into Panorama

Although Kang had proposed a scene model based on a vanishing circle, spherical panorama is a kind of uneven sampling representation, there are two disadvantages:(1) Easily leads to distortion, especially on the North and South Pole. (2) Needs additional transformation from non-linear map into linear map. Therefore, this paper proposes a kind of NPR walkthrough model based on 3D cubic panorama, which is easier in projection and simpler in calculation.

For cubic panorama, there is no important content on the top and the bottom sides in general. Thus the system just simplifies to model the four sides using vanishing line based TIP techniques. Supposed that the center of cube (viewpoint) is positioned at the origin. Fig. 5 illustrates the local model of one side of cube. We make local model for the four sides of cube respectively. Since the input image is panorama, the vanishing line on each of the four sides will have the same height above the bottom side. So the four local models will put together to form a global scene model. Then, the system just needs to project the top and the bottom sides of cubic panorama on the top and the bottom of the global model respectively. Finally, we get a close hexahedron scene model.

With the above, the system renders the input image and the background image using fast artistic style transfer algorithm mentioned in Section 3 and generates textures. Finally, 3D interactive real-time walkthrough with different artistic style can be realized by using texture mapping which maps the textures onto formerly reconstructed 3D scene model. Since system deals with each side severally, maybe there is aberration between the local models. That can be rectified using the method in Section 3.2.

5 Experiment Results

Our system works on Pentium III 1000MHz PC using Microsoft Visual C++ 6.0 and the OpenGL texture mapping routines. The interactive real-time non-photorealistic scene navigation is possible for an output of full-screen display. The system has been applied to the panoramic walkthrough of Kunming horticultural

Fig. 6. panorama walkthrough with expressionist style

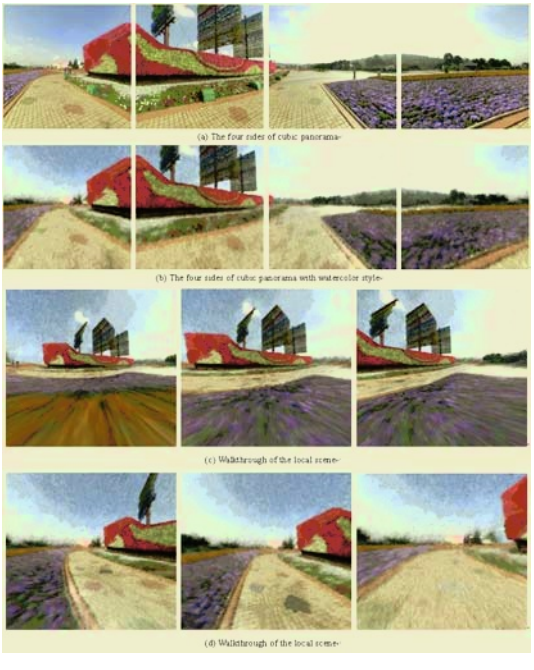


Fig. 7. panorama walkthrough with watercolor style

expo garden in China. In order to present much more artistic style to users and enhance flexibility and extensibility, we have integrated Hertzmann's algorithm [12] into our system. Fig. 6 and Fig. 7 are some results of our system.

6 Conclusions

As yet, whether we see NPR for 3D scene from the speed or from the quality of rendering, its effects are still not very perfect. In this paper, we describe a system called NP-TIP. It has below key features: 1) Painting styles are represented as one or more blocks of sample textures randomly selected from the example painting; 2) The algorithm allows some overlap between blocks during the synthesis process to eliminate the boundary effect between blocks. This method can also eliminate the aberration between the local models of cubic panorama. 3) By copying the color information of the $\alpha\beta$channels from the source image to the target image, NP-TIP can easily keep the color characteristic of target image consistent with the source image. 4) With the modeling scheme for TIP based on the cubic panorama, the system not only overcomes the problem of fixed viewpoint in panorama walkthrough based on image series, but also models easily. 5) Since cubic panoramic image is a line-to-line projective mapping, we can directly use it as a texture map and don't need an additional transformation from the non-linear map into a linear map, which reduces the complexity of computation. Our system is not only fit to the fast rendering of the small-scale scene, but also to that of the complex large-scale scene. Furthermore, according to users' selection from example paintings of different artistic style, it can provide interactive and real-time panorama walkthrough of corresponding style.

Acknowledgements

This work is supported by Province and University Science & Technology Cooperation Projects of Yunnan Province and Zhejiang University (2001JAALA02A022).

References

1. A. W. Klein, W. Li, M. M. Kazhdan, W. T. Corra, A. Finkelstein, and T. A. Funkhouser. Non-photorealistic virtual environments. In Proceedings of ACM SIGGRAPH 2000 (27th International Conference on Computer Graphics and Interactive Techniques), pages 527-534, 2000.
2. Horry, Y., K. Anjyo and K. Arai. "Tour Into the Picture: Using a spidery mesh user interface to make animation from a single image," In SIGGRAPH 97 conference proceedings, Computer Graphics, August 1997, pp. 225-232.
3. H. W. Kang, S. H. Pyo, K. Anjyo, S. Y. Shin, "Tour Into the Picture using a Vanishing Line and its Extension to Panoramic Images", Eurographics 2001, Volume 20, No. 3, 2001.

4. Robert D. Kalnins, Lee Markosian, Barbara J. Meier, Michael A. Kowalski, Joseph C. Lee, Philip L. Davidson, Matthew Webb, John F. Hughes and Adam Finkelstein. WYSIWYG NPR: Drawing Strokes Directly on 3D Models. ACM Transactions on Graphics. 21(3):755-762, July 2002.
5. Meier, B. J.: Painterly Rendering for Animation. In: Proceedings of SIGGRAPH 96, pp. 477-484, 1996
6. Mario Costa Sousa and Kevin Foster and Brian Wyvill and Faramarz Samavati. Precise Ink Drawing of 3D Models. Special issue: Proceedings of EuroGraphics 2003. Computer Graphics Forum, vol.22, no.3,(2003), pp.369-379
7. Hertzmann, A., Jacobs, C. E., Oliver, N., Curless, B., and Salesin, D. H. Image analogies. In Proceedings of ACM SIGGRAPH 2001, ACM Press/ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, 2001. 327-340.
8. Ashikhmin, M. Fast Texture Transfer IEEE Computer Graphics and Applications, vol.23 no. 4 (2003), pp.38-43
9. B. Wang, W. Wang, H.P. Yang, and J.G. Sun, Efficient example-based painting and synthesis of 2D directional texture. IEEE Transactions on Visualization and Computer Graphics, vol.10, no. 3, (2004), pp. 266-277.
10. Alexei A. Efros, William T. Freeman. Image quilting for texture synthesis and transfer. In the Proceedings of the 28th annual conference on Computer graphics and interactive techniques.2001, 341-346
11. Erik Reinhard, Michael Ashikhmin, Bruce Gooch, Peter Shirley. Color Transfer between Images. IEEE Computer Graphics and Applications. September 2001. (Vol. 21, No. 5), pp.34-41
12. Hertzmann, A. 1998. Painterly rendering with curved brush strokes of multiple sizes. In Proceedings of SIGGRAPH 98, ACM SIGGRAPH/Addison Wesley, Orlando, Florida, Computer Graphics Proceedings, Annual Conference Series, 453-460
13. Y.L.Zhao, Real-Time Navigation Algorithm Research Based on Panoramic Image [MS. Thesis]. Yunnan University,Kunming, 2004 (in Chinese with English abstract)
14. M.W.Chen, Research for the Panorama Mosaics [MS. Thesis]. Yunnan University, Kunming,2004 (in Chinese with English abstract)

Image Space Silhouette Extraction Using Graphics Hardware

Jiening Wang[1], Jizhou Sun[2], Ming Che[2], Qi Zhai[3], and Weifang Nie[2]

[1] CAAC ATM Research Base, Civil Aviation University of China,
Tianjin 300300, P.R.China
jnwang@cauc.edu.cn
http://www.cauc.edu.cn
[2] IBM Computer Technology Center, Department of Computer Science,
Tianjin University, Tianjin 300072, P.R.China
http://ibm.tju.edu.cn
[3] School of Material Science and Engineering,
Tianjin University, Tianjin 300072, P.R.China

Abstract. In computer graphics, silhouette extracting and rendering has an important role in a number of application. Special features such as silhouette and crease of a polygonal scene are usually displayed by identifying for the corresponding geometry especially in non-photorealistic rendering (NPR). We present an algorithm for extracting and rendering silhouette outlines and crease edges of 3D polygonal meshes in image space. This algorithm is simple and suitable for modern programmable graphics hardware implementation. It can generate NPR image synthesis in real-time. Our experimental results show it satisfies real-time interactive needs.

1 Introduction

The field of computer graphics is concerned with techniques for creating images via computer in order to communicate visual information. In view of this idea, many algorithms and techniques have been proposed to generate photorealistic image synthesis. These include reflectance light models such as Cook-Torrance[1]and Ashikhmin[2]models, ray-tracing[3]and radiosity [4]etc. However photorealistic imagery might not be the best choice for communicating information about 3D scenes. One reason is that sharp features such as silhouette could convey a grate deal of information in a CAD diagrams. These facts induce a new rendering approach called non-photorealistic rendering (NPR), which can create compelling line drawings from 3D models. The NPR applications include illustrations, painterly renderings, CAD/CAM 3D model representation and cartoon production. The most common features in NPR are silhouettes and creases, which can enhance the visual results in very efficient way. For polygonal meshes, the silhouette edges are usually the visible segments of all edges that connect back-facing polygons to front-facing polygons. Meanwhile a crease edge is a ridge if the dihedral angle between adjacent polygons is less than a threshold.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 284–291, 2005.

Generally, the NPR algorithms are classified into object-based, image-based or hybrid approaches according to the techniques on how to find out the silhouette outlines and crease edges. Up to now, there are some techniques have been explored and some approaches can do silhouettes extraction at interactive rates that were described in[5][6][7][8][9]. The traditional approach is what called "brute force" that traverse through the scene polygon graph and for each edge finds the relationship between adjacent polygons to determine whether it is a silhouette or a crease edge. Unfortunately, explicitly identifying such edges is a burdensome process and not supported by current graphics hardware. However, image space algorithms turn to solve the problem in 2D image processing. One of the most important approaches is a geometric buffer, known as G-buffers, which preserve geometric properties of scenes objects in image space such as normals, depth, or object identifiers[10]. A challenging task is that how to find appropriate implementations of G-buffer generation and extract silhouette outlines and crease edges in real-time rendering. Recent developments in programmable graphics hardware armed with powerful GPU use programmable graphics pipeline instead of traditional fixed function graphics pipeline. This leads to new rendering algorithm.

Our algorithm is largely dependent on G-buffer techniques. To achieve real-time performance, we construct G-buffer data using vertex programs and fragment programs which are built in graphics hardware. Then we also do 2D image processing to find out silhouette outlines and crease edges on graphics hardware directly. Finally we blend these edges and rendering results of 3D scenes in frame buffer.

The remainder of this paper is structured as follows. Section 2 gives brief introduction of current graphics hardware architecture. Section 3, the main contributions of this paper, illustrates the main features and implementations of the algorithm. Section 4 figures out the lighting model used in this paper. Section 5 draws conclusion.

2 Modern Graphics Hardware Architecture

Most contemporary programmable graphics hardware architectures have migrated standard pipeline into programmable one as shown in Fig.1. The transform and lighting computation integrate into one step in which the position, color and lighting of scene objects are determined by vertex program written in a custom shader program. Similarly, the texturing and color composing stages are also collapsed into a single abstract stage where the outputs are determined by shader programs. The GPU in graphics hardware could execute all these programs through 4-way SIMD instructions. That is the current programmable graphics hardware can do computations in parallel mechanism and we can achieve more efficiency by writing our algorithm with graphics hardware instructions.

Originally, hardware programming interface is low-level; it provides an assembly-language-like interface or an explicit pipeline-configuration model. This

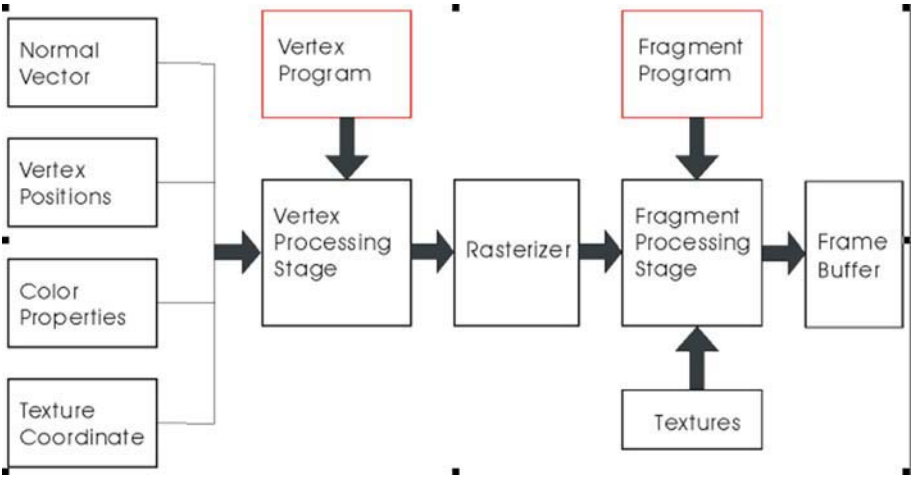


Fig. 1. The graphics processing pipeline, the red box functions mean programs can be run on GPU

interface often related to specified hardware vendors, for example, Nvidia's GPU gives a set of vertex program vector-based assemble operators, a set of texture shader operators (executed by TS1.0) and a set of register combiner operators (executed by RS1.0). Along with the developing of programmable graphics hardware, many high-level shader language have been introduced, the most common is Nvidia's Cg (C for graphics) language[11]. The Cg language is based on both the syntax and the philosophy of C language, and is a hardware-oriented language. Cg defines an interface for data transferring with OpenGL extension. Our works are largely dependent on Cg programming.

3 Silhouettes and Crease Edges Extraction in Image Space

As described before, the silhouette outlines and crease edges extraction and their renderings are very critical steps in NPR. They convey the primary features of objects in a 3D scene and the observer can achieve the outlines of 3D objects with the help of them. Firstly we give the definitions of silhouette and crease edges:

Definition of a Silhouette. Given $E(u,v)$ as the eye vector, a point on a surface $S(u,v)$ with surface normal $N(u,v)$ is a silhouette point if $E(u,v) \cdot N(u,v)=0$.

Definition of Silhouettes for Polygonal Models. Given $E(u,v)$ as the eye vector, a set of polygons $S=\{N_1, N_2, \ldots, N_n\}$, where N_i represents the normal vector of ith polygon. We assume that polygon normals point outward from surface, then it can be said if $E \cdot N_i < 0$ the polygon is front-facing, if $E \cdot N_i > 0$ the

polygon is back-facing. A silhouette edge is an edge adjacent to a front-facing polygon and a back-facing polygon.

Definition of a Crease Edge. Given two front-facing polygons (or back-facing polygons) $P_1=\{N_1\}$ and $P_2=\{N_2\}$, a crease edge is an edge between P_1 and P_2 if $N_1 \cdot N_2 < \delta$, where δ is some threshold.

Generally, we have the facts that an abrupt change in the z-depth value occurs at silhouette profiles. This induces that we can detect silhouette edges from z-depth map of a typical 3D scene by 1^{th} order differential operator in image space. Additionally abrupt changes in the normal-buffer occur at crease edges. Also we can detect them by inspecting the difference of two adjacent normal values in image space. Actually the normal-buffer detecting approach can supplement the z-depth methods. The overview of our algorithm is below:

Firstly, using vertex program and fragment program to generate G-buffer data including normal-buffer and depth-buffer in a single texture object, then with the image space edge detecting algorithm implemented in graphics hardware to extract the silhouette outlines and crease edges directly.

Secondly, to use the same techniques to render the 3D scenes with NPR lighting model and to blend the detecting edges with the rendering results using projecting texture mapping techniques.

This algorithm presented in this paper can satisfy real-time interactive NPR applications.

3.1 G-Buffer Constructions and Storage Mechanism

The G-buffer includes normal-buffer and z-depth buffer. In the first, objects in 3D scenes are rendered into a 2D texture with fragment normal values and z-depth values. In order to have this step done, we turn to some hardware graphics techniques that described bellow:

Upon using a normalization cube map texture, to render eye-space normalized fragment normals of objects with Cg fragment program. The key Cg statement is: float3 normalcolor = (texCUBE(normalMap, normalVector)).rgb, here normalMap is a normalization cube map texture.

Upon using a 1D GL-LUMINANCE texture, to render eye-space z-depth value of objects also with Cg fragment program. The key Cg statement is: float4 zColor = tex1D(zDepthMap, zValue), here zDepthMap is a 1D grey-scale texture.

Upon using Render-To-Texture (RTT) techniques[12], we can store above rendering results into a pixel-buffer. The pixel-buffer is constructed and reconfigured according to current drawing window size dynamically and can be converted into a texture object TexObj. The TexObj uses RGBA format and each texel content is {normalcolor.r, normalcolor.g, normalcolor.b, zColor.w}.

3.2 Edge Detection in Image Space Based-on Graphics Hardware Programming

Modern programmable graphics hardware was primarily designed for 3D object rendering pipeline, it is not intended to solve image-processing problem with

its hardware architectures. One way to realize image processing is to copy the contents in a pixel-buffer into main memory, and to process it in normal way supported by CPU. The post-processing result would be converted to an OpenGL texture object and transferred it back into texture memory for further use. This will lead the performance problem. In this paper, we choose to do image processing with graphics hardware instructions. Since the graphics pipeline provides 4-way vector parallel streaming-oriented processing techniques, it will be more efficient in image processing. Considering the fact that Z-depth buffer image is gray-scale and its abrupt changes primarily caused by silhouette profiles, which often have strong connectivity features. We adopt *Kirsh* operator to detect discontinuity edges in Z-depth buffer. On the other hand, the normal buffer contains RGB that represents the fragment normal vector; we can use dot product value of neighbor texels to inspect the abrupt changes of normal.

Z-Depth Buffer Edge Detection Using Kirsh Operator

Considering the texel neighborhood relationship of (s_0,t_0) in a typical texture object (see Fig. 2), we will use Kirsh operator to determine the abrupt changes of this texel (s_0,t_0). The Kirsh operator is formularized as bellow:

$$K(s_0,t_0)=\max(1, \max(5s_i\text{-}3t_i)) \ (i=1,2,\ldots,8) \tag{1}$$

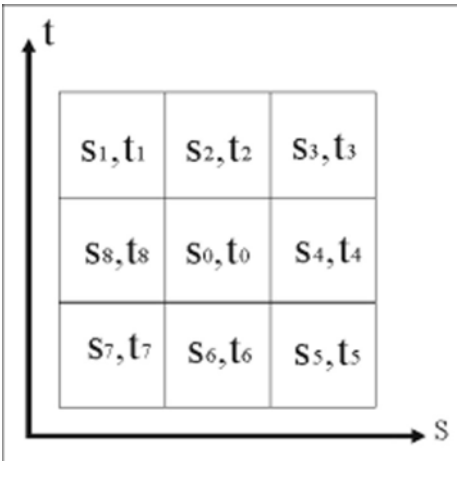


Fig. 2. Neighbors of texel (s_0,t_0)

Where s_i is first three neighbors sum, and t_i is the sum of other neighbors, give a threshold H_k, if $K(s_0,t_0) > H_k$ then the texel (s_0,t_0) is an edge texel. We implement the algorithm with Cg fragment program based on intermediate rendering passes, which render a textured window-aligned quad using multitexturing. The core Cg code is listed bellow:

```
float2 sampleroffset[8]; float kirsh(uniform sampler2D tex, float2
texCoord, int offset) {
  float2 texCoords[8];  float2 texColor[8];
  float  k1;   float  k2;
  for(int i=0; i<9; i++) { texCoords[i]=
     texCoord+sampleroffset[(offset) fmod (8.0)]; }
  for(i=0; i<9; i++) {
    texColor[i]=tex2D(tex, texCoords[i]); }
    k1=5.0f*(texColor[0].w+texColor[1].w+texColor[3].w);
    k2=3.0*(texColor[4].w+texColor[5].w+ texColor[6].w+
       texColor[7].w)+texColor[8].a;
    return (k1-k2);
 }
```

Normal-Buffer Edge Detection

Normal-buffer contains fragment normals of object in a 3D scene that represent in *(R, G, B)*value. To achieve the abrupt changes of normal value we inspect (s_2,t_2), (s_4,t_4), (s_6,t_6) and (s_8,t_8) of texel (s_0,t_0). As described before, firstly we sample these texel color, let them are *texColor0, texColor1, texColor2, texColor3* and *texColor4* respectively. Here we calculate dot product between *texColor0* and other neighbor *texColor(i) (i=1,2,3,4)* respectively to determine the edge. If the dot product less than a threshold H the texel is an edge texel. The core Cg code is below:

```
float4 detec_val(dot(expand(texColor0),expand(texColor2));
float4 detec_val(dot(expand(texColor0),expand(texColor4));
float4 detec_val(dot(expand(texColor0),expand(texColor6));
float4 detec_val(dot(expand(texColor0),expand(texColor8));
float4 threshold(H,H,H,H);
float4 test = detect_val - threshold;
float  final_color = step(float4(0,0,0,0),test);
```

Now through the above two passes edge detection, we can achieve black-white texture object in texture memory that can be used by NPR techniques. Fig. 3(c) and Fig. 3(d) show the experimental results. We find that normal-buffer edge detection would produce more aliased edges.

Finally we can combine the result texture with NPR image by projective texture mapping techniques.

4 NPR Lighting Model

In this paper we simply use following lighting model for 3D scene rendering:

$$I=((1+L\cdot N)/2)k_{cool}+(1-((1+L\cdot N)/2))k_{warm} \quad (2)$$

Where L is normalized light vector, N is the normalized normal vector, $k_{cool}=k_{blue}+\alpha k_d$, $k_{warm}=k_{yellow}+\beta k_d$, $k_{blue}=(0,0,b)$ $0\leq b\leq 1$, $k_{yellow}=(\gamma,\gamma,0)$

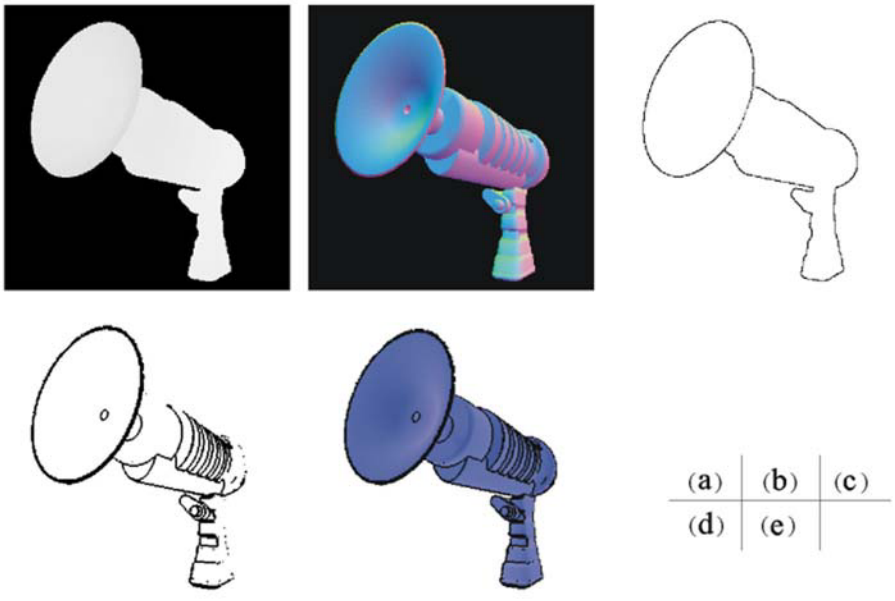


Fig. 3. Experimental rendering results, (a) Z-depth buffer (b) normal-buffer (c) Z-depth buffer detecting silhouette outlines (d) normal-buffer detecting crease edges (e) final blending result

$0 \leq \gamma \leq 1$, k_d is diffuse color coefficient. Here b and γ determine the temperature shift by giving the maximum "blueness" and "yellowness". The other two values, α and β, determine the amount of which actual object color is visible.

We implement this lighting model using Cg vertex program, its performance also can satisfy real-time rendering needs. Fig. 3(e) shows the final render result.

5 Conclusion

The presented algorithm solves the silhouette outlines and crease edges extraction and rendering with graphics hardware acceleration. Our experimental results show it satisfies real-time interactive application in NPR. We implemented our algorithm totally based-on Cg programming and completely fitted to today's programmable graphics hardware.

The main disadvantage of the algorithm is how to do anti-aliasing of the detection texture. In near future we will turn to image erosion techniques to get bone of detect image and fitted with Bezier curve to solve the problem.

References

[1] Cook, R. L., Torrance, K. E. : A Reflection Model for Computer Graphics. Computer Graphics. **22(3)** (1982) 7–24

[2] Ashikhmin, M., and Shirley, P. : An Anisotropic Phong BRDF Model. Journal of Graphics. **5(2)** (2000) 25–32
[3] Whitted, T. : An Improved Illumination Model for Shaded Display. Comm. ACM. **23(6)** (1980) 343–349
[4] Nishita T. and Nakamae E. : Continuous Tone Representation of Three-Dimensional Objects Taking Account of Shadows and Interreflection. Computer Graphics. **19(3)** (1985) 23–30
[5] Jarek Rossignac, Maarten van Emmerik : Hidden Contours on a framebuffer. Proceedings of the 7th Workshop on Computer Graphics Hardware, Eurographics. (Setp. 1994)
[6] Lee Markosina, Michael A. Kowalski, Samuel J. Trychin, Lubomir D. Bourdev, Daniel Goldstein, and John F. Hughes : Real-Time Non-photorealistic Rendering. Proceedings of 1993 Parallel Rendering Symposium. SIGGRAPH'97 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addsion Wesley (1997), 415-420
[7] Ramesh Raskar, and Michael Chhen : Image Precision Silhouette Edges. ACM Symposium on Interactive 3D Graphics (1999), 135-140
[8] Bruce Gooch, Peter-Pike J. Solan, Amy Gooch, Peter Shirley and Richard Riesenfeld : Interactive Technical Illustration. Proceedings of the 1999 Symposium on Interactive 3D Graphics (1999), 31-38
[9] A. Hertzmann : Introduction to 3D Non-photorealistic Rendering: Silhouettes and Outlines. SIGGRAPH Course Notes (1999)
[10] Saito, T. and Tahahashi, T. : Comprehensible Rendering of 3D Shapes. In Computer Graphics (Proceedings of SIGGRAPH'90). **24(4)** (1990) 197–206
[11] William R. Mark, R. Steven Glanville, Kurt Akeley, Mark J. Kilgard : Cg: A System for programming Graphics Hardware in a C-like Language. In Proceedings of ACM SIGGRAPH 2003, 896-907.
[12] Chris Wynn : OpenGL Render-to-Texture. http://developer.nvidia.com GDC 2002.

Adaptive Fuzzy Weighted Average Filter for Synthesized Image

Qing Xu, Liang Ma, Weifang Nie, Peng Li, Jiawan Zhang, and Jizhou Sun

School of information, Tianjin University

Abstract. Monte Carlo is a powerful tool for the computation of global illumination. However noises, which are resulted from the low convergence of Monte Carlo, are noticeable for the synthesized image with global illumination effects and can be regarded as the combination of Gaussian noise and impulse noise. Filter is a cheap way to eliminate noises. In this paper, we investigate nonlinear filtering techniques to reduce the mixturing noise. Based on fuzzy theory, we present an adaptive weighted average filter to optimize the weighs of the filters. Analysis and the computational results, which have been obtained from experiments for noise attenuation and edge preservation, indicate that the new algorithm is promising.

Keywords: Fuzzy set, membership function, mixture noise, adaptive filter, Monte Carlo, global illumination.

1 Introduction

Monte Carlo based methods are quite suitable for the calculation of global illumination problem when highly complicated scenes with very general and difficult reflection models are rendered. Especially, Monte Carlo is applied as a method of last resort when all other analytical or numerical methods fail [1]. However, the synthesized image generated by using Monte Carlo based global illumination algorithms is very noisy. Usually, we can reduce the noise through taking more samples within each pixel to get a visually acceptable image, but a lot of rendering time must be spent because of the slow convergence of the Monte Carlo technique.

Filtering is a cheap and good way to reduce noises for the synthesized images. Noise resulted from Monte Carlo can be considered as the combination of Gaussian noise and impulse noise. L-filter [2] and alpha-trimmed filter [3] are simple cases for this kind of filter.

Fuzzy theory [4, 5] has been widely used in the fields of pattern recognition, image processing [6, 7, 8, 9, 10, 11] and automatic control [12, 13]. In this paper, we propose an adaptive fuzzy weighted average filter (AFWA) based on fuzzy theory. The new filter regards the pixels in the filter window as a fuzzy set. Every pixel in the filter window can be depicted by membership function, which is actually the weight of this pixel. The optimal weight can be derived from

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 292–298, 2005.

fuzzy iteration. The novel filter can cope with impulse noise and Gaussian noise soundly, in the mean time, they can preserve edge perfectly.

This paper is organized as follows. In section 2 we introduce the fundamental of fuzzy set. In Section 3 we detail the AFWA. Experimental results are presented in section 4. And finally, conclusion is drawn in section 5.

2 Fuzzy Set and Image

Membership, in the N-dimensional Euclidean space E^N, of a crisp set A, can be defined in terms of the characteristic function $\mu_A :\rightarrow \{0,1\}$

$$\mu_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \tag{1}$$

Membership of fuzzy set is characterized by its membership function $\mu_A :\rightarrow \{0,1\}$. The value $\mu_A(x)$of the membership at point x of the N-dimensional Euclidean space denotes the degree that point x belongs to set A.

The basic set operations, i.e. intersection, union, complement and difference are defined as follows:

$$\mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)] \tag{2}$$

$$\mu_{A\cup B}(x) = \max[\mu_A(x), \mu_B(x)] \tag{3}$$

$$\mu_{A^c}(x) = 1 - \mu_A(x) \tag{4}$$

$$\mu_{A\setminus B}(x) = \min[\mu_A(x), 1 - \mu_B(x)] \tag{5}$$

Finally, the subset relation is

$$A \subseteq B \Leftrightarrow \mu_A(x) = \min[\mu_A, \mu_B] \Leftrightarrow \mu_B(x) = \max[\mu_A, \mu_B] \tag{6}$$

Let Z^2 be the Cartesian grid and $G = \{0, 1/L, 2/L, ..., L/L\}$the set of L+1 normalized gray-levels. A normalized image A is defined as a mapping $A : Z^2 \rightarrow G$.

An image can be considered as an array of fuzzy singletons, each with a membership function equal to the normalized gray-level value of the image at that point. A fuzzy singleton is a fuzzy set whole support is a single point.

3 The AFWA Filter

3.1 The Fuzzy Weighted Average Filter (FWA)

The FWA filter [14] regards the pixels in the filter window as elements of a fuzzy set A. Each pixel in the window can be described by membership function

$\mu_A(d_i)$, where d_i is the different between the pixel x_i and the center c of the set. The membership function $\mu_A(d_i)$ is defined by a decreasing function whose domain is $[0, +\infty)$ and output range is [0, 1]. Here we use the Gaussian function as the membership:

$$\mu(x) = e^{-\frac{(x-c)^2}{2\beta^2}} \tag{7}$$

Where β is the scale parameter.

$S(c, \beta)$ is the sum of the membership of all the pixels in the filter window:

$$S(c, \beta) = \sum_{i=1}^{N} \mu_A(x_i) \tag{8}$$

Obviously, when the function $S(c, \beta)$ get maximal S_{max} with a fixed scale parameter β, we can find the optimal fuzzy set and the fuzzy center by iteration:

$$c = \sum_{i=1}^{N} e^{-\frac{(x_i-c)^2}{2\beta^2}} x_i / \sum_{i=1}^{N} e^{-\frac{(x_i-c)^2}{2\beta^2}} \tag{9}$$

Note that the fuzzy set center c is the output of the filter.

3.2 The AFWA Filter

The FWA filter utilizes the fixed scale parameter to do filtering. But the scale parameter should adapt to the different cases because the image to be is non-stable. For example, the scale parameter should be small when the area to be filtered is of edge or impulse noise, while the scale parameter should be large when the area is constant and is not of impulse noise. In this section, we address the optimization of the scale parameter β for the FWA filter.

Actually, the scale parameter β is a boundary between the edge pixels or pixels corrupted by impulse noise (heavy-tail noise) and constant pixels or pixels corrupted by Gaussian noise (short-tail noise). When there do not exist any heavy-tail noise and edges in the filter window, the scale parameter trends to infinity, while when there do not exist any short-tail noise in the window, the scale parameter β should be small enough to preserve the edges and eliminate heavy-tail noise.

For the pixels in the filter window, we can define sets S and B

$$S = \{x_i \, | |x_i - c| < \beta\} \tag{10}$$

$$B = \{x_i \, | |x_i - c| > \beta\} \tag{11}$$

The constant pixels and the pixels corrupted by short-tail noise are in set S, while the edge pixels and pixels corrupted by heavy-tail noise are in set B. When the set B is empty, we will use average to replace the output. And when the local derivation of the set S is equal to zero, we will use the median of the

pixels to replace the output. When the two conditions are not satisfied, we will optimize the scale parameter β during the iteration algorithm.

To optimized the scale parameter β, we first get the local derivation of the set S

$$\hat{\sigma}^{=}\sqrt{\frac{1}{M}\sum_{x_i\in S}(x_i-c)^2} \tag{12}$$

where M is the number of the pixels in the set S. Then we get the average difference in the set B

$$D=\frac{1}{N-M}\sum_{x_i\in B}|x_i-c| \tag{13}$$

To compromise the needs for smoothing short-tail noise and eliminate heavy-tail noise, we get the following equation

$$De^{-\frac{D^2}{2\beta^2}}=\lambda\hat{o} \tag{14}$$

where $0<\lambda<1$. From equation (14) we get

$$\beta=\frac{1}{2}\sqrt{\frac{D^2}{\ln\frac{D}{\lambda\hat{\sigma}}}} \tag{15}$$

And from equation (15), we note that when $\hat{\sigma}$ is increasing, the scale parameter β is increasing, and then the smoothing capability is increasing. On the other hand, when D is increasing, the rate $\frac{D}{\lambda\hat{\sigma}}$ is also increasing, and then the increasing rate of scale parameter β is more slowly than D or even it is decreasing. Therefore the edge preservation and attenuation of heavy-tail noise capability of the algorithm is increasing with the increase of the D and the decrease of scale parameter β.

To guarantee the convergence of the filter, we add another constraint: when the new difference of D and $\hat{\sigma}$ is smaller than the old one, the scale parameter β is not changed.

The AFWA algorithm is described as follows:

1. Let c_0, which is equal to the average of the set, is the initial output center. And we denote the initial β,$\beta_0=\min(\sigma,50)$. From c_0 and β_0, we get memberships of the elements in the set from membership function and this center: $\mu_{0i}=\mu(x_i,c_0,\beta_0)$.
2. Get the result of the weighted averaging of the set in which the weight of every element is equal to the membership of the fuzzy set. Let the output be new center c_1. From c_1 and β_0, we change the pixels in the filter window into two sets, and from (32), we get new scale parameter β_1. Comparing the new difference of D and $\hat{\sigma}$ is with the old one, if it is smaller than the old one, keep the scale parameter β unchanged.
3. Compare the new center with the old center. If the difference of the two centers is smaller than the pre-established value ε, the iteration ends; otherwise go to 2 with the old center replaced by the new center.
4. If the iteration ends, the final center is the output of the set.

4 Experimental Results

In this section, we study the performance of the AFWA filter to reduce different noises.

Table.1 indicate that the smoothing properties of different filters for mixture noise. Here, NMSE (Normalized Mean Square Error) is defined as follows:

$$NMSE = \frac{\sum_{i,j} [x(i,j) - y(i,j)]^2}{\sum_{i,j} x(i,j)^2} \tag{16}$$

Where, $x(i,j)$ and $y(i,j)$ is the origin image without noise and the image processed by filters.

Fig. 1 is an image corrupted by mixture noise which includes Gaussian ($\sigma = 20$) and impulse noise (p=10%). Fig. 2 is the result of median filter of the noisy image. Fig. 3 is the result of the AFWA filter with scale parameter $\beta = 50$, and Fig. 4 is the result of the AFWA filter whose parameter $\lambda = 0.9$. From these figures and Table 1, we can find that the performance of AFWA filter is better than other filters for reducing mixture noise. We also compare the AFWA filter, the average filter and the median filter for Gaussian noise with different derivation (Fig. 5). From Fig. 5, we can find that except for extreme conditions (the noise is too small or too heavy) the AFWA filter is better than the average filter and the median filter.

Table 1. NMSE of images processed by kinds of filter

Noise	FWA filter					AFWA filter
	$\beta = 30$	$\beta = 40$	$\beta = 45$	$\beta = 50$	$\beta = 60$	
Gaussian	0.009464	0.008602	0.008360	0.008244	0.008170	0.008093
mixture	0.012033	0.011248	0.010985	0.010951	0.011291	0.010773

Fig. 1. Image corrupted by mixture noise

Fig. 2. Resultant image processed by the median filter

Fig. 3. Resultant image processed by the FWA filter

Fig. 4. Resultant image processed by the AFWA filter

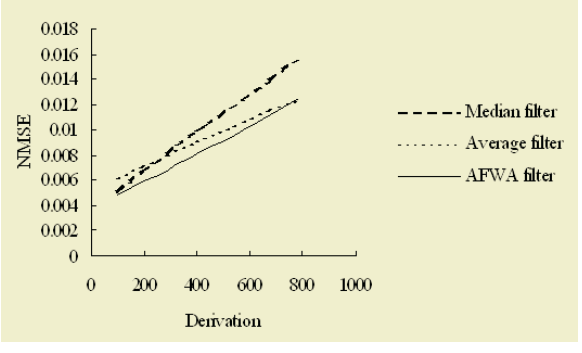


Fig. 5. Comparison of different filter to Gaussian noise with different dervation

5 Conclusion

We have presented the new adaptive fuzzy weight average filtering scheme based on fuzzy theory. Fuzzy membership function is introduced as a basis to obtain optimal weight by using iteration. The convergence of the new filter can be proved soundly in theory. Also, the experimental results for image filtering and edge preservation show an improvement over the classic methods.

In the near future, we will study on the more elaborate iteration procedure falling on some sound theoretical framework, for instance, information theory.

References

[1] P. Bekaert. Hierarchical and Stochastic Algorithms for Radiosity. Ph.D. Dissertation, Katholieke Universiteit Leuven, Belgium, December 1999.
[2] F. Hampel, E. Ronchetti, P. Rousseeuw, W. Stahel, "Robust Statistics", John Wiley, 1986.

[3] I. Pitas, and A. N. Venetsanopoulos, "Nonlinear digital filters: principles and applications", Kluwer academic publishers, 1990.
[4] Guo guirong, Zhuang zhaowen"Fuzzy Technology in Information Processing" Publishing company of university of national defense technology 1996
[5] He xingui "Theory and Technology of the Processing of Fuzzy Knowledge" Publishing company of national defense industry 1998
[6] YoungSik Choi and Raghu Kirshnapuram "A Robust Approach to Image Enhancement Based on Fuzzy logic" **IEEE Trans. Image Processing**, vol. 6, No. 6, pp.808-825, Jun 1997.
[7] H. D,. Cheng, C. H. Chen and H. H. Chui, " Fuzzy homogeneity approach to multilevel thresholding", **IEEE Trans. on IP**, Vol. 7 No.7, pp.1084-1088, Jul. 1998.
[8] A. Gasteratos, I. Andreadis and Tsalides, "Fuzzy soft mathematical morphology", **IEEE Proc.-Vis Image Signal Processing**. Vol.145, No.1, pp.41-49, Feb. 1998.
[9] F. Russo and G. Ramponi, "Nonlinear fuzzy operators for image processing", Signal Processing 38(1994), pp.429-440.
[10] V. Chatizs and I. Pitas, "Nonlinear location and scale estimatiors of fuzzy numbers", **IEEE Trans. on SP**, Vol.46 No.1, pp.231-234, Jan. 1998.
[11] T. Lwe, H. Itoh and H. Seki, "Image filtering, edge detection, and edge tracing using fuzzy reasoning", **IEEE Trans. on PAMI**, Vol.18, No.5, pp.481-491, May 1996.
[12] C. C. Lee, "Fuzzy logical in control systems: fuzzy logical controller-Part". **IEEE Trans. on SMC**, Vol. 20 No. 2, pp.404-418, Mar. 1990.
[13] C. C. Lee, "Fuzzy logical in control systems: fuzzy logical controller-Part". **IEEE Trans. on SMC**, Vol. 20 No. 2, pp.419-435, Mar. 1990.
[14] Q. Xu et. al. "Fuzzy Weighted Average Filtering for Mixture Noises". International Conference on image and graphics 2004, Hong Kong, Dec. 2004.

The Binary Multi-SVM Voting System for Protein Subcellular Localization Prediction

Bo Jin[1], Yuchun Tang[1], Yan-Qing Zhang[1], Chung-Dar Lu[2], and Irene Weber[2]

[1] Department of Computer Science, Georgia State University, Atlanta, GA 30302,USA
[2] Department of Biology, Georgia State University, Atlanta, GA 30303,USA
{bojin, ytang}@gsu.edu, yzhang@cs.gsu.edu

Abstract. Support Vector Machine (SVM) as a learning system has been widely employed for pattern recognition and data classification tasks such as biological data classification. Choosing appropriate parameters are essential for SVM to achieve a high global performance. In this paper, we propose a new binary multi-SVM voting system without difficult parameter selection for protein subcellular localization prediction. The sufficient experimental results demonstrate that the multi-SVM voting system can achieve higher average prediction accuracies for the protein subcellular localization prediction than the traditional single-SVM system.

1 Introduction

Since Support Vector Machine (SVM) was first proposed by Vapnik [1]-[3], it has been widely employed for pattern recognition and data classification tasks, such as bioinformatics[4]. Based on the Structural Risk Minimization (SRM) theory, SVM generally has better generalization capabilities than other traditional Empirical Risk Minimization (ERM) based learning algorithms. In SRM [1,3,5,6], the concept of VC confidence $Conf$ is proposed, which decides the actual risk bound R, together with the empirical risk R_{emp} in machine learning.

$$R \le R_{emp} + Conf \tag{1}$$

A SRM based learning machine is expected to have the ability of finding an optimal tradeoff point between low capacity and low empirical risk. The capacity is measured using VC dimension, which is defined as the maximum number of points "shattered" by the machine. At the tradeoff point, the learning machine will work with the lowest actual risk. Fig.1 [3,5,6] shows the bound on the risk of a learning machine under the SRM theory.

In SVM learning, the machine is expected to construct a hyperplane to separate the training data with the largest margin. To solve linear non-separable and nonlinear separable problems, the regularization parameter C and a nonlinear kernel function

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 299 – 308, 2005.

with tunable parameters are introduced into SVM. With these selectable kernel functions and tunable parameters, SVM is more flexible to handle various problems. On the other hand, it becomes more difficult to select appropriate parameters to make SVM work at the lowest risk point, such as the point h^* in Fig.1.

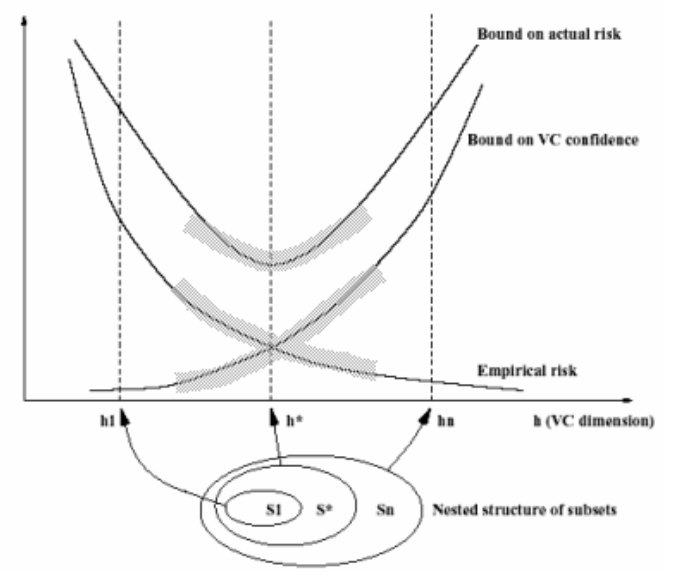


Fig. 1. Relationship among the actual risk, VC confidence and empirical risk

Some methods have been proposed to tuning SVM parameters. In [7], Chapelle et al. presented a gradient descent algorithm to minimize some estimates of the generalization error. In [8], Wang et al. used a fisher discrimination method to compute the optimal parameters and demonstrated some parameter range exists, within which the SVM generalization is stable. In [9], Duan et al. studied the VC dimension bound factor as a functional to tune the SVM parameters. In this paper, we present a binary Multi-SVM voting system. Instead of tuning parameters to minimize some error estimates, we sample multiple SVM classifiers with different capacities in a wide risk range (for example, the shadow ranges in Fig.1) and then use them to make a voting system for future classification tasks. Although this kind of voting system hardly reaches the optimal risk tradeoff point, it will work at a more stable and reliable risk point comparing to the system with only one SVM classifier.

The binary SVM voting system is used for the prediction of protein subcellular localization and the experimental results indicate that the system can achieve higher prediction accuracies comparing to the average accuracies achieved by single SVM systems.

The rest of the paper is organized as follows. Section 2 describes the basic SVM classification theories. In Section 3, the binary Multi-SVM voting system is presented. Section 4 shows the experiments and results. Finally, Section 5 gives conclusions.

2 SVM Theories

In this section, we review the basic SVM theories for data classification. For a binary classification problem, given a training data set

$$(\vec{x}_1, y_1), \cdots, (\vec{x}_l, y_l) \tag{2}$$

where $\vec{x}_i \in R^n, i = 1, \cdots, l$ are training example vectors with n-dimension and $y_i \in \{-1, +1\}, i = 1, \cdots, l$ is the class label associated with $\vec{x}_i$. SVM tries to find an optimal hyperplane,

$$\langle \vec{w}, \vec{x}_i \rangle + b = 0 \tag{3}$$

where $\vec{w} \in R^n$ is constructed using some training example vectors and $b \in R$. The hyperplane is unique, which separates the training examples with the maximum margin under the condition of

$$y_i(\langle \vec{w}, \vec{x}_i \rangle + b) \geq 1 \tag{4}$$

The decision function is

$$f(\vec{x}) = sign(\langle \vec{w}, \vec{x} \rangle + b) \tag{5}$$

In practice, the optimal hyperplane is calculated by solving the following constrained optimization problem,

$$\text{Minimize } \frac{1}{2} \| \vec{w} \|^2 + C \sum_i \xi_i \tag{6}$$

$$\text{Subject to } y_i(\langle \vec{w}, \vec{x}_i \rangle + b) \geq 1 - \xi_i \tag{7}$$

where ξ_i are nonnegative slack variables used to penalize training errors and C is the regularization parameter. The problem described by Eq. (6) and (7) is a QP problem, which can be transformed to

$$\text{Maximize } \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \langle \vec{x}_i, \vec{x}_j \rangle \tag{8}$$

$$\text{Subject to } \sum_i \alpha_i y_i = 0 \tag{9}$$

where $0 \leq \alpha_i \leq C, i, j = 1, \cdots, l$ and those $\vec{x}_i$ with $\alpha_i \neq 0$ are called support vectors. Generally kernel functions $K(\cdot, \cdot)$ are used instead of the inner products $\langle \cdot, \cdot \rangle$ in (8). The kernel function maps input data from original space R^n into a higher dimensional feature space H, where a linear classification decision is made. (8) and (5) are rewritten as

$$\text{Maximize } \sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j K(\vec{x}_i, \vec{x}_j) \tag{10}$$

$$f(\vec{x}) = \text{sgn}\left(\sum_i \alpha_i y_i K(\vec{x}, \vec{x}_i) + b \right) \tag{11}$$

The following are some nonlinear kernel functions commonly used in SVM,

$$\text{Polynomial function } K(\vec{x}, \vec{y}) = (\omega \vec{x} \bullet \vec{y} + 1)^d \tag{12}$$

$$\text{RBF } K(\vec{x}, \vec{y}) = \exp(-\gamma \| \vec{x} - \vec{y} \|^2) \tag{13}$$

3 Binary Multi-SVM Voting System

In this section, we present the binary multi-SVM voting system. Building the system involves selecting a risk range, sampling the SVM classifiers.

3.1 Risk Range Selection

In SVM, the capacity is measured using the VC dimension, which is bounded by the following inequality [3]

$$h \le \left(\min\left\lceil D^2 A^2 \right\rceil, n\right) + 1 \tag{14}$$

where D is the radius bound of the feature space ball, A is the bound of $\|w\|$ and n is the dimension of the feature space H. As suggested in [10], the longest support vector is used as the approximate value of D. If the VC dimension is too large, the actual risk will be high (see Fig.1). In the voting system, the risk range is evaluated by the ratio r of the training error Et to the VC dimension bound $\left(\min\left\lceil D^2 A^2 \right\rceil, n\right) + 1$. The number of misclassified training data is chosen as the training error.

$$r = \frac{Et}{\left(\min\left\lceil D^2 A^2 \right\rceil, n\right) + 1} \tag{15}$$

where $r \in [0, r_0]$ and r_0 is a linear SVM classifier's r.

3.2 SVM Classifiers Sampling

$[0, r_0]$ is evenly separated into many subranges α_i. To generate the classifier sample candidates the classifiers are trained with a series of parameters and values r are calculated. Sometimes the large parameter values may result in long-time training. In practice we restrict the RBF's γ chosen from the set $\{2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8 2^9, 2^{10}\}$, Polynomial function's d chosen from $\{1,2,3,4,5,6,7,8,9,10,11\}$ and C chosen from $\{I^B\}$, where

$I \in \{1,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9\}$ and $B \in \{-5,-4,\ldots,9,10\}$. During training if the classifier's r locates at a subrange α_i, where no other r locate at that time, the classifier will be chosen as a sample. The Multi-SVM voting system sampling procedure is described in Table 1. In the procedure, T represents the training set, i is the number of subrange of $[0, r_0]$, P is the kind of kernel functions and $m\ (\le i)$ is the number of samples. The returned set Ω involves all SVM classifier samples.

Table 1. Multi-SVM voting system Training Procedure

Multi-SVM-Sampling (T, i, P, m)
{
  Initialize r_0;
  Separate $[0, r_0]$ into i subranges α_i;
  Choose a kernel function;
  For each regularization parameter C
    For each kernel parameter P
      Train the SVM classifier (j) with C and P;
      Calculate $r = Et / h$;
      If (r belongs to α_i and no other r has been located at α_i)
      Then {
        Put the SVM classifier (j) into the sample set Ω;
        Increase count;
        }
      If (count >= m) **Then** Stop sampling;
    For End
  For End
  Return set Ω;
}

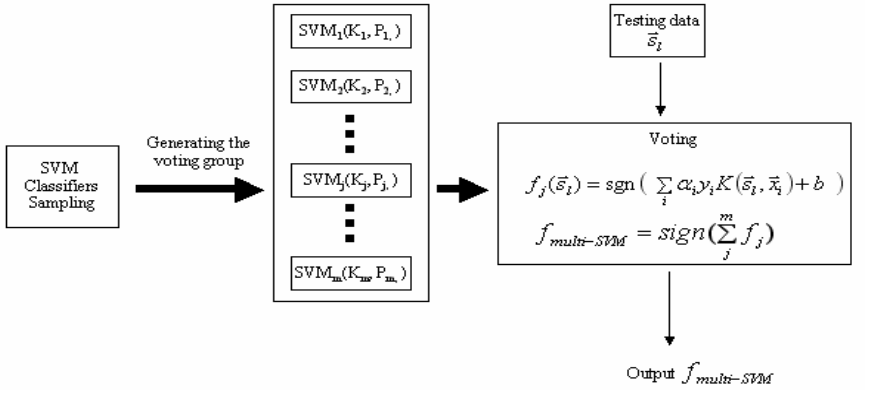


Fig. 2. The binary Multi-SVM voting system framework

3.3 Binary Multi-SVM Voting System Framework

The binary Multi-SVM voting system framework is described in Fig. 2. $SVM_j(K_j, P_j)$ are selected SVM classifiers during the sampling, where K_j is a kind of nonlinear kernel function, such as RBF or polynomial kernel and P_j is a group of kernel and regularization C parameters used by SVM_j.

4 Experiments for the Prediction of Protein Subcellular Localization

A protein's subcelluar location in a cell plays an important role in deciding the protein's function. The prediction of protein's subcellular localization may help us understand its functions and interactions with other molecules [11]. Some computational methods have been developed to solve the problem in recent years. Most of them can be classified into two categories [12]-[14]. One kind of prediction is based on amino acid composition [15]-[18]; the other is based on the individual sorting signals [19]-[20]. In [18], Hua and Sun developed a tree based SVM voting system and used a 5-fold cross validation to select SVM parameters. In Hua's system k SVMs were constructed for k-class classification. The difference between our system and Hua's system is that we use the multiple SVM classifiers voting to make one class classification each time. In the paper, we only consider the binary prediction based on amino acid composition.

4.1 Data Sets Processing and System Evaluation

We use two data sets collected by Reinhardt and Hubbard [15] to do the experiments. One is prokaryotic data set S_{pro}, which includes 997 prokaryotic sequences belonging to three location categories (Cytoplasmic set S_{pro1} 688**,** Periplasmic set S_{pro2} 202**,** Extracellular set S_{pro3} 107). The other is eukaryotic data set S_{eu}, which has 2427 eukaryotic sequences belonging to four categories (Nuclear set S_{eu1} 1097, Cytoplasmic set S_{eu2} 321, Mitochondrial set S_{eu3} 684, Extracellular set S_{eu4} 325). The numerical data are generated based on amino acid composition of each protein sequence. 5-fold cross validation method is used to make an evaluation in the experiments. Table 2 shows an evaluation procedure on the Prokaryotic Cytoplasmic set S_{pro1}. Other sets are trained and tested in a similar way. The voting system is built over the Joachims' SVM^{light} [10] software.

4.2 Experimental Results

The experimental results are analyzed and evaluated based on the confusion matrix table (Table 3). To make compare the multi-SVM voting system with basic SVM system, we evaluate the system with RBF kernel and polynomial kernel separately. The results of evaluation on prokaryotic data sets are listed in Table 4 and 5. From the

Table 2. 5-Fold cross Validation Evaluation on the Prokaryotic Cytoplasmic Category S_{pro1}

[Procedure]
Label positive on the data of S_{pro1};
Label negative on the data of S_{pro2} and S_{pro3};
For $j = 1,\cdots,5$
Create the training set $T_j = S_{pro} - S^j_{pro1} - S^j_{pro2} - S^j_{pro3}$;
Create the testing set $V_j = S^j_{pro1} + S^j_{pro2} + S^j_{pro3}$;
Ω' =**Multi-SVM-Training** $(T_j, 500, RBF, 121)$ **;**
For each item v in V_j
Predict the item v using each samples in Ω';
The member of Ω' take a vote and judge if v belongs S_{pro1};
For end
For end

tables, it can been seen most of the prediction accuracies of the multi-SVM voting system using RBF and polynomial kernel are better than the average accuracies of121 single SVM classifiers. In each 5-fold average case, multi-SVM voting system is always better than 121 single SVM classifiers. Using the RBF kernel, in the total average evaluation on prokaryotic set , the accuracy of the multi-SVM voting system is better than those of 121 single SVM classifiers by 2.12% on OA, 6.31% on Precision and 5% on Recall. Using the polynomial kernel, in the total average evaluation on prokaryotic set, the accuracy of the multi-SVM voting system is better than those of 121 single SVM classifiers by 1.26% on OA, 5.0% on Precision and 6.51% on Recall. The multi-SVM voting system using RBF kernel is a little bit better than that using polynomial kernel. The prediction performance of the system does not reach that of the best one of 121 single SVM classifiers, but it really better than the 121 single classifiers in average case. It looks like the new system performance locate in middle of the average 121 single classifiers' performance and the best single one's.

The results of evaluation on eukaryotic data sets are listed in Table 6 and 7. From the tables, We can see the prediction accuracies of the multi-SVM voting system using RBF and polynomial kernel are also better than the average accuracies of 121 single SVM classifiers in most cases. In each 5-fold average case, the recall of multi-SVM voting system is lower than that of 121 single SVM classifiers at Mitochond location. But for the total average evaluation of eukaryotic set, the multi-SVM voting system is better than the 121 single SVM classifiers in the average case by 1.7% on OA, 11% on Precision and 1.8% on Recall, when using the polynomial kernel. If using polynomial kernel, the improvements are about 0.6% on OA, 16% on Precision and 1% on Recall in the total average evaluation of eukaryotic set. It also looks like the new system performance locate in middle of the average 121 single classifiers' performance and the best single one's.

Table 3. Confusion Matrix in Classification

	Belonging to category S	Not belonging to category S
Classified as category S	TP (True positive)	FP (False positive)
Classified as non-category S	FN (False negative)	TN (True negative)

Overall Accuracy (OA) = $(TP + TN)/(TP + FP + FN + TN)$

Precision (Pre.)= $TP/(TP + FP)$

Recall (Rec.)= $TP/(TP + FN)$.

Table 4. Evaluation results on prokaryotic data set S_{pro} using RBF (%)

	Average Prediction of 121 single SVM classifiers			The best prediction among 121 single SVM classifiers			Prediction of Multi-SVM voting system		
	OA	Pre.	Rec.	OA	Pre.	Rec.	OA	Pre.	Rec.
S_{pro1}	91.27	91.06	97.27	95.57	95.54	100	93.87	93.3	98.09
S_{pro2}	94.83	84.98	63.10	96.98	100	83.15	95.88	90.97	69.14
S_{pro3}	89.04	83.36	54.56	93.86	100	80.59	91.75	94.08	62.66
Ave.	91.71	86.47	71.64	95.47	98.51	87.91	93.83	92.78	76.63

Table 5. Evaluation results on prokaryotic data set S_{pro} using polynomial kernel (%)

	Average Prediction of 121 single SVM classifiers			The best prediction among 121 single SVM classifiers			Prediction of Multi-SVM voting system		
	OA	Pre.	Rec.	OA	Pre.	Rec.	OA	Pre.	Rec.
S_{pro1}	91.64	91.51	97.06	94.27	94.32	100	92.46	92.35	97.09
S_{pro2}	94.68	80.23	63.48	96.68	98.66	82.25	95.57	84.76	73.05
S_{pro3}	87.28	75.09	47.12	92.36	96.63	75.25	89.34	84.59	57.05
Ave.	91.20	82.28	69.22	94.43	96.54	85.83	92.46	87.23	75.73

Table 6. Evaluation results on eukaryotic data set S_{eu} using RBF kenrel (%)

	Average Prediction of 121 single SVM classifiers			The best prediction among 121 single SVM classifiers			Prediction of Multi-SVM voting system		
	OA	Pre.	Rec.	OA	Pre.	Rec.	OA	Pre.	Rec.
S_{eu1}	92.39	82.56	50.25	96.24	100	81.88	94.22	97.87	58.79
S_{eu2}	88.64	69.10	27.69	91.46	100	61.65	89.28	89.79	21.36
S_{eu3}	82.91	74.78	59.53	87.83	89.64	74.66	85.23	79.95	63.49
S_{eu4}	85.87	85.32	83.17	89.15	89.72	92.56	87.95	88.65	84.10
Ave.	87.45	77.94	55.16	91.17	94.84	77.69	89.17	89.06	56.93

Table 7. Evaluation results on eukarytic data set S_{eu} using polynomial kernel (%)

	Average Prediction of 121 single SVM classifiers			The best prediction among 121 single SVM classifiers			Prediction of Multi-SVM voting system		
	OA	Pre.	Rec.	OA	Pre.	Rec.	OA	Pre.	Rec.
S_{eu1}	90.08	70.25	31.30	93.81	97.38	68.3	90.68	91.02	33.51
S_{eu2}	87.43	60.28	7.85	88.99	100	40.84	87.50	100	5.55
S_{eu3}	80.38	70.72	49.52	84.08	85.02	68.26	81.48	74.02	53.17
S_{eu4}	84.38	86.15	77.81	87.46	89.19	83.76	84.78	86.47	78.5
Ave.	85.57	71.85	41.62	88.58	92.90	65.29	86.11	87.87	42.68

5 Conclusion

In this paper, we proposed a binary Multi-SVM voting system and use it to predict the protein subcellular localization. The experimental results of protein's subcelluar localization prediction demonstrate the method of multiple SVM classifiers voting within a wide risk range can work at a more stable and reliable risk point and give better generalization comparing to the system with only one SVM classifier system in the average case.

Acknowledgment

This work is supported in part by NIH under Grant P20 GM065762.

References

1. V. Vapnik, *The Nature of Statistical Learning Theory*, New York, Springer-Verlag, 1995.
2. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
3. V. Vapnik, *Statistical Learning Theory*, New York, John Wiley and Sons, 1998.
4. Noble, William Stafford, "Support vector machine applications in computational biology," Kernel Methods in Computational Biology, B. Schoelkopf, K. Tsuda and J.-P. Vert, ed., MIT Press, 2004.
5. Smith, N. D, "Support vector machines applied to speech pattern classification", Master's thesis, Cambridge University, Dept. of Engineering, 1998.
6. K.K. Chin, "A Support Vector Machines applied to Speech Pattern Classification," Dissertation, Cambridge University, Dept. of Engineering, 1999.
7. O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing Multiple Parameters for Support Vector Machines," *Machine Learning*, vol. 46, pp. 131-159, 2002.
8. W.J. Wang, Z.B. Xu, W.Z. Lu, and X.Y. Zhang, "Determination of the spread parameter in the Gaussian kernel for classification and regression," *NeuroComputing*, vol. 55, pp. 643-663, 2003.
9. K. Duan, S.S. Keerthi and A.N. Poo, "Evaluation of simple performance measures for tuning SVM hyperparameters", *Neurocomputing,* Vol. 51, 2003, pp. 41-59.
10. T. Joachims, "Making Large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning," MIT Press, 1999.
11. Rong She, "Association-Rule-Based Prediction of Outer Membrane Proteins," Thesis, Simon Fraser University, 2003.
12. Chou, K.C., "Review: Prediction of Protein Structural Classes and Subcellular Location," *Curr. Protein Peptide Sci*., 1, 171-208, 2000.
13. K. Nakai, "Protein sorting signals and prediction of subcellular localization," *Adv. Protein Chem.*, 54, pp. 277-344, 2000.
14. Ying Huang and Yanda Li, "Prediction of protein subcellular locations using fuzzy k-NN method", *Bioinformatics*, Vol. 20 no. 1, pp. 21-28,2004.
15. Reinhardt, A. and Hubbard, T. "Using neural networks for prediction of the subcellular location of proteins", *Nucleaic Acid Research*, Vol. 26, No. 9, pp. 2230-2236, 1998.
16. Chou, K.C. and Elrod, D., "Protein subcellular location prediction," *Protein Eng*., Vol. 12, No. 2, pp. 107-118, 1999.
17. Yuan, Z., "Prediction of protein subcellular locations using Markov chain models," *FEBS Lett*., Vol. 451, No. 1, pp. 23-26, 1999.
18. Hua, S.J. and Sun, Z.R., "Support Vector Machine Approach for Protein Subcellular Localization Prediction*," Bioinformatics*, vol. 17, No. 8, pp. 721-728, 2001.
19. Nakai, K. and Kanehisa, M, "Expert system for predicting protein localization sites in Gram-negative bacteria," *Proteins Struct. Funct. Genet.,* 11, 95-110, 1991.
20. Nielsen, H., Brunak, S. and von Heijine, G., "Machine learning approaches for the prediction of signal peptides and other protein sorting signals," *Protein Eng.,* 12, pp. 3-9, 1999.

Gene Network Prediction from Microarray Data by Association Rule and Dynamic Bayesian Network

Hei-Chia Wang* and Yi-Shiun Lee

Institute of Information Management, National Cheng Kung University,
Tainan, 701, Taiwan
{hcwang, r7692102}@mail.ncku.edu.tw

Abstract. Using microarray technology to predict gene function has become important in research. However, microarray data are complicated and require a powerful systematic method to handle these data. Many scholars use clustering algorithms to analyze microarray data, but these algorithms can find only the same expression mode, not the transcriptional relation between genes. Moreover, most traditional approaches involve all-against-all comparisons that are time consuming. To reduce the comparison time and find more relations, a proposed method is to use an a priori algorithm to filter possible related genes first, which can reduce number of candidate genes, and then apply a dynamic Bayesian network to find the gene's interaction. Unlike the previous techniques, this method not only reduces the comparison complexity but also reveals more mutual interaction among genes.

Keywords: dynamic Bayesian network, apriori algorithm, association rule, gene network.

1 Introduction

The Human Genome Project (HGP) has been in operation for a decade. The project originally planned to be completed in 15 years, but rapid technological advances accelerated its completion to 2003. The HGP's ultimate goal was to identify all the functions of the 30,000 human genes and render them accessible for further biological study. After the announcement that all human genes had been sequenced, the next stage was to investigate functional genomics (Hieter and Boguski, 1997; Eisenberg *et al.,* 2000). To achieve this purpose requires the application of a powerful technique. Microarray technology is one of them.

The development of microarray technology has been phenomenal in the past few years. It has become an important tool in many genomics research laboratories because it has revolutionized the approach to finding the relations among genes. Instead of working on a gene-by-gene basis, scientists can use microarray to study tens of thousands of genes at once. Many benefits result from this technology. One is determining how the expression of any particular gene might affect the expression of other genes. Another is determining the gene network. Another is determining what genes are expressed as a result of certain cellular conditions, for example, those expressed in diseased cells that are not expressed in healthy cells.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 309 – 317, 2005.

Currently, the most widely used algorithm for analyzing microarray data is hierarchical clustering (Eisen *et al.*, 1998; Tavazoie *et al.*, 1999; Tamayo *et al.*, 1999). Although this method has been successful in showing that genes participating in the same biological processes have the same expression profiles, several problems remain. First, genes that are biologically related often are not related in their expression profiles and hence will not be clustered together (Torgeir *et al.,* 2002;Creighton *et al.,* 2003). Second, clustering genes into disjoint clusters will not capture the fact that many gene products participate in more than one biological process (Becquet *et al.,* 2003;Creighton *et al.,* 2003). Third, no relationship can be inferred between the different members of a group (Becquet *et al.,* 2003). Finally, the all-against-all comparison in clustering is time consuming.

Here we describe our attempt to solve the above problems with an apriori algorithm to filter possible related genes first, which can reduce number of candidate genes, and then application of a dynamic Bayesian network to find the gene's interaction. First, an association rule algorithm is applied to filter unimportant genes. Since the microarray data we used was limited and the apriori algorithm requires lot of data, we thought variation over time may reveal some association of genes. Therefore, we used an across time-finding algorithm to expand the microarray data. When the association rules were confirmed, genes could be listed under more than one rule. This means that a gene may be involved in more than one syn-expression group. Second, after the candidate gene associations were selected, a dynamic Bayesian network (DBN) was applied to construct a gene network for biologists. However the use of DBN is time consuming, irrelevant genes should be filtered to reduce the impact (Zou *et al.,* 2004). In this research, we use an apriori algorithm in first step to filter irrelevant genes to reduce the impact and find extra information.

2 Background

Two important concepts, association rule by a priori algorithm and DBN, are used in this model. We explain them briefly in the following sections.

2.1 Association Rule

An association rule is presented in the form $LHS \Rightarrow RHS$, where *LHS* and *RHS* both are "itemsets." Itemsets can be defined in terms of "transactions," which in the retail industry refers to customer transactions (a customer purchases one or more items at the checkout counter in a single transaction). Here we adopt the definition for itemsets and association rules from Doddi *et al.* (2001) to implement our program:

DEFINITION 1:

1. Given a set S of items, any nonempty subset of S is called an itemset.
2. Given an itemset I and a set T of transactions, the "support" of I with respect to T , denoted by support$_T(I)$, is the number of transactions in T that contain all the items in I.

3. Given an itemset I, a set T of transactions and a positive integer α, I is a "frequent itemset" with respect to T and α if support$T(I) > \alpha$. We refer to α as the "minimum support."

DEFINITION 2:

1. An "association rule" is a pair of disjointed itemsets. If *LHS* and *RHS* denote the two disjoint itemsets, the association rule is written as $LHS \Rightarrow RHS$.
2. The "support" of the association rule $LHS \Rightarrow RHS$ with respect to a transaction set T is the support of the itemset $LHS \cup RHS$ with respect to T.
3. The "confidence" of the rule $LHS \Rightarrow RHS$ with respect to a transaction set T is the ratio of support ($LHS \cup RHS$) to support (LHS).

2.2 Dynamic Bayesian Network

DBNs can be viewed as an extension of Bayesian networks (BNs). In contrast to BNs, which are based on static data, DBNs use time series data for constructing causal relationships among random variables. Suppose that we have n microarrays and each microarray measures expression levels of p genes. The microarray data can present $n * p$ matrix. $X = (x_1, x_2, ...x_n)^T$ whose i'th row vector $x_i = (x_{i1}, x_{i2}, ...x_{ip})^T$ corresponds to a gene expression level vector measured at time t. The DBN model assumes that each gene's responses depend only on the status of the previous time unit. In other words, the status vector of time i (in Figure 1) depends only on that of time i (in figure 1) depends only on that of time i-1 (Kim et al., 2003, 2004; Murphy and Mian, 1999)

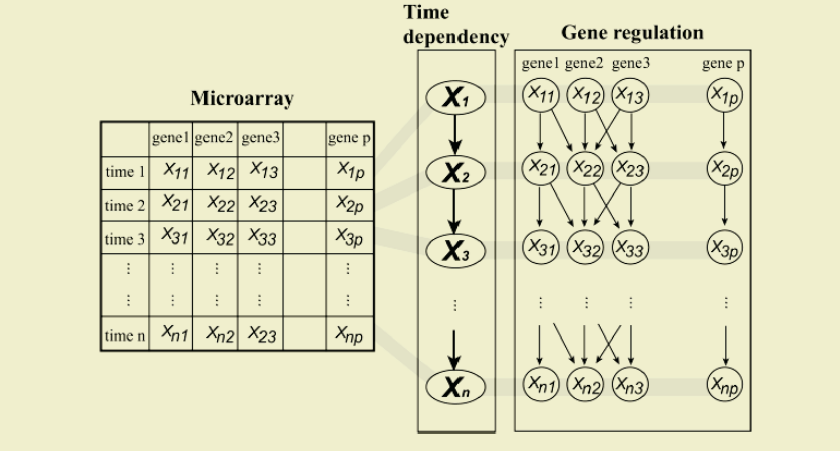


Fig. 1. DBN model

Based on this assumption, the joint probability can be decomposed as equation (1):

$$P(X_{11},....,X_{np}) = P(X_1)P(X_2 \mid X_1)\times...\times P(X_n \mid X_{n\text{-}1}) \qquad (1)$$

where $X_i = (x_{i1},...,x_{ip})^T$ is a p-dimensionalrandom variable

According to the time dependency, only forward edges (i.e., edges from time *i-1* to *i*) are allowed in the network, seen in Figure 2.

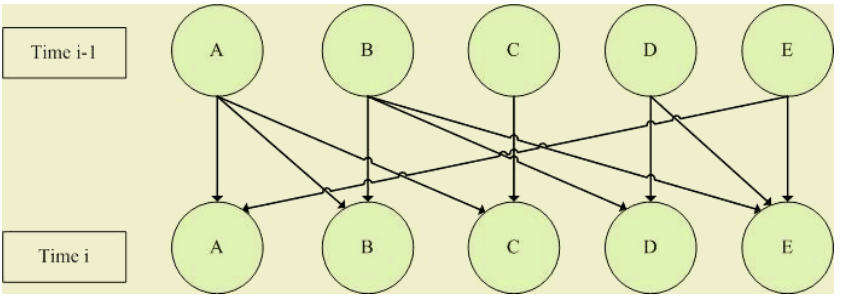


Fig. 2. DBN interaction diagram

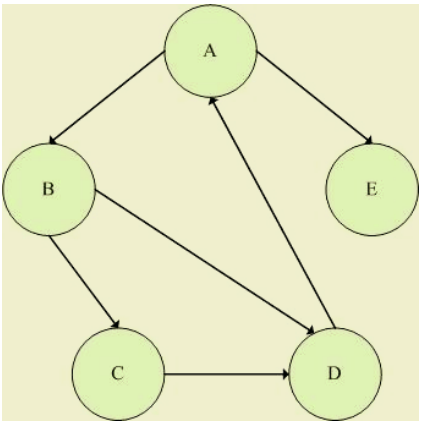


Fig. 3. DBN used to solve cycles

With DBN, the interaction between genes can be divided into 2 slices (shown as Figure 3). Hence, use of DBN can solve the cycle situation that BN cannot (Kim et al., 2003, 2004). From the structure in Figure 3, $P(X_i \mid X_{i-1})$ can be decomposed to:

$$P(X_i \mid X_{i\text{-}1}) = P(X_{i1} \mid P_{i\text{-}1,1}) \times \ldots \times P(X_{ip} \mid P_{i\text{-}1,p}) \quad (2)$$

where $P_{i\text{-}1,j} = (P^{(j)}_{i\text{-}1,1}, \ldots P^{(j)}_{i\text{-}1,q_j})^T$

is a random variable vector of parent genes of jth geneat time i -1

We can then combine (1) and (2) for the DBN model as follows:

$$f(x_{11}, \ldots, x_{np}) = \prod_{i=1}^{n} \prod_{j=1}^{p} g_j(x_{ij} \mid p_{i-1,j})$$

where $p_{oj} = \varnothing$

3 Method

3.1 Research Model

In this research, candidate data from a microarray is considered as input data. The output information is the gene relations represented in a network chart. By these relations, reactions of catalysis or restraint between genes can be obtained to help biologists find relations among genes.

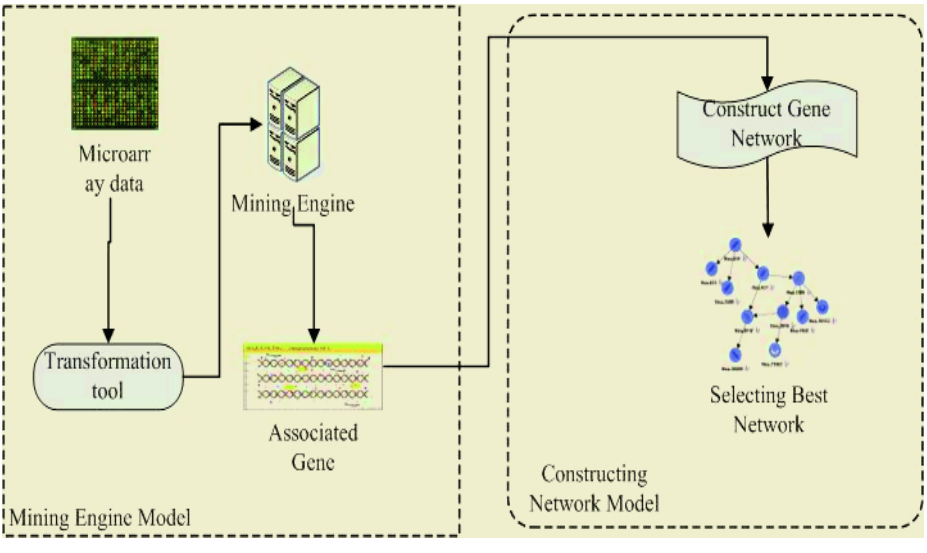


Fig. 4. System infrastructures

3.2 The Mining Engine Module

Currently, most microarray data are analyzed by a clustering algorithm, which can find only the same model and may cause some information loss (i.e., transcription factor loss). To find this information, we adopt time series microarray data to make the variation between 2 slices more meaningful. We first expand the microarray data to find variation between 2 slices, then we use an a priori algorithm to analyze the data to find the complete relations among genes. Table 1 shows microarray data calculated by cy5/cy3.

Table 1. Microarray data

	A	B	C	F
1	2.5	1.5	1.1	1.3
2	1.8	2.1	0.5	2.2
3	1.3	2.7	1.0	2.8
4	0.9	2.6	0.4	2.1

Table 2. Microarray transporting data array

	A	B	C	F
1-2	***D***	***U***	***D***	***U***
1-3	***D***	***U***	***N***	***U***
1-4	***D***	***U***	***D***	***U***
2-3	***N***	***U***	***U***	***U***
2-4	***D***	***U***	***N***	***U***
3-4	***D***	***N***	***D***	***U***

We then use the following steps to find rules by different time slices.

1. Produce all possible situations from different time slices.
2. Using the data in Table 1 and obtain the cy5 front value and background value for equation (1) to calculate the expression value to find the trend between the time slices. The results are shown in Table 2.
3. Use the results to run an apriori algorithm to determine possible associations.

$$Expression\_Value = \log_2 \left[\frac{\left(Cy5_i(F) - Cy5_i(B)\right) - \left(Cy5_j(F) - Cy5_j(B)\right)}{\left(Cy3_i(F) - Cy3_i(B)\right) - \left(Cy3_j(F) - Cy3_j(B)\right)} \right] \quad (4)$$

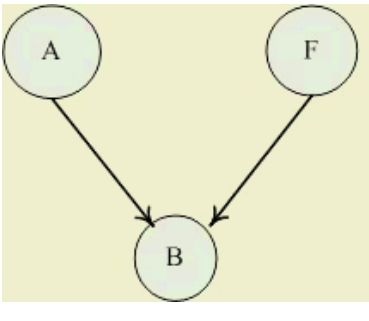
After step 2, 3 possible levels could be found : "***U***" for "expression", "***N***" for "no reaction" and "***D***" for "deterrence." They are judged by the rule: $Value = Log(Cy5/Cy3)$. When the expression value is **larger than 2**, the data will be transferred to "***U***" as the expression status. When the value is **between 1 and 2**, the data will be transferred to "***N***", which means that the gene has no significant reaction in this process. When the value is **less than 1**, the data will be transferred to "***D***", which means that the gene is in deterrence status at this time slice.

After this step, the apriori algorithm is applied to find the possible rules. In this example, we find that:

$$\text{If } (A = \boldsymbol{D} \text{ and } F = \boldsymbol{U}) \text{ then } B = \boldsymbol{U}$$

3.3 Constructing the Gene Network

After obtaining the rules, DBN is then used to build a gene network. We select all genes in all rules and ignore the rest. In a previous example, genes A, B, and F will be selected to go through the DBN. Figure 5 shows a possible result. In this example, suppose the probability of node B is as shown in Figure 5, which is calculated by the DBN model. In this figure, we use the sign "+" and "-" to present status. For example, "–F"shows that the value of F is ***D***. +A shows that the value of A is ***U***. It can be observed that B is ***U*** when A is ***D*** and F is ***U*** .So, genes A and F are 2 the factors that will cause the gene B reaction.



	+A,−F	+A,-F	-A,+F	-A,-F
B	0.4	0.6	0.9	0.5
-B	0.6	0.4	0.1	0.5

Fig. 5. Network topology

4 Conclusion

This paper presents a way to use gene association rules to determine 2 types of gene interactions and to generate a gene network of microarray data. In this method, we

find the relational gene first by apriori algorithm and then adopt a DBN to find the gene network. Currently, we are using the microarray data from TaiwanOrchid (http://www.taiwanorchid.net) research team and evaluate the performance of our method. This method can find not only positive relations but also negative relations. Although it can find extra relations that cannot be found by similar methods, the method could be improved upon.

1. The DBN still takes a lot of time to build, so we try to verify the rules before the DBN step.
2. Some extra resources, such as annotations or literature, which contain biology meaning should be used to refine the gene network.

References

Becquet, C. Blachon, S., Jeudy, B., Boulicaut, J.F. and Gandrillon, O. (2003). Strong-association-rule mining for large-scale gene-expression data analysis: a case study o human SAGE data. *Genome Biol,* 12, 1-16.

Creighton,C. and Hanash,S. (2003). Mining gene expression databases for association rules. Bioinformatics, 19, 79-86.

Doddi, S., Marathe, A., Ravi, S.S. and Torney, D.C. (2001). Discovery of association rules in medical data. *Med Information Internet Med.*, **26**, 25–33.

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Science (USA?),* 95, 14863-14868.

Eisenberg, D., Marcotte, M.,E., Xenarios, I. and Yeates, O.,T. (2000). Protein function in the post-genomic era. *Nature*, 405, 823-826.

Ewing, B. and Green, P. (2000). Analysis of expressed sequence tags indicates 35,000 human genes. *Nature Genet,* 25, 232-234.

Hieter, P. and Boguski, M. (1997). Functional genomics: it's all how you read it. *Science*, 278, 601-602.

Kim, S.Y., Imoto, S. and Miyano, S. (2003). Inferring gene networks from time series microarray data using Dynamic Bayesian Networks. *Briefing Bioinformatics*, 4(3), 228-235.

Kim, S.Y., Imoto, S. and Miyano, S. (2004) Dynamic Bayesian networks and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. *Biosystems,* 75, 57-65.

Murphy, K. and Mian, S. (1999) Modeling gene expression data using dynamic Bayesian networks. Technical Report, Computer Science Division, University of California, Berkeley, CA.

Tamayo,P., Slonim,D., Mesirov,J., Zhu,Q., Kitareewan,S., Dmitrovsky,E., Lander,E.S. and Golub,T.R. (1999). Interpreting patterns of gene expression with self-organizing maps methods and application to hematopoietic differentiation. Nature Genetics, 96, 2907–2912.

Tavazoie,S., Hughes,J.D., Campbell,M.J., Cho,R.J. and Church,G.M. (1999). Systematic determination of genetic network architecture. Nature Genetics, 22, 281-285.

Torgeir, R.H., Astrid, L. and Jan, K. (2002). Learning rule-based models of biological process from gene expression time profiles using Gene Ontology. *Bioinformatics,* 19, 1116-1123.

Zou, M. and Conzen, S.D. (2004) A new dynamic Bayesian network approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics,* Advance Access published on August 12, 2004, 1-29.

Protein Interaction Prediction Using Inferred Domain Interactions and Biologically-Significant Negative Dataset

Xiao-Li Li, Soon-Heng Tan, and See-Kiong Ng

Institute For Infocomm Research,
21 Heng Mui Keng Terrace, Singapore 119613
{xlli, soonheng, skng}@i2r.a-star.edu.sg

Abstract. Protein domains are evolutionarily-conserved structural or functional subunits in proteins that are suggestive of the proteins' propensity to interact or form a stable complex. In this paper, we propose a novel domain-based probabilistic classification method to predict protein-protein interactions. Our method learns the interacting probabilities of domain pairs based on domain pairing information derived from both experimentally-determined interacting protein pairs and carefully-chosen non-interacting protein pairs. Unlike conventional approaches that use random pairing to generate artificial non-interacting protein pairs as negative training data, we generate biologically meaningful non-interacting protein pairs based on the proteins' biological information. Such careful generation of negative training data set is shown to result in a more accurate classifier. Our classifier predicts potential interaction between any pair of proteins based on the probabilistically inferred domain interactions. Comparative results showed that our probabilistic approach is effective and outperforms other domain-based techniques for protein interaction prediction.

1 Introduction

Cellular processes are biochemical events that are typically achieved by the interactions of proteins with one another. The elucidation of protein interactions is therefore the necessary first-step for understanding the biology of cellular processes. Many experimental methods have been developed to detect protein-protein interactions, however, none of the current experimental methods is adequate to interrogate the entire interactome [11, 15]. It is therefore useful to develop complementary computational methods for predicting new protein-protein interactions.

Several computational techniques have been proposed to predict protein-protein interactions. For example, potential protein interactions can be derived from gene context analysis such as gene neighborhood [3, 13], gene fusion [5, 9], and gene co-occurrences and phylogenetic profiles [8, 14]. Alternatively, the physiochemical properties or tertiary structure of proteins can also be used for predicting interactions[1, 10].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 318–326, 2005.

Recently, however, there is an increased focus on using *protein domains* to predict protein-protein interactions [4, 6, 7, 12, 16]. Protein domains are evolutionarily-conserved structural or functional subunits in proteins found across different proteins. They are often found to participate in intermolecular interactions with one another. The existence of certain domains in proteins can therefore suggest the possibility of interaction between two proteins. As such, the analysis of many protein-protein interactions can be reduced to understanding the underlying domain-domain interactions between two proteins.

Domain-based protein interaction prediction methods generally consist of two main steps: 1) inferring domain-domain interactions from known protein interactions, 2) predicting protein interactions based on the inferred domain-domain interaction information. A few domain-based interaction detection techniques have recently been proposed. Deng *et al.* described a Maximum Likelihood estimation technique to infer domain-domain interactions that was then used to predict protein interactions [4]. Wan *at al.* presented a alternative statistical scoring system as a measure of the interaction probability between domains [16]. Ng *et al.* devised an integrative approach to infer the protein domain interactions [12] from other data sources in addition to experimentally determined protein interactions. Han *et al.* designed a probabilistic framework that takes domain combinations instead of single domains as basic units of protein interactions [6, 7].

These proposed techniques can be grouped into two main paradigms in terms of the way they infer domain-domain interactions. The domain interactions that are used for predicting protein-protein interactions are learned either (1) from an interacting protein set or positive class only [4, 12, 16], or (2) from both an interacting protein set and an artificially generated non-interacting protein set as negative set; the latter being generated by randomly pairing the proteins [6, 7]. In the case of (1) where learning is conducted only from an interacting protein set, many false positive domain pairs may be derived because these domain pairs may occur in the (unavailable) negative set with high frequency. In the case of (2), the use of a putative negative data set helps alleviate this problem. However, using artificially generated non-interacting protein set as negative set is inadequate for inferring domain-domain interactions because the randomly generated negative dataset may contain interacting protein pairs. In addition, if the artificially generated negative dataset is subsequently used in evaluating the performance of classifier, it will lead to inaccurate computation of the actual sensitivity and specificity of the technique.

In this paper, we propose a novel probabilistic technique to infer domain-domain interactions using both positive and negative training datasets. Our probabilistic model was able to outperform other domain-based techniques in predicting potential protein interactions. Unlike conventional approaches that use random pairing to generate artificial non-interacting protein pairs as negative training data, we generate biologically meaningful non-interacting protein pairs based on the proteins' biological information, namely, proteins are most unlikely to interact if they are from different cellular locations and functional categories.

We showed that the performance of classifier is improved with the more confident negative dataset.

2 Methods

Our proposed approach classifies a protein pair to be either interacting or non-interacting based on inferred underlying domain-domain interactions. The approach consists of three steps as follows: 1) generate the negative set N (non-interacting protein pairs); 2) infer domain-domain interactions based on the interacting proteins pair set I and the negative set N; 3) build a classifier based on the interacting probabilities of domain pairs. Below, we present the methods for these three steps in turn.

Generate the Negative Set: Proteins are most unlikely to interact if they are from different cellular locations and functional categories. So our generated negative set only pairs those proteins located at different locations and with different functions. Algorithm 1 shows how to generate non-interacting protein pairs (negative set).

Algorithm 1 Generate non-interacting protein pairs

1: **Input**: interacting set I, protein set P;
2: **Output**: negative set N;
3: BEGIN
4: Set $N = \emptyset$;
5: **for** all the protein $p_i \in P$ **do**
6: Search p_i's locations (l) and functional categories (c);
7: **end for**
8: Combine all protein pairs into a set PS: $PS = \{(p_i, p_j) | p_i \in P, p_j \in P, i \neq j\}$;
9: **repeat**
10: **for** each protein pair $(p_i, p_j) \in PS$ **do**
11: **if** $(p_i, p_j) \notin I$ **then**
12: **if** $((p_i.l \neq p_j.l) \wedge (p_i.c \neq p_j.c))$ **then**
13: $N = N \cup \{(p_i, p_j)\}$;
14: **end if**
15: **end if**
16: $PS = PS - \{(p_i, p_j)\}$;
17: **end for**
18: **until** $(PS = \emptyset)$
19: END

In Algorithm 1, for each protein in P, the set of proteins of interest, we retrieve the biological information about its locations and functional categories (Steps 5-6) from the MIPS database[1]. Then, from Step 9 to Step 18, we check

[1] http://mips.gsf.de/genre/proj/yeast/index.jsp

each protein pair (p_i, p_j) in protein pair set PS: if it is already in the interacting protein set I , we eliminate it from PS; otherwise, if p_i and p_j are located at different cellular locations and from different functional categories, we add them into negative set N.

Note that in Step 12, a protein (p_i or p_j) may be located at multiple locations and has multiple functions. We consider (p_i, p_j) to be non-interacting only if none of p_i's locations and functions match the p_j's locations and functions. In addition, because the proteins' functional classifications given in MIPS are hierarchical, we will only regard two proteins to have different functions at the highest possible level of the MIPS functional hierarchy(Level 1). Such strict selection strategy helps us get a much purer negative set for training our classifier.

Infer Domain-Domain Interactions: The objective of this next step is to assign interaction probabilities to each domain pair based on its occurrence in the protein-protein interacting set I and the negative set N. For a protein pair $(p_i, p_j) \in I$, we infer that domain $d_{i,r}$ potentially interacts with domain $d_{j,s}$ with a probability of $1/(|p_i| * |p_j|)$, where $|p_i|$ and $|p_j|$ are the number of domains in proteins p_i and p_j respectively; $d_{i,r}$ and $d_{j,s}$ are the r-th and s-th domains of proteins p_i and p_j respectively.

Given that a domain pair (d_x, d_y) may occur in many interacting protein pairs of I, the interacting frequency of (d_x, d_y) in I is defined as:

$$N((d_x, d_y), I) = \sum_{i=1}^{|I|} \lambda_i(d_x, d_y) * \frac{1}{|p_x^i| * |p_y^i|} \tag{1}$$

where (p_x^i, p_y^i) is the i-th protein pair in I and $\lambda_i(d_x, d_y)$ is the total number of occurrences of the domain pair (d_x, d_y) in (p_x^i, p_y^i). We compute $N((d_x, d_y), N)$, the interacting frequency of (d_x, d_y) in N, in a similar way:

$$N((d_x, d_y), N) = \sum_{i=1}^{|N|} \lambda_i(d_x, d_y) * \frac{1}{|p_x^i| * |p_y^i|} \tag{2}$$

Let a set of pre-defined classes be $C = \{I,\ N\}$ and all the domain pairs set be DP. For any domain pair $(d_x, d_y) \in DP$, their interacting probability $P((d_x, d_y)|c_e)$, with Laplacian smoothing and $c_e \in C$, is defined as:

$$P((d_x, d_y)|c_e) = \frac{1 + N((d_x, d_y), c_e)}{|DP| + \sum_{k=1}^{|C|} N((d_x, d_y), c_e)} \tag{3}$$

For a domain pair (d_x, d_y), the greater the interacting probability $P((d_x, d_y)|I)$, the more frequent it occurs in the interacting set I. However, since such a domain pair may also be chanced occurrences in class I, it is necessary to check its interacting probability in N: $P((d_x, d_y)|N)$. Obviously, if $P((d_x, d_y)|I)$ is significantly larger than $P((d_x, d_y)|N)$, then the domain pair (d_x, d_y) is likely to be a genuine domain-domain interaction. Otherwise, if $P((d_x, d_y)|N)$ is similar or even bigger than $P((d_x, d_y)|I)$, then the domain pair is unlikely to be inter-

acting. In other words, to check if a domain pair (d_x, d_y) interacts, we compute its interacting probabilities in both interacting set I and negative set N.

Note that the purity of N can affect the accuracy of inferred domain-domain interactions. If N were generated from randomly paired proteins, the false negative protein pairs in N will result in the inference of many domain pairs that should have occurred only in interacting protein set (i.e. positive class). This will result in assigning inaccurate interacting probabilities to domain pairs and subsequently affect the accuracy of the eventual classifier to infer protein interactions.

Build a Protein Interaction Classifier: Given a protein pair (p_i, p_j), in order to perform classification (i.e. to judge whether the proteins may interact with each other or not), we compute the posterior probability $P(c_e|(p_i, p_j)), c_e \in C$. The prior probability $P(c_e)$ of class c_e is defined as:

$$P(c_e) = \frac{\sum p(c_e, (p_i, p_j)), (p_i, p_j) \in I \cup U}{|I| + |N|} \tag{4}$$

Based on Equations (4) and (3), our proposed technique uses the joint probabilities of domain pairs and classes to estimate the probabilities of classes given a protein pair. Our classifier is described as follows:

$$P(c_e|(p_i, p_j)) = \frac{p(c_e) * \prod_{m=1}^{|p_i|*|p_j|} p((d_{i,r}, d_{j,s})|c_e)}{\sum_{k=1}^{|C|} p(c_e) * \prod_{m=1}^{|p_i|*|p_j|} p((d_{i,r}, d_{j,s})|c_e)} \tag{5}$$

For a protein pair (p_i, p_j), the class with highest $P(c_e|(p_i, p_j))$ is assigned as its final class label. In other words, if $I = argmax_{c_e} P(c_e|(p_i, p_j))$, then the protein pair (p_i, p_j) will be classified as an interacting pair. Otherwise, it is classified as non-interacting.

3 Evaluation

In this section, we evaluate the proposed technique for predicting protein interactions. Positive and negative datasets are employed to train a classifier and to evaluate the performance of our method. For positive datasets, interacting proteins are retrieved from DIP[2]—a comprehensive curated catalog of about 44,482 experimentally determined protein-protein interactions in over 110 organisms. We select all *yeast* interactions in DIP to construct our positive dataset I as this species is particularly well-studied. The *yeast* positive dataset consists of 15,658 interactions among 4,749 *yeast* proteins. The negative set of non-interacting protein pairs used in this work is constructed using Algorithm 1 described in the previous section. Proteins are paired up only if they are not from the same cellular location and functional category. This results in a very large negative set of 213,560 protein pairs. To avoid size bias between the positive and negative

[2] http://dip.doe-mbi.ucla.edu/dip/

datasets, we randomly assembled a negative set N with the same number of protein pairs as I.

The domain information of proteins are obtained from the **Pfam** database [2], which contains a large collection of multiple sequence alignments and profile hidden Markov models of protein domains. Both **Pfam-A** and **Pfam-B** are used to ensure sufficient coverage. We first infer the domain-domain interactions from both positive set I and negative set N. Each domain pair gets an interacting probability for I and N using Equation (3).

Table 1. Top 10 interacting and non-interacting domain pairs

Interacting domain pairs	Non-interacting domain pairs
(PF07719, PF00515)	(PF00153, PF00400)
(PF02985, PF02985)	(PF00560, PF00172)
(PF00515, PF00515)	(PF00137, PF00400)
(PF00400, PF00118)	(PF00172, PF07714)
(PF07719, PF07719)	(PF00023, PF00153)
(PF02985, PF00514)	(PF00036, PF00400)
(PF00400, PF00514)	(PF00560, PF00096)
(PF00036, PF00612)	(PF00400, PF00122)
(PF00076, PF00514)	(PF00702, PF00400)
(PF00432, PF01239)	(PF04082, PF00153)

For illustration, Table 1 shows the top 10 interacting and top 10 non-interacting domain pairs respectively. The top interacting domain pairs have maximal values of $P((d_x, d_y)|I)/P((d_x, d_y)|N)$. In other words, these are domain pairs with biggest $P((d_x, d_y)|I)$, while smallest $P((d_x, d_y)|N)$. The top non-interacting domain pairs show those with significant occurrence in non-interacting protein pairs. As we know, not all domain pairs derived from protein-protein interactions are truly interacting as some could occur in interacting proteins by chance. This could lead to false positive domain-domain predictions if we learn from the positive class I only. For example, domain pair (PF07714, PF00400) and (PF00515, PF00806) occurred 170 and 70 times in I respectively. If we just learn from I, it is natural to infer them to be interacting domain pairs since they have high occurrence in interacting set. However, with the help of our biological refined negative class N, we were able to eliminate them since both domain pairs also occurred 1141 and 160 times in N respectively. Furthermore, since our negative class N is more biologically significant than randomly paired proteins, we can estimate the interacting probabilities of each domain pair more precisely and thus result in a more accurate classifier.

For evaluation, we use the inferred domain-domain interactions to classify protein pairs. A 5-fold cross validation is performed to test the accuracy of the classifier described in Equation (5). We compare our results with the reported results using the "Hybrid Classification" technique from reference [7] and the

"Possibility Ranking" technique from reference [6], both of which used positive and negative training datasets for improved protein interaction prediction. Table 2 shows the comparison results of four domain-based protein predication techniques in terms of sensitivity and specificity. The first two techniques were from references [7] and [6]. The other two are our probabilistic technique with two different negative sets, namely, randomly paired negative set (random pairs) and the biologically significant negative set (biological refinement).

Table 2. Classification results of different techniques

Techniques	Specificity	Sensitivity
Hybrid Classification [7]	56.00	86.00
Possibility Ranking [6]	75.00	84.36
Our technique with random pairs	83.71	84.80
Our technique with biological refinement	90.21	87.52

Compared with the techniques in [7] and [6], our probabilistic technique was able to achieve much higher specificity at similar sensitivity regardless of whether it has been trained with random protein pairs as the negative set or the refined negative set assembled using biological domain knowledge. Our classifier that was trained with the biologically refined negative dataset gave the best performance, obtaining an increase of 6.5% and 3.3% in specificity and sensitivity respectively as compared to the same probabilistic classifier trained with negative dataset of randomly paired proteins. This shows that the use of biological domain knowledge for negative dataset construction can benefit the prediction performance of the eventual classifier built on the training data.

The techniques from [7] and [6] were not tested with cross-validation. They randomly selected 20% DIP data as test set and the remaining 80% as training set. Then they repeated their experiments 3 times and got the average results. In fact, as reported in [7], their specificities was rather fluctuating according to the selected test sets. Our method is more robust as our results fluctuated only within 3% in each division of cross validation.

Finally, we also investigate how the size of negative set may affect the performance of our classifier. We systematically increase the size of N by 2 to 10 times. The results are shown in Table 3. With the increase in the size of the negative set, the specificity of our classifier increases while the sensitivity decreases.

Table 3. Performance of classifier with the different size of N

Size ratio	2	4	6	8	10
Specificity	93.00	95.51	96.07	96.27	96.32
Sensitivity	83.21	75.54	73.14	71.50	71.30

One reason that sensitivity has decreased is that the imbalance of positive and negative training set makes our classifier biased towards the negative class N. However, we believe that it is possible to get better performance through intelligently selecting negatives, and we will leave this problem as our future study. Another possible future work involves integrate other biological features such as "amino acid composition" with the domain information in the prediction of protein interactions.

4 Conclusion

In this paper, we predict protein-protein interactions based on domain information. Our learning algorithm first constructs a biologically meaningful negative set based on biological domain knowledge. It then infers the underlying domain interactions based on their probabilities in both interacting class and non-interacting class. A probabilistic classifier for predicting protein interactions is then built upon the inferred probabilistic domain interactions. Our experimental results show that our probabilistic approach is effective and outperforms other similar domain-based techniques for protein interaction prediction.

References

1. J. R. Bock and D.A. Gough. Prediction of protein-protein interaction from primary structure. *Bioinformatics*, 17:455–460, 2001.
2. F. Corpet, F. Servant, J. Gouzy, and D. Kahn. Prodom and prodom-cg: tools for protein domain analysis and whole genome comparisons. *Nucleic Acids Res*, 28:267–269, 2000.
3. T. Dandekar, B. Snel, M. Huynen, and P. Bork. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem. Sci.*, 23:324–328, 1998.
4. M. Deng, F. Sun S. Metha, and T. Chen. Inferring domain-domain interactions from protein-protein interactions. *Genome research*, 12:1540–1548, 2002.
5. A. J. Enright, I. Illiopoulos, N. C. Kyrpides, and C.A. Ouzounis. Protein interaction maps for complete genomes based on gene fusion events. *Nature*, 402:86–90, 1999.
6. D. Han, H. Kim, W. Jang, and S. Lee. Domain combination based protein-protein interaction possibility ranking method. In *IEEE Fourth Symposium on Bioinformatics and Bioengineering (BIBE2004)*, pages 434–441, 2004.
7. D. Han, H. Kim, J. Seo, and W. Jang. Domain combination based probabilistic framework for protein-protein interaction predication. *Genome Informatics*, 14:250–259, 2003.
8. M. A. Huynen and P. Bock. Measuring genome evolution. *Natl Acad. Sci*, 95:5849–5856, 1998.
9. E.M. Marcotte. Detecting protein function and protein-protein interactions from genome sequences. *Science*, 285:751–753, 1999.
10. S. Martin, D. Roe, and J.L. Faulon. Predicting protein-protein interactions using signature products. *Bioinformatics*, 20:1–9, 2004.
11. S.K. Ng and S.H. Tan. Discovering protein-protein interactions. *Journal of Bioinformatics and Computational Biology*, 1(4):711–741, 2004.

12. S.K. Ng, Z. Zhang, and S.H. Tan. Integrative approach for computationally inferring protein domain interactions. *Bioinformatics*, 19:923–929, 2003.
13. R. Overbeek, M. Fonstein, M. D'Souza, G.D. Pusch, and N.Maltsev. The use of gene clusters to infer functional coupling. *Natl Acad. Sci.*, 96:2896–2901, 1999.
14. M. Pellegrini, E.M. Marcotte, M.J. Thompson, D. Eisenberg, and T.O. Yeastes. Assigning protein functions by comparative genome analysis: Protein phylogenetic profiles. *Natl Acad. Sci*, 96:4285–4288, 1999.
15. C. von Mering, R. Krause, B. Snel, et al. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature*, 417(6887):399–403, 2002.
16. K. K. Wan and P. Jong. Large scale statistical prediction of protein-protein interaction by potentially interacting domain (pid) pair. *Genome Informatics*, 13:45–50, 2002.

Semantic Annotation of Biomedical Literature Using Google

Rune Sætre[1], Amund Tveit[1,3], Tonje S. Steigedal[2], and Astrid Lægreid[2]

[1] Department of Computer and Information Science
[2] Department of Cancer Research and Molecular Medicine
[3] Norwegian Center for Patient Record Research, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
{rune.saetre, amund.tveit}@idi.ntnu.no
{tonje.strommen, astrid.laegreid}@medisin.ntnu.no

Abstract. With the increasing amount of biomedical literature, there is a need for automatic extraction of information to support biomedical researchers. Due to incomplete biomedical information databases, the extraction is not straightforward using dictionaries, and several approaches using contextual rules and machine learning have previously been proposed. Our work is inspired by the previous approaches, but is novel in the sense that it is using Google for semantic annotation of the biomedical words. The semantic annotation accuracy obtained - 52% on words not found in the Brown Corpus, Swiss-Prot or LocusLink (accessed using Gsearch.org) - is justifying further work in this direction.

Keywords: Biomedical Literature Data Mining, Semantic Annotation.

1 Introduction

With the increasing importance of accurate and up-to-date databases for biomedical research, there is a need to extract information from biomedical research literature, e.g. those indexed in MEDLINE [34, 33, 15]. Examples of information databases are LocusLink, UniGene and Swiss-Prot [24, 23, 3].

Due to the rapidly growing amounts of biomedical literature, the information extraction process needs to be (mainly) automated. So far, the extraction approaches have provided promising results, but they are not sufficiently accurate and scalable.

Methodologically all the suggested approaches belong to the *information extraction field* [8], and in the biomedical domain they range from simple automatic methods to more sophisticated, but manual, methods. Good examples are: Learning relationships between proteins/genes based on co-occurrences in MEDLINE abstracts (e.g. [16]), *manually* developed information extraction rules (e.g. [35]), information extraction (e.g. protein names) classifiers trained on *manually* annotated training corpora (e.g. [4]), and our previous work on classifiers trained on *automatically* annotated training corpora [32]).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 327–337, 2005.

Examples of biological name entities in a textual context

1. "duodenum, a **peptone** meal in the"
2. "subtilisin plus leucine **amino-peptidase** plus prolidase followed"
3. "predictable hydrolysis of **[3H]digoxin-12alpha** occured in vitro"

Semantic Annotation

An important part of information extraction is to *know* what the information is, e.g. knowing that the term "gastrin" is a protein or that "Tylenol" is a medication. Obtaining and adding this knowledge to given terms and phrases is called *semantic tagging* or *semantic annotation*.

1.1 Research Hypothesis

Our hypothesis is based on ideas from our preliminary experiments using *Google* to generate features for protein name extraction classifiers in [**?**], i.e. using the number of search hits for a word as a feature.

Fig. 1. Google is among the biggest known "information haystacks"

– Google is probably the world's largest available source of heterogeneous electronically represented information. *Can it be used for semantic tagging of textual entities in biomedical literature?* And if so, how?

The rest of this paper is organized as follows. Section 2 describes the materials used, section 3 presents our method, section 4 presents empirical results, section 5 describes related work, section 6 discusses our approach, and finally the conclusion and future work.

2 Materials

The materials used included biomedical (sample of MEDLINE abstract) and general English (Brown) textual corpora, as well as protein databases. See below for a detailed overview.

MEDLINE Abstracts - Gastrin-Selection

The US National Institutes of Health (NIH) grants a free academic licence for PubMed/MEDLINE. It includes a local copy of 6.7 million abstracts, out of the 12.6 million entries that are available on their web interface. As subject for the expert validation experiments we used the collection of 12.238 gastrin-related MEDLINE abstracts that were available in September 2004.

Biomedical Information Databases

As a source for finding already known protein names we used a web search system called Gsearch, developed at Department of Cancer Research and Molecular Medicine at NTNU. It integrates common online protein databases, e.g. Swiss-Prot, LocusLink and UniGene, [24, 23, 3].

The Brown Corpus

The Brown repository (corpus) is an excellent resource for training a Part Of Speech (POS) tagger. It consists of 1,014,312 words of running text of edited English prose printed in the United States during the calendar year 1961. All the tokens are manually tagged using an extended Brown Corpus Tagset, containing 135 tags (Lancester-OsloBergen-tagset). The Brown corpus is included in the Python NLTK data-package, found at Sourceforge.

3 Our Approach

We have taken a modular approach where every submodule can easily be replaced by other similar modules in order to improve the general performance of the system. There are five modules connected to the data gathering phase, namely data selection, tokenization, POS-tagging, Stemming and Gsearch. Then the sixth and last module does a Google search for each extracted term. See figure 2.

1. **Data Selection.** The data selection module uses PubMed Entrez online system to return a set of PubMed IDs (PMIDs) for a given protein, in our case "gastrin" (symbol GAS). The PMIDs are matched against our local copy of MEDLINE, to extract the specific abstracts.
2. **Tokenization.** The text is tokenized to split it into meaningful tokens, or "words". We use the WhiteSpaceTokenizer from NLTK with some extra processing to adapt to the Brown Corpus, where every special character (like ( ) " ' - , and .) is treated as a seperate token. Words in parentheses are clustered together and tagged as a single token with the special tag *Paren*.
3. **POS tagging.** Next, the text is tagged with Part-of-Speech (POS) tags using a Brill tagger trained on the Brown Corpus. This module acts as an advanced stop-word-list, excluding all the everyday common American English words from our protein search. Later, the actually given POS tags are used also as context features for the neighboring words.

Fig. 2. Overview of Our Approach (named Alchymoogle)

4. **Porter-Stemming.** We use the Porter Stemming Algorithm (also from NLTK) to remove even more everyday words from the "possibly biological term" candidate list. If the stem of a word can be tagged by the Brill tagger, then the word itself is given the special tag "STEM", and thereby transferred to the common word list.
5. **Gsearch.** Identifies and removes already known entities from the search, but after the lookup in Gsearch, there are still some unknown words that are not yet stored in our dictionaries or databases, so in order to do any reasoning about these words it is important to know which class they belong to. Therefore, in the next phase they are subjected to some advanced Google-searching, in order to determine this.
6. **Google Class Selection.** We have a network of 275 nouns, arranged in a semantic network on the form "X is a kind of Y". These nouns represent the classes that we want to annotate each word with. The input to this phase is a list of hitherto unknown words. From each Word a query on the form in the example below is formed (query syntax: *Word is (an|a)"*).

   Then these queries are fed to the PyGoogle module which allows 1000 queries to be run against the Google search engine every day with a personal

password key. In order to maximize the use of this quota, the results of every query are cached locally, so that each given query will be executed only once. If a solution to the classification problem is not present among the first 10 results returned, the resultset can be expanded by 10 at a time, at the cost of one of the thousand quota-queries every time.

Each returned hit from Google contains a "snippet" with the given query phrase and approximately 10 words on each side of it. We use some simple regular grammars to match the phrase and the words following it. If the next word is a noun it is returned. Otherwise, adjectives are skipped until a noun is encountered, or a "miss" is returned.

4 Empirical Results

The table below shows the calculated classification scores for the expert evaluation phase. The first column shows *correct* predictions (True Positives and Negatives), the second column shows *incorrect* predictions (False Positives and Negatives), the third column gives Precision and Recall, the fourth gives the standard (balanced) F-Score number, and the last column presents the overall classification accuracy (correct classifications vs. incorrect ones).

Table 1. Semantic classification of *untagged* words

Classifier	TP/TN	FP/FN	Prec/Rec	F-score	CA
Alchymoogle	24/80	31/65	43.6/27.0	33.3	52.0

5 Related Work

Our specific approach was on using Google for *direct* semantic annotation (searching for is-a relations) of tokens (words) in biomedical corpora. We haven't been able to find other work that does this, but Dingare et al. is on using the number of Google hits as input features for a maximum entropy classifier used to detect protein and gene names [10, 11]. Our work differs since we use Google to *directly determine* the semantic class of a word (searching for is-a relationships and parsing text (filtering adjectives) after *(a/an)* in *"Word is (a|an)*, as opposed to Dingare et al.'s *indirect* use of Google search as a feature for the information extraction classifier. A second difference between the approaches is that we search for explicit semantic annotation (e.g. "word is a protein") as opposed to their search for hints (e.g. "word protein"). The third important difference is that our approach does *automatic* annotation of corpuses, whereas they require pre-tagged (manually created) corpuses in their approach.

Other related works include extracting protein names from biomedical literature and some on semantic tagging using the web. Under, a brief overview of related work is given.

Work describing approaches for semantic annotation using the Web can be found in [27, 12, 18, 19, 9, 22].

Semantic Annotation of Biomedical Literature

Other approaches for (semantic) annotation (mainly for protein and gene names) of biomedical literature include:

- Rule-based discovery of names (e.g. of proteins and genes), [13, 29, 36, 35]
- Methods for discovering relationships of proteins and genes, [2, 16].
- Classifier approaches (machine learning) with textual context as features, [4, 5, 6, 14, 1, 20, 30, 21, 17]
- Other approaches include generating probabilistic rules for detecting variants of biomedical terms, [31]

A comprehensive overview of such methods is provided in [28].

The paper by Cimiano and Staab [7] shows that a system (PANKOW) similar to ours works, and can be taken as a proof that automatic extraction using Google is a useful approach. Our systems differ in that we have 275 different semantic tags, while they only use 59 concepts in their ontology. They also have a table explaining how the number of concepts in a system influences the recall and precision in several other semantic annotation systems.

6 Discussion

In the following section we discuss our approach step-by-step. (The steps as presented in fig. 2.)

1. **Data selection.** Since the results were inspected by cancer researchers the focus was naturally on proteins with a role in cancer development, and more specifically cancer in the stomach. One such protein is gastrin, and even though a search in the online PubMed Database returned more than eighteen thousand abstract IDs, only twelve thousand of these were found in our local academic copy of MEDLINE. Therefore only 12.238 abstracts were used as input to the tokenizer. Another important question is if the gastrin collection is representative for MEDLINE in general or for the "molecular biology" part of MEDLINE in particular.
2. **Tokenization into "words".** The tokenization algorithm is important in the sense that it dictates which "words" you have to deal with later in the pipeline. Our choice of using the Brown Corpus for training the Unigram and Brill taggers also influences our choice of tokenizing algorithm. For example, in the Brown Corpus all punctuation characters like comma, full stop, hyphen and so on are written with whitespace both before and after them. This turns them into separate tokens, disconnected from each other and from the other tokens. How to deal with parentheses is another question. Sometimes they are important parts of a protein name (often part of "formulae" describing the protein), and other times they are just used to state that the words within

them aren't that important. We decided to keep the contents of parentheses as a single token, but this kind of parentheses clustering is a hard problem, especially if the parentheses aren't well balanced (like smiley and "1), 2), 3)" style paragraph numbering). Parentheses in MEDLINE are usually well balanced, but still some mistokenization was introduced at this point. Other tokens that require special attention are the multi-word-tokens. They can sometimes be composed using dash, bracket etc. as glue, but are at other times single words separated with whitespaces, even though they should really be one single token. One example is protein names, such as g-protein coupled receptor (GPCR).

3. **a) Brown Corpus and tagging.** To train the Unigram and Brill taggers, an already tagged text is needed as a training set. We used the Brown Corpus, an American English corpus made from texts from 1961. They are rather old, and might not be as representative of "MEDLINE English" as we want. There is also the challenge of how quote symbols and apostrophes are used for protein names in MEDLINE abstracts, e.g. as a marker for the five-prime or three-prime end of a DNA formula. Also, there are only one million words in the corpus, so not all lowercase and capital letter combinations of every word are present.
   b) POS tagging with Brill algorithm and the Brown Corpus. The Brill tagger doesn't tag perfectly, so maybe classifier-based taggers such as SVM could perform better. The performance of the Brill tagger could be better if we used a higher-ordered tagger than the unigram tagger as input to Brill, but the memory need for n-gram taggers are $O(m^n)$, where m is the number of words in the dictionary. So with million word training- and test sets, even the use of just a bi-gram tagger gets quite expensive in terms of memory and time-use. Tagging itself may also introduce ambiguous tags (e.g. superman is a protein, but it may be tagged as a noun/name earlier in the pipeline, because that's the most common sense mentioned in the Brown Corpus).
4. **Porter-stemming.** turns out to work poorly on protein and biological names, since they are often rooted in Latin or have acronyms as their name or symbol. E.g. the symbol for gastrin is GAS, and the porter stem of GAS becomes GA, which is wrong, and too ambiguous.
5. **Gsearch.** The indexing algorithm of Gsearch also contains some stemming of the search terms, leading to some "strange" results when removing well-known proteins from the unknown words list. It should be extended with a larger selection of databases and dictionaries covering biological terms, so that protein names like "peptone" could also be found in the database. In other words there are "precision and recall" issues also at this stage, but our program should be able to solve "half of this problem" automatically. The worst problem is actually how to handle names with "strange characters" like ([]) in them, since these characters are usually not taken into account during the index-building in systems like Gsearch (or Google).
6. **Google Search.** The precision of (positive) classification and the total classification accuracy is close to 50%, which is really good considering that no

context information has been used in the classification process. By using context information in the way that is done in [**?**] it should be possible to increase the classification accuracy further. We had a lower recall than expected (24/89 = 27.0%), mainly because a lot of our unknown words are parts of a multi-word-tokens, and can only be sensibly classified using the context which contains the rest of the multi-word-unit. Also, many of the words are not nouns, so they are not suitable class names in the first place, but still expert biologists often think of them in a concrete way. One example of this is "extracardiac", which were tagged as a place (outside the heart), even though nobody would actually write "extracardiac is a place outside the heart". (Except, I just did! And that really illustrates the problem of freedom, when dealing with Natural Language Understanding.)

We did another test using 1500 semantic classes, instead of the 275 strictly molecular biology related classes. Then we got more hits among the 200 words, so this may be a method to increase the coverage of our system. It is of course much harder to manually evaluate these results, and there is also the danger of lowering the precision this way.

7 Conclusion and Future Work

This paper presents a novel approach - Alchymoogle - using Google for semantic annotation of entities (words) in biomedical literature.

We got empirically promising results - 52% semantic annotation accuracy ((TP+TN)/N, TP=24,TN=80,N=200) in the answers provided by Alchymoogle compared to expert classification performed by a molecular biologist. This encourages further work possibly in combination with other approaches (e.g. rule- and classification based information extraction methods), in order to improve the overall accuracy (both with respect to precision and recall). Disambiguation is another issue that needs to be further investigated. Other opportunities for future work include:

- Improve tokenization. Just splitting on whitespace and punctuation characters is *not* good enough. In biomedical texts non-alphabetic characters such as brackets and dashes need to be handled better.
- Improve stemming. The Porter algorithm for English language gives mediocre results on biomedical terms (e.g. protein names).
- Do spell-checking before a query is sent to Google, e.g. allowing minor variations of words (using the Levenshtein Distance).
- Search for other semantic tags using Google, e.g. "is a kind of" and "resembles", as well as negations ("is not a").
- Investigate whether the Google ranking is correlated with the accuracy of the proposed semantic tag. Are highly ranked pages better sources than lower ranked ones?
- Test our approach on larger datasets, e.g. *all* available MEDLINE abstracts.
- Combine this approach with more advanced natural language parsing techniques in order to improve the accuracy, [25, 26].

– In order to find multiword tokens, one could extend the search query *(" X is (an|a) ")* to also include neighboring words of X, and then see how this affects the number of hits returned by Google. If there is no reduction in the number of hits, this means that the words are "always" printed together and are likely constituents in a multiword token. If you have only one actual hit to begin with, the certainty of the previous statement is of course very weak, but with increasing number of hits, the confidence is also growing.

Acknowledgements

We would like to thank Waclaw Kusnierczyk for proposing additional biomedical information databases for inclusion in future work, and Tore Amble for continuous support. We would also like to thank Martin Thorsen Ranang for proposing improvements for future work. And finally a thanks to the Gsearch developers Jo Kristian Bergum, Hallgeir Bergum and Frode Jünge.

References

1. Steffen Bickel, Ulf Brefeld, Lukas Faulstich, Jrg Hakenberg, Ulf Leser, Condrad Plake, , and Tobias Scheffer. A Support Vector Machine classifier for gene name recognition. In *Proceedings of the EMBO Workshop: A Critical Assessment of Text Mining Methods in Molecular Biology*, March 2004.
2. C. Blaschke, MA. Andrade, C. Ouzounis, and A. Valencia. Automatic Extraction of biological information from scientific text: Protein-protein interactions. In *Proceedings of International Conference on Intelligent Systems for Molecular Biology*, pages 60–67. AAAI, 1999.
3. B. Boeckmann, A. Bairoch, R. Apweiler, MC. Blatter, A. Estreicher, E. Gasteiger, MJ Martin, K Michoud, C. O'Donovan, I. Phan, S. Pilbout, and M. Schneider. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research*, 31(1):365–370, January 2003.
4. Razvan Bunescu, Ruifang Ge, Rohit J. Kate, Edward M. Marcotte, Raymond J. Mooney, Arun Kumar Ramani, and Yuk Wah Wong. Comparative Experiments on Learning Information Extractors for Proteins and their Interactions. *Journal Artificial Intelligence in Medicine: Special Issue on Summarization and Information Extraction from Medical Documents (Forthcoming)*, 2004.
5. Razvan Bunescu, Ruifang Ge, Rohit J. Kate, Raymond J. Mooney, Yuk Wah Wong, Edward M. Marcotte, and Arun Kumar Ramani. Learning to Extract Proteins and their Interactions from Medline Abstracts. In *Proceedings of the ICML-2003 Workshop on Machine Learning in Bioinformatics*, pages 46–53, August 2003.
6. Razvan Bunescu, Ruifang Ge, Raymond J. Mooney, Edward Marcotte, and Arun Kumar Ramani. Extracting Gene and Protein Names from Biomedical Abstracts. Unpublished Technical Note, Machine Learning Research Group, University of Texas at Austin, USA, March 2002.
7. Philipp Cimiano and Steffen Staab. Learning by Googling. *SIGKDD Explorations Newsletter*, 6(2):24–34, December 2004.
8. J. Cowie and W. Lehnert. Information Extraction. *Communications of the ACM*, 39(1):80–91, January 1996.

9. Stephen Dill, Nadav Eiron, David Gibson, Daniel Gruhl, R. Guha, Anant Jhingran, Tapas Kanungo, Sridhar Rajagopalan, Andrew Tomkins, John A. Tomlin, and Jason Y. Zien. SemTag and seeker: bootstrapping the semantic web via automated semantic annotation. In *Proceedings of the Twelfth International World Wide Web Conference, WWW2003*, pages 178–186. ACM, 2003.
10. Shipra Dingare, Jenny Finkel, Christopher Manning, Malvina Nissim, and Beatrice Alex. Exploring the Boundaries: Gene and Protein Identification in Biomedical Text. In *Proceedings of the BioCreative Workshop*, March 2004.
11. Shipra Dingare, Jenny Finkel, Christopher Manning, Malvina Nissim, Beatrice Alex, and Claire Grover. Exploring the Boundaries: Gene and Protein Identification in Biomedical Text. Submitted to BMC Bioinformatics, 2004.
12. Oren Etzioni, Michael Cafarella, Doug Downey, Ana-Maria Popescu, Tal Shaked, Stephen Soderland, Daniel S. Weld, and Alexander Yates. Unsupervised Named-Entity Extraction from the Web: An Experimental Study. Submitted to Artificial Intelligence, 2004.
13. K. Fukuda, A. Tamura, T. Tsunoda, and T. Takagi. Toward information extraction: identifying protein names from biological papers. In *Proceedings of Pacific Symposium on Biocomputing*, pages 707–718, 1998.
14. Filip Ginter, Jorma Boberg, Jouni Jarvinen, and Tapio Salakoski. New Techniques for Disambiguation in Natural Language and Their Application to Biological Texts. *Journal of Machine Learning Research*, 5:605–621, June 2004.
15. Jun ichi Tsuji and Limsoon Wong. Natural Language Processing and Information Extraction in Biology. In *Proceedings of the Pacific Symposium on Biocomputing 2001*, pages 372–373, 2001.
16. Tor-Kristian Jenssen, Astrid Lægreid, Jan Komorowski, and Eivind Hovig. A literature network of human genes for high-throughput analysis of gene expression. *Nature Genetics*, 28(1):21–28, May 2001.
17. Sittichai Jiampojamarn. Biological term extraction using classification methods. Presentation at Dalhousie Natural Language Processing Meeting, June 2004.
18. Vinay Kakade and Madhura Sharangpani. Improving the Precision of Web Search for Medical Domain using Automatic Query Expansion. Online, 2004.
19. Udo Kruschwitz. Automatically Acquired Domain Knowledge for ad hoc Search: Evaluation Results. In *Proceedings of the 2003 Intl. Conf. on Natural Language Processing and Knowledge Engineering (NLP-KE'03)*. IEEE, 2003.
20. Sougata Mukherjea, L. Venkata Subramaniam, Gaurav Chanda, Sriram Sankararaman, Ravi Kothari, Vishal Batra, Deo Bhardwaj, and Biplav Srivastava. Enhancing a biomedical information extraction system with dictionary mining and context disambiguation. *IBM Journal of Research and Development*, 48(5/6):693–701, September/November 2004.
21. M. Narayanaswamy, KE Ravikumar, and K Vijay-Shanker. A biological named entity recognizer. In *Proceedings of the Pacific Symposium on Biocomputing 2003*, pages 427–438, 2003.
22. David Parry. A fuzzy ontology for medical document retrieval. In *Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation - Volume 32*, pages 121–126. ACM Press, 2004.
23. JU. Pontius, L. Wagner, and GD. Schuler. *The NCBI Handbook*, chapter UniGene: a unified view of the transcriptome. National Center for Biotechnology Information, 2003.
24. KD Pruitt and DR Maglott. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Research*, 29(1):137–140, January 2001.

25. Rune Sætre. GeneTUC, A Biolinguistic Project. (Master Project) Norwegian University of Science and Technology, Norway, June 2002.
26. Rune Sætre. Natural Language Processing of Gene Information. Master's thesis, Norwegian University of Science and Technology, Norway and CIS/LMU Munchen, Germany, April 2003.
27. Urvi Shah, Tim Finin, and Anupam Joshi. Information Retrieval on the Semantic Web. In *Proceedings of CIKM 2002*, pages 461–468. ACM Press, 2002.
28. Hagit Shatkay and Ronen Feldman. Mining the Biomedical Literature in the Genomic Era: An Overview. *Journal of Computational Biology*, 10(6):821–855, 2003.
29. Lorraine Tanabe and W. John Wilbur. Tagging gene and protein names in biomedical text. *Bioinformatics*, 18(8):1124–1132, 2002.
30. Manabu Torii and K. Vijay-Shanker. Using Unlabeled MEDLINE Abstracts for Biological Named Entity Classification. In *Proceedings of the 13th Conference on Genome Informatics*, pages 567–568, 2002.
31. Yoshimasa Tsuruoka and Jun'ichi Tsuji. Probabilistic Term Variant Generator for Biomedical Terms. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 167–173. ACM, July/August 2003.
32. Amund Tveit, Rune Sætre, Tonje S. Steigedal, and Astrid Lægreid. ProtChew: Automatic Extraction of Protein Names from . In *Proceedings of the International Workshop on Biomedical Data Engineering (BMDE 2005, in conjunction with ICDE 2005)*, Tokyo, Japan, April 2005. IEEE Press (Forthcoming).
33. Limsoon Wong. A Protein Interaction Extraction System. In *Proceedings of the Pacific Symposium on Biocomputing 2001*, pages 520–530, 2001.
34. Limsoon Wong. Gaps in Text-based Knowledge Discovery for Biology. *Drug Discovery Today*, 7(17):897–898, September 2002.
35. Hong Yu, Vasileios Hatzivassiloglou, Carol Friedman, Andrey Rzhetsky, and W. John Wilbur. Automatic Extraction of Gene and Protein Synonyms from MEDLINE and Journal Articles. In *Proceedings of the AMIA Symposium 2002*, pages 919–923, 2002.
36. Hong Yu, Vasileios Hatzivassiloglou, Andrey Rzhetsky, and W. John Wilbur. Automatically identifying gene/protein terms in MEDLINE abstracts. *Journal of Biomedical Informatics*, 35(5/6):322–330, October 2002.

Fast Parallel Algorithms for the Longest Common Subsequence Problem Using an Optical Bus*

Xiaohua Xu[1], Ling Chen[2,3], Yi Pan[4], and Ping He[2]

[1] Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R.China
JavaArter@21cn.com
[2] Department of Computer Science, Yangzhou University, Yangzhou 225009, P.R.China
lchen@yzcn.net, angeletx@citiz.net
[3] National Key Lab of Novel Software Tech, Nanjing University, Nanjing 210093, P.R.China
[4] Department of Computer Science, Georgia State University, Atlanta, GA 30303, U.S.A.
pan@cs.gsu.edu

Abstract. A parallel algorithm for the longest common subsequence problem on LARPBS is presented. For two sequences of lengths m and n, the algorithm uses p processors and costs $O(mn/p)$ computation time where $1 \le p \le \max\{m, n\}$. Time-area cost of the algorithm is $O(mn/p)$ and memory space required is $O((m+n)/p)$ which all reach optimal. We also show this algorithm is scalable when the number of processors p satisfies $1 \le p \le \max\{m, n\}$. To the best of our knowledge this is the fastest and cost-optimal parallel algorithm for LCS problem on array architectures.

1 Introduction

The problem of longest common subsequence (LCS) is a fundamental problem in bioinformatics. Aho et al. [1] have obtained a lower bound of $\Omega(mn)$ on time for the LCS problem using a decision tree model. Using dynamic programming, it is shown in [2] that the problem can be solved using $O(mn)$ time and $O(mn)$ space. To further reduce the computation time, a lot of parallel algorithms have been proposed for the LCS problem on different models. On CREW-PRAM model, Aggarwal [3] and Apostolico et al [4] independently proposed an $O(\log m \log n)$ time algorithm using $mn/\log m$ processors. On the CRCW-PRAM model, Apostolico et al [4] gave a $O(\log n(\text{loglog } m)^2)$ time algorithm using $mn/\text{loglog } m$ processors. Babu and Saxena [5] improved these algorithms on the CRCW-PRAM model. They designed an $O(\log m)$ algorithm with mn processors and an $O(\log^2 n)$ time optimal parallel algorithm. Unfortunately, these algorithms are designed on the theoretical models and need large number of processors (not less than $mn/\log m$) which are hardly applied in realistic bioin-

* Supported in part by the National Natural Science Foundation under grant No. 60473012, Science Foundation of Educational Commission Jiangsu Province.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 338 – 348, 2005.

formatics problems. Several other algorithms have been proposed on systolic arrays. Chang et al [6] proposed a parallel algorithm with $n+5m$ steps using $m(m+1)$ processing elements. Luce et al [7] designed a systolic array with $m(m+1)/2$ processing elements and $n+3m+q$ steps where q is the length of the common subsequence. Freschi and Bogliolo [8] addressed the problem of computing the LCS between run-length-encoded (RLE) strings. Their algorithm executes in $O(m+n)$ steps on a systolic array of $M+N$ units, where M and N are the lengths of the original strings and m and n are the number of runs in their RLE representation.

Since these algorithms are designed on either theoretical models such as PRAM, or specialized hardware such as systolic arrays, they are not practical. Therefore, it is necessary to develop parallel algorithms for the LCS problem on a realistic computational model. Due to advances in optical technology, optical communication has been implemented in parallel computing systems. LARPBS presented by Pan and Li in [9] is a computational model based on optical bus technology. It has been shown that LARPBS is a realistic model and has been used as a standard computational architecture for parallel algorithm designing by many researchers [10][11][12][13].

In this paper, we propose a scalable LCS algorithm on LARPBS model. Our algorithm uses p ($1 \le p \le \max\{m, n\}$) processors to compute the LCS problem in $O(mn/p)$ time. Our algorithm has several advantages over the algorithms on systolic arrays. First, it is scalable with different values of p. Second, p can be as large as the maximum of m and n. This means that our algorithm can reduce the time complexity to $O(\min\{m,n\})$, while the systolic algorithms all have a time complexity of $O(\max\{m,n\})$. Third, time-area cost of our algorithm is $O(mn)$ which reaches optimal.

The rest of this paper is organized as follows. In Section 2, we explain pipelined optical buses and linear array with reconfigurable pipeline bus system (LARPBS). Primitive communication operations on LARPBS are also summarized. In Section 3, a cost-optimal parallel LCS algorithm on LARPBS is presented and its time and area complexities are analyzed. A scalable cost-optimal parallel LCS algorithm on LARPBS is given in Section 4. Section 5 concludes the paper.

2 Structure and Primitive Operations of LARPBS

Recently, arrays with reconfigurable optical bus systems [10] have been proposed and have drawn much attention from the researchers. In these systems, messages can be transmitted concurrently on a bus in a pipelined fashion and the bus can be reconfigured dynamically under program control to support different algorithmic requirements. LARPBS is one of such model where any processor involvement is not allowed during a bus cycle, except setting switches up at the beginning of a bus cycle. Hence, it can exploit the high bandwidth of optical buses used to connect processors there. Many algorithms have been designed for basic data movement operations, sorting and selection, computational geometry, and PRAM simulation on the LARPBS model [10].

A pipelined optical bus system uses optical waveguides instead of electrical signals to transfer messages among processors. Besides the high propagation speed of light, optical signal transmission on an optical bus has two other important characteristics: they are unidirectional propagation and predictable propagation delay. These advan-

tages of using waveguides enable synchronized concurrent accesses of an optical bus in a pipelined fashion.

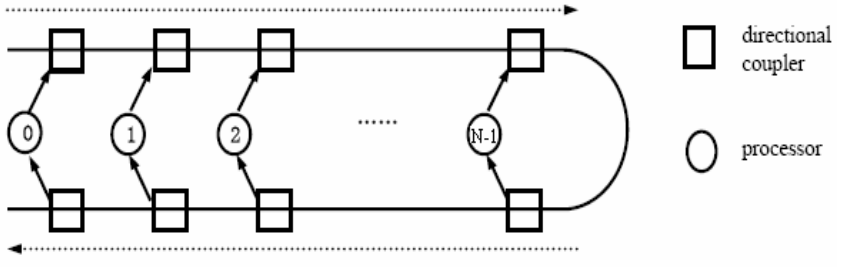


Fig. 1. An linear array of n processors with an optical bus

Fig.1 shows a linear array of N processors connected via an optical bus. Each processor is connected to the bus with two directional couplers. Optical signals propagate unidirectionally from left to right on the upper segment and from right to left on the lower segment. An optical bus contains three identical waveguides, i.e., the massage waveguide for carrying data, the reference waveguide and the select waveguide for carrying address information, as shown in Fig.2. To operate an optical bus in a pipelined fashion so that multiple data transfer can be performed in parallel, one unit delay ω (shown as a loop in Fig. 2) is added between two consecutive processors on the receiving segments of the reference waveguide.

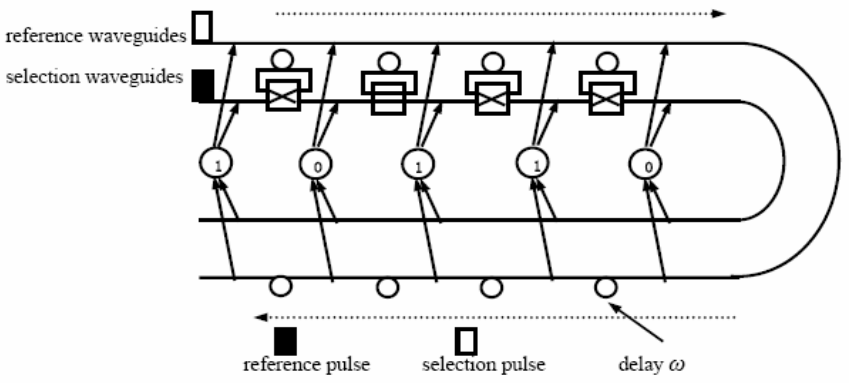


Fig. 2. An optical bus

The coincident pulse addressing technique [10] can be applied on an optical bus system to implement the primitive operations of data transmission. Suppose that a source processor $P(i)$ wants to send a message to a destination processor $P(j)$. Processor $P(i)$ first sends a reference waveguide at time t_{ref}, the beginning of a bus cycle, and a message frame at time t_{ref} on the message waveguide, which propagates synchronously with the reference pulse sent by $P(i)$. Processor $P(i)$ also sends a select pulse at

time $t_{sel}(j)$ on the select waveguide. Whenever processor $P(j)$ detects a coincidence of a reference pulse and a select pulse, it reads the message frame. Thus, in order for processor $P(i)$ to send a message to $P(j)$, we need to have the two pulse coincide at $P(j)$. This happens if and only if $t_{sel}(j) = t_{ref} + (N\text{-}j\text{-}1)\omega$, namely, the select pulse is delayed for $(N\text{-}j\text{-}1)\omega$ time relative to the reference pulse. Since the reference pulse goes through $N\text{-}j\text{-}1$ delays on the receiving segments, the two pulse will meet just at processor $P(j)$.

Based on the coincident pulse addressing technique, the pipelined optical bus systems can support a massive volume of communications simultaneously and are particularly appropriate for applications that involve intensive communication operations such as broadcasting, one-to-one communication, multicasting, compression, split, and many irregular communication patterns. Here we describe implementation details of several primitive operations.

(1) One-to-one communication. Assume that processors $P(i_k)$ (k=1,2,…,M) are senders and processors $P(j_k)$ (k=1,2,…,M) are receivers. The operation is presented as:

for $k \leftarrow 1$ ***to*** M ***pardo*** $R(j_k) \leftarrow R(i_k)$

(2) Broadcasting. In this operation, the source processor $P(i)$ wants to broadcast a message to all the other processors $P(i_k)$ (k=1,2,…,M) in the array:

$R(i_1), R(i_2),\ldots, R(i_M) \leftarrow R(i)$

(3) Multicasting. The implementation of this operation is similar to that of broadcasting. Processor $P(i_k)$ (k=1,2,…,M) wants to broadcast a message to $P(j_{k,1})$, $P(j_{k,2})$, $P(j_{k,3})$,…. Namely:

for $k \leftarrow 1$ ***to*** M ***pardo*** $R(j_{k,1}),R(j_{k,2}),R(j_{k,3}),\ldots \leftarrow R(i_k)$

(4) Element pair-wise operations. Assume $\square$ is a binary operator. Element pair-wise operations are presented as:

for $k \leftarrow 0$ ***to*** N-1 ***pardo*** $R(i+k) \leftarrow R(i+k) \square R(j+k)$

(5) Binary prefix sum. Assume every processor $P(i)$ (i=0,1,…,N-1) has a register $R(i)$ which holds a binary value. The binary prefix sum operation is presented as:

for $i \leftarrow 0$ ***to*** N-1 ***pardo*** $R(i) \leftarrow R(0)+R(1)+\ldots+R(i)$

(6) Extraction and compression. In this operation every processor $P(i)$ (i=0,1,…,N-1) has a value $x(i)$, and we wish to extract those $x(i)$ that have certain property and compact them to the beginning of the linear array. Suppose there are M $x(i_k)$ (k=1,2,…,M) that have such property and they stored in $R(i_1)$, $R(i_2)$,…, $R(i_M)$ ($i_1<i_2<,\ldots,<i_M$), the operation can be presented as:

for $k \leftarrow 1$ ***to*** M ***pardo*** $R(N\text{-}1\text{-}M+k) \leftarrow R(i_k)$

The reader is referred to [10] for implementation details of other primitive operations on the LARPBS model. It has been shown [10] that by using the coincident pulse addressing technique, all the above primitive operations take $O(1)$ bus cycles, where the bus cycle length is the end-to-end message transmission time over a bus . To avoid controversy, let us emphasize that in this paper, by '$O(f(n))$ time' we mean $O(f(n))$ bus cycles for communication plus $O(f(n))$ time for local computation. This approach of computation time measurement is commonly used in other computational models with optical connected bus systems such as AROB.

In addition to supporting fast communications, an optical bus itself can be used as a computing device for global aggregation. It was proven in [10] that by using n processors, the summation of n integers or real numbers with bounded magnitude and

precision, the prefix sums of n binary values, the logical-or and logical-and of n Boolean values can be calculated in constant number of bus cycles. It was also shown in [10] that the minimum or maximum of n numbers can be found in $O(\text{loglog } n)$ time on a LARPBS of size n.

In addition to the tremendous communication capabilities, a LARPBS can also be partitioned into several independent subarrays. The subarrays can operate as regular linear arrays with pipelined optical bus systems, and all subarrays can be used independently for different computations without interference. Hence, this architecture is very suitable for many divide-and-conquer problems. The basic communication, data movement, and aggregation operations provide an algorithmic view on parallel computing using optical buses, and also allow us to develop, specify, and analyze parallel algorithms by ignoring optical and engineering details. These powerful primitives that support massive parallel communications plus the reconfigurability of optical buses make the LARPBS computing model very attractive in solving problems that are both computation and communication intensive. Although LARPBS has great potential of computational ability, it is still a theoretical computational model at present. But unlike many other theoretical models, such like PRAM, the LARPBS model is implementable and practical using current optical technologies.

3 Parallel Longest Common Subsequence Algorithm

Let X and Y be two sequences with lengths of m and n respectively ($m \le n$). The length of X is denoted as $|X|$, and $X[i]$ is the ith character of X. Let $A = (a[i, j])$ be an $m \times n$ matrix, and $a[i, j]$ be the length of the longest common subsequence of $X[1: i]$ and $Y[1: j]$. Matrix A can be computed with the following well-known recursive formula.

$$a[i, j] = \begin{cases} 0, & \text{if either } i = 0 \text{ or } j = 0 \\ a[i-1, j-1] + 1, & \text{if } X[i] = Y[j] \\ \max\{a[i-1, j], a[i, j-1]\}, & \text{if } X[i] \neq Y[j] \end{cases} \tag{1}$$

We assume that all the indexes of the characters in Y matched with X[i] are $0 < j_1 < j_2 < \ldots < j_r < n+1$. Let $j_0 = 0$ and $j_{r+1} = n+1$, we can get the following formula.

$$a[i, j] = \begin{cases} 0, & \text{if either } i=0 \text{ or } j = 0 \\ a[i-1, j], & j \in [0, j_1) \\ a[i-1, j-1] + 1, & j \in \{j_1, j_2, \cdots, j_r\} \text{ where } Y[j_k] = X[i] (k = 1, 2, \cdots, r) \\ \max\{a[i-1, j], a[i-1, j_k - 1] + 1\}, & j \in (j_k, j_{k+1}) \end{cases} \tag{2}$$

The proof for (2) is omitted due space limits.

3.1 Framework of the Parallel LCS Algorithm

Based on (2), we can compute the elements of $a[i, j]$ row by row sequentially and the entries of each row can be computed in parallel using the entries of the previous row. In the process of the ith row, the elements of previous row $a[i\text{-}1, j]$ (j=1,2,…,n) are stored in the registers $d[j]$ and $q[j]$ of $P(j)$ (j=1,2,…,n), the value of $a[i\text{-}1, j_k\text{-}1]+1$ are stored in the registers $q[j_k]$ of $P(j_k)$ (k=1,2,…,r), and elements of $b[i, j]$ (j=1,2,…,n) are stored in the registers $e[j]$ of $P(j)$. The value of $a[i, j]$, i.e. the new value of $d[j]$ can be obtained by parallel computation of $d[j]=\max\{ d[j], q[j] \}$($j$=1,2,…,$n$). The value of $q[j]$ in $P(j)$ can be set by broadcasting $a[i\text{-}1, j_k\text{-}1]+1$ to all the $P(j)$ where j satisfies $j_k \le j < j_{k+1}$. Suppose $m \le n$, in the LARPBS there are n processors which are denoted as $P(1)$, …, $P(n)$. Elements of sequence $X=X[1 : m]$ and $Y=Y[1 : n]$ are stored in the processors: $X[i]$ is in $P(i)$ (i=1,2,…,m) and $Y[j]$ in $P(j)$ (j=1,2,…,n). Suppose in the ith row, $b[i, j_k]=1$ (k=1,…,r), $P(j_k)$ are called critical processor. The framework of the algorithm is as follows.

Algorithm *PLCS* *//Parallel Algorithm for LCS Problem*
Input : sequence $X=X[1 : m]$ and $Y=Y[1 : n]$ are stored in the processors: $X[i]$ is in $P(i)$ (i=1,2,…m) and $Y[j]$ in $P(j)$ (j=1,2,…,n).
Output : The length of LCS of X and Y is in $d[n]$ of $P(n)$.

```
01    for j ← 1 to n pardo          // Initialize the length array d[]
02       d[ j] ← 0      // P(j) does
03    end for
04    for i ← 1 to m do
05       P(i) broadcast X[i] to all processors in the array.
06       for j ← 1 to n pardo       //Compute e[j]
07          e[j] ← (X[i]=Y[j])  //P(j) computes e[j] using X[i] and Y[j] stored in P(j)
08       end for
09       for j ← 1 to n pardo       // Initialize q[]
10          q[j] ← 0  // P(j) initializes the register q[j] 0
11       end for
12       for j ← 1 to n-1 pardo     //One–to-One
13          q[j+1] ← d[j]   //P(j) sends d[j] to P(j+1), and P(j+1) stores d[j] as q[j+1]
14       end for
15       for j ← 1 to n pardo       //Compute d[j] where e[j] = 1
16          if e[j] = 1 then d[j] ← q[j] + 1  // P(j) does
17       end for
18       for j ← 1 to n pardo       //Re-initialize q[]
19          q[j] ← d[j]    // P(j) does
20       end for
21       for j ← 1 to n pardo       //Compress
22          if e[j] = 1 then
23             P(j) is a critical processor, denote the kth critical processor as P(j_k).
               P(j_k) (k=1,2,…,r) compress the index j_k and store j_k in P(N-1-r+k).
24          end if
25       end for
26       for k ← 2 to r pardo       //One–to-One
27          P(N-1-r+k) sends j_k to P(N-1-r+k-1), hence P(N-1-r+k-1) holds j_{k-1} and j_k
```

```
28      end for
29      for k ← 1 to r pardo      //One–to-One
30        P(N-1-r+k) sends the index j_{k+1} to P(j_k) , where we let j_{r+1} be n + 1.
31      end for
32      for k ← 1 to r pardo      // Multicasting
33        //Critical processor P(j_k) broadcasts the value of q[j_k] to q[j]
          //for all j's satisfying j_k ≤ j < j_{k+1}.
          q[j_k], q[j_k+1], …, q[j_{k+1}-1] ← q[j_k]
34      end for
35      for j ← 1 to n pardo      //Compute d[]
36        d[j] ← max{d[j], q[j]}      // P(j) does
37      end for
38    end for
39    return d[n]
```

3.2 Complexity Analysis of the PLCS Algorithm

For the two sequences of sequence $X = X[1 : m]$ and $Y = Y[1 : n]$ $(m \le n)$, it is obvious that the algorithm above uses n processors in the array. In each processor, only constant number of registers is required.

Next we show time complexity of the algorithm is $O(m)$ using n processors. In the algorithm, lines 1-3 are simply operation of assignment which takes $O(1)$ time. In each iteration of the i-loop, line 5 is a primitive operation of broadcast which takes $O(1)$ time. In lines 6-8 every processor computes the $e[j]$ in parallel, obviously this also requires $O(1)$ time. Lines 12-14, lines 26- 28 and lines 29-31 consist of primitive operations of one-to-one communication, they also take $O(1)$ time. Lines 21-24 use a compress operation which requires $O(1)$ time. Lines 32-34 use a multicasting operation which also requires $O(1)$ time. The other lines are simply operations of assignment which require $O(1)$ time. Therefore, the time complexity of each i-loop is $O(1)$. Since the algorithm can be completed after m iterations of i-loops, time complexity of algorithm is $O(m)$. Therefore, time-area cost of the algorithm is $O(mn)$ which is optimal for a parallel algorithm to computes the length of the two sequences with lengths of m and n .

4 Scalability of the Parallel Algorithm

While speed is an important motivation of parallel computing, there is another issue in realistic parallel computing, namely, scalability, which measures the ability to maintain speedup linearly proportional to the number of processors. It is less practical to build an LARPBS system with n processors. For a system with fixed size, a large sequence should be partitioned into subsequences to fit into the fix-sized array. We say that a parallel algorithm is scalable in the range $[p_1, p_2]$, if linear speedup can be achieved for all $p_1 \le p \le p_2$. In the other words, suppose the time complexity of a scalable parallel algorithm be presented as $O(T(n))$ using p processors where n is the size of the problem, for some constant r great then 0, if p/r processors are used, the time complexity must be $rO(T(n))$. Now we show that for $1 \le r \le n$, our algorithm is scal-

able, namely, in case the number of processors is $p(1 \le p \le n)$, the algorithm can be implemented in $O(mn/p)$ time.

To implement the algorithm in $p(1 \le p \le n)$ processors, we first partition and store the elements of sequences X and Y into the processors. Assuming $N = \text{ceil}(n / p)$, each processor has N registers to store the elements of X, here 'ceil' is the ceiling function. Denote the hth such register in $P(g)$ as $X[g, h]$. The element $X[i]$ is stored into $X[g, h]$, here $g = (i\text{-}1) \bmod p + 1$, and $h = \text{ceil}(i / p)$. Elements of sequence Y are stored in the M registers for Y of the processors in the same way, here $M = \text{ceil}(m/p)$. Similarly each processor has N registers for each of the variables d, q and e. Denote the hth such registers in $P(g)$ as $d[g, h]$, $q[g, h]$ and $e[g, h]$ respectively.

Algorithm *SPLCS* *//Scalable Parallel Algorithm for the LCS Problem*
Input : sequence $X=X[1 : m]$ and $Y=Y[1 : n]$ are stored in the processors, each processor has N of X elements and M *of* Y elements stored. In processor $P(p)$, set $d[p, 0] = 0$.
Output : $d[(n\text{-}1) \bmod p + 1, N]$ of $P((n\text{-}1) \bmod p + 1)$

```
01    M ←ceil(m / p), N ←ceil(n / p)
02    for h ← 1 to N do
03       for g ← 1 to p pardo
04          d[g, h] ← 0
05       end for
06    end for
07    for i ← 1 to m do
08       g_i ← (i-1) mod p +1,  h_i ← ceil(i / p)
09       P(g_i) broadcasts X[g_i, h_i] to all processors in the array, and X[g_i, h_i] is
         stored in register Xt[g] of P(g ) (g = 1,2,…,p).
10       for h ← 1 to N do
11          for g ← 1 to p pardo
12             e[g, h] ← (Xt[g] = Y[g, h])
13          end for
14       end for
15       for h ← 1 to N do
16          for g ← 1 to p pardo
17             q[(g mod p) · g + 1, h] ← d[g, h- g/p]
18          end for
19       end for
20       for h ← 1 to N do
21          for g ← 1 to p pardo
22             if e[g, h] = 1 then d[g, h]← q[g, h] + 1
23          end for
24       end for
25       for h ← 1 to N do
26          for g ← 1 to p pardo
27             q[g, h]← d[g, h]
28          end for
29       end for
30       for h ← 1 to N do
```

```
31        P(p) sends d[p, h-1] to q[1, h] of P(1)
32        q[1, h] ← max{q[1, h], d[1, h]},  e[1, h] ← 1
33        for  g ← 1  to  p  pardo
34          if  e[g, h] = 1  then
35            P(g) is a critical processor. Denote the kth critical processor as P(j_k).
              P(j_k) (k=1,2,…, r_h) compress the index j_k and store j_k in P(p-1- r_h +k).
36          end if
37        end for
38        for  k ← 2  to  r_h  pardo      //One–to-One
39          P(p-1-r_h+k) sends j_k to P(p-1- r_h+k-1)
40        end for
41        for  k ← 1  to  r_h  pardo      //One–to-One
42          P(p-1-r_h+k) sends the index j_{k+1} to P(j_k) , here we let j_{rh+1} be p+1.
43        end for
44        for  k ← 1  to  r_h  pardo      // Multicasting
45          // Critical processor P(j_k) broadcasts the value of q[j_k, h] to q[j, h] for
               all j's satisfying  j_k ≤ j < j_{k+1}.
            q[j_k, h], q[j_k+1, h], …, q[j_{k+1}-1, h] ← q[j_k, h]
46        end for
47        for  g ← 1  to  p  pardo       //Compute d[: , h]
48          d[g, h] ← max{d[g, h], q[g, h]}       // P(g) does
49        end for
50      end for  // of  Line 30
51    end for  // of  Line 3
52    return d[(n-1) mod p +1, N]
```

4.1 Complexity Analysis of the SPLCS Algorithm

In each iteration of the *i*-loop in the algorithm, each of lines 1 and 8 takes only $O(1)$ time, and line 9 is a primitive operation of broadcast which takes $O(1)$ time. Lines 2-6 sequentially execute the *h*-loop. Obviously each execution of its loop body also requires $O(1)$ time. Since the body of *h*-loop is executed for N different h values, lines 2-6 takes $O(N)$ time. Similarly, it is obvious that lines 10-29 also requires $O(N)$ time. Lines 30-50 is also a *h*-loop for N different h values. In the body of the loop, lines 38-43 consist of primitive operations of one-to-one communication, and can be executed in $O(1)$ time. Lines 44-46 use a multicasting operation which requires $O(1)$ time. Therefore lines 30-50 also requires $O(N)$ time. Since the algorithm can be completed after m iterations of *i*-loops, time complexity of the algorithm is $O(Nm) = O(mn/p)$, here the number of processors p satisfies $1 \leq p \leq \max\{m,n\}$. Therefore, time-area cost of the algorithm is also $O(mn)$ which is optimal for two sequences with lengths of m and n. Comparisons of the cost of our algorithm with that of other parallel algorithms are shown in Table 1. It is obvious that our algorithm has the least time-area cost and hence the highest efficiency.

Table 1. Comparisons of our algorithm with that of other parallel algorithms

Algorithm	Number of processors	Time complexity	Time-area cost
Robert et al [14]	$m(m+1)$	$n+5m$	$O(m^3+m^2n)$
Chang et al [6]	mn	$4n+2m$	$O(m^2n+mn^2)$
Luce et al [7]	$m(m+1)/2$	$n+3m+q$	$O(m^3+m^2n)$
Freschi et al [8]	$M+N$	$O(m+n)$	$O((M+N)(m+n))$
Ours	m	$O(n)$	$O(mn)$
Ours	n	$O(m)$	$O(mn)$
Ours	$1\leq p \leq \max\{m,n\}$	$O(mn/p)$	$O(mn)$

5 Conclusions

In this paper, we present a parallel algorithm on LARPBS for computing the length of the longest common subsequence of two given sequences. For two sequences of lengths m and n, the algorithm uses p processors for $1\leq p \leq \max\{m,n\}$ and costs $O(mn/p)$ computation time by using the method of row iteration instead of traditional diagonal iteration. Time-area cost of our algorithm is $O(mn)$ which reaches optimal, and space cost is $O((m+n)/p)$ which also reaches optimal. Therefore it is superior to the other parallel algorithms in systolic array [6][7][8][14] in the aspect of efficiency. We also show this algorithm is scalable, namely, when the number of processors is p satisfying $1\leq p \leq \max\{m,n\}$ time complexity is $O(mn/p)$. These algorithms can be modified to compute the edit distance so as to speed up the process of sequences alignment which is useful in bioinformatics.

References

1. Aho, A., Hirschberg, D.S., Ullman, J.: Bounds on the Complexity of the Longest Common Subsequence Problem. J. Assoc. Comput. Mach., Vol. 23, No. 1, pp. 1–12, Jan. 1976
2. Gotoh, O.: An improved algorithm for matching biological sequences, J. Molec. Biol. 162 (1982) 705–708
3. Aggarwal, A., Park, J.: Notes on Searching in Multidimensional Monotone Arrays, Proc. 29th Ann. IEEE Symp. Foundations of Comput. Sci. 1988, pp. 497–512
4. Apostolico, A., Atallah, M., Larmore, L., Mcfaddin, S.: Efficient Parallel Algorithms for String Editing and Related Problems, SIAM J. Computing, vol. 19, pp. 968-988, Oct. 1990
5. Babu, K.N., Saxena, S.: Parallel Algorithms for the Longest Common Subsequence Problem, 4th International Conference on High Performance Computing December 18-21, 1997 - Bangalore, India
6. Chang, J.H., Ibarra, O.H., Pallis, M.A.: Parallel Parsing on a one-way array of finite-state machines, IEEE Trans. Computers C-36 (1987) 64–75
7. Luce, G., Myoupo, J.F.: Systolic-based parallel architecture for the longest common subsequences problem. Integration 25(1): 53-70 (1998)
8. Freschi, V., Bogliolo, A.,: Longest common subsequence between run-length-encoded strings: a new algorithm with improved parallelism, Information Processing Letters, Volume 90 , Issue 4 (May 2004), pp. 167-173

9. Pan, Y., Li, K.: Linear Array with a Reconfigurable Pipelined Bus System – Concepts and Applications, Journal of Information Science 106 (1998) 237-258
10. Li, K., Pan, Y., Zheng, S.-Q eds.: Parallel Computing Using Optical Interconnections, Kluwer Academic Publishers, Boston, USA, Hardbound, ISBN 0-7923-8296-X, October 1998
11. Li, K., Pan, Y., Hamdi, M.: Solving graph theory problems using reconfigurable pipelined optical buses, Parallel Computing, Vol. 26, No. 6, May 2000, pp. 723-735
12. Li, K., Pan, Y., Zheng, S.-Q.: Efficient Deterministic and Probabilistic Simulations of PRAMs on Linear Arrays with Reconfigurable Pipelined Bus Systems, The Journal of Supercomputing, vol. 15, no. 2, pp. 163-181, February 2000.
13. Chen, L., Pan, Y., Xu, X.: Scalable and Efficient Parallel Algorithms for Euclidean Distance Transform on the LARPBS Model. IEEE Trans. Parallel Distrib. Syst. 15(11): 975-982 (2004).
14. Robert, Y., Tchuente, M.: A Systolic Array for the Longest Common Subsequence Problem, Inform. Process. Lett.21 (1985) 191–198

Estimating Gene Networks from Expression Data and Binding Location Data via Boolean Networks

Osamu Hirose[1], Naoki Nariai[1], Yoshinori Tamada[2], Hideo Bannai[1], Seiya Imoto[1], and Satoru Miyano[1]

[1] Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, 108-8639, Japan
[2] Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, 611-0011, Japan

Abstract. In this paper, we propose a computational method for estimating gene networks by the Boolean network model. The Boolean networks have some practical problems in analyzing DNA microarray gene expression data: One is the choice of threshold value for discretization of gene expression data, since expression data take continuous variables. The other problem is that it is often the case that the optimal gene network is not determined uniquely and it is difficult to choose the optimal one from the candidates by using expression data only. To solve these problems, we use the binding location data produced by Lee *et al.* [8] together with expression data and illustrate a strategy to decide the optimal threshold and gene network. To show the effectiveness of the proposed method, we analyze *Saccharomyces cerevisiae* cell cycle gene expression data as a real application.

1 Introduction

Using gene expression data measured by microarrays, estimating gene networks has received considerable attention in the field of bioinformatics. Boolean networks [1, 10] are a mathematical model for representing causal relationships between variables. While the Boolean network model succeeded in estimating gene regulations from expression data in many cases, some practical problems still remain to be solved: (1) Boolean network models require expression data to be discretized into binary values and the threshold of the discretization could affect the resulting gene networks. (2) Boolean network models often suggest a number of gene networks as optimal. However, it is difficult to determine the optimal gene network from the candidates by using expression data only. In this paper, we address the above two issues by using the binding location data observed by Lee *et al.* [8] together with expression data.

The choice of the threshold for discretization of expression data is an essential problem for the Boolean network modeling. Therefore, we have to choose an appropriate threshold, or the estimated gene network is not reliable. The mutual

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3480, pp. 349–356, 2005.

information is sometimes used for selecting the parent set for each gene in the context of Boolean networks [9]. However, we cannot use the mutual information for the threshold selection, because it can only evaluate the goodness of the parent sets for given binarized expression data. Therefore, we newly define a score function for choosing the optimal threshold based on the binding location data. After determining the optimal value of the threshold, we obtain the optimal gene network that gives the best fitting to the expression data and the binding location data. We demonstrate the proposed method by using the *Saccharomyces cerevisiae* cell cycle gene expression data of Spellman *et al.* [13] and estimate a cell cycle-related gene network.

2 Estimation of Gene Networks

2.1 Boolean Networks and Best-Fit Extension

For using the Boolean network model, we need to discretize expression data into binary values by a threshold. We convert an expression value x to 1 for $|x| \geq \theta$, or 0 for $|x| < \theta$, where θ is a threshold for discretization. That is, genes with expression values greater than θ or less than $-\theta$ are considered as significantly changed. A Boolean network model is constructed based on this binarized expression data.

In the Boolean network modeling, Shmulevich *et al.* [11] provides a useful framework for estimating gene networks in the context of the best-fit extension problem, which is originally defined by Boros *et al.* [2]. We call that algorithm the best-fit extension algorithm. In this paper, we first apply the best-fit extension algorithm to the binarized expression data. We briefly explain the best-fit extension algorithm: Let f_i be a Boolean function that models the values of the ith gene, g_i, depending on the expression pattern of its parents, i.e. formally $f_i(v) = 1$ or 0 for $v \in \{0,1\}^{K_i}$, where K_i is the number of parent genes of g_i. First, we construct the set of positive expression patterns T_i and the set of negative expression patterns F_i for g_i from the binarized expression data, where a Boolean function f_i is expected to take 1 or 0 for an expression pattern included in T_i or F_i, respectively. Next, for each f_i, we can define the set $T_i(f_i)$ of expression patterns, called the true set, that f_i takes 1 for an expression pattern of $T_i(f_i)$. The false set $F_i(f_i)$ can also be defined by the expression patterns that f_i takes 0. The loss function $\varepsilon(f_i)$ for selecting the Boolean function f_i and the parent genes of g_i is defined by

$$\varepsilon(f_i) = |T_i(f_i) \cap F_i| + |F_i(f_i) \cap T_i|, \tag{1}$$

where $|A|$ is the number of elements of the set A. We can choose the optimal parents of g_i and the Boolean function f_i by minimizing $\varepsilon(f_i)$. Note that it is a possible case that we find more than one set of parents that give the minimum $\varepsilon(f_i)$ [7]. Moreover, it is possible that we obtain different results from the best-fit extension algorithm for different thresholds for discretization. In the next section, we describe a strategy for choosing the best parents and optimal threshold for discretization.

2.2 Scoring Estimated Gene Networks

The best-fit extension algorithm often gives more than one set of parent genes as optimal for each gene. Let $\mathrm{Pa}_j(i)$ be the jth parent set of g_i, where $i = 1, \ldots, p$ and $j = 1, \ldots, m_i$, and let $p(\mathrm{g}_k, \mathrm{g}_l)$ be the p-value of the relationship "$\mathrm{g}_l \rightarrow \mathrm{g}_k$" given by the binding location data. The best-fit extension algorithm chooses the best parent set for g_i in terms of the loss function $\varepsilon(f_i)$ defined in the previous section. Therefore, we can divide the model estimation process into the optimization of each local structure defined by a gene and its direct parents. Hence, the score function $S_{\mathrm{BLD}}(\theta)$ (Score based on Binding Location Data) for evaluating the threshold for discretization of expression data

$$S_{\mathrm{BLD}}(\theta) = \sum_{i=1}^{p} s_i(\theta), \tag{2}$$

where we define the local score $s_i(\theta)$ that is related to g_i and its parents by

$$s_i(\theta) = \frac{1}{m_i} \sum_{j=1}^{m_i} \sum_{\mathrm{g} \in \mathrm{Pa}_j(i)} \psi(\alpha - \beta \log_{10} p(\mathrm{g}_i, \mathrm{g})) \tag{3}$$

with $\psi(x) = 1/\{1 + \exp(-x)\}$. Here, α and β are parameters ($\beta > 0$). The logistic function $\psi(x)$ can be regarded as a transformation of the p-value to the posterior probability of the edge "$\mathrm{g} \rightarrow \mathrm{g}_i$" conditional on $p(\mathrm{g}_i, \mathrm{g})$, that is $\Pr[\text{"}\mathrm{g} \rightarrow \mathrm{g}_i\text{"} | p(\mathrm{g}_i, \mathrm{g})]$.

Note that we can define the optimal parent set $\widehat{\mathrm{Pa}(i)}$ of g_i by

$$\widehat{\mathrm{Pa}(i)} = \mathrm{argmax}_{\mathrm{Pa}_j(i)} \{ \sum_{\mathrm{g} \in \mathrm{Pa}_j(i)} \psi(\alpha - \beta \log_{10} p(\mathrm{g}_i, \mathrm{g})) \} \tag{4}$$

and $\sum_{\mathrm{g} \in \widehat{\mathrm{Pa}(i)}} \psi(\alpha - \beta \log_{10} p(\mathrm{g}_i, \mathrm{g}))$ can be considered as an alternative of $s_i(\theta)$. However, it is a possible case that the difference between the optimal parents and nearly optimal parents is not significant. For such a case, to obtain a more robust choice of θ, we average $\sum_{\mathrm{g} \in \mathrm{Pa}_j(i)} \psi(\alpha - \beta \log_{10} p(\mathrm{g}_i, \mathrm{g}))$ against possible $\mathrm{Pa}_j(i)$. Therefore, the score function $s_i(\theta)$ defined in (3) can be considered as the averaged score when we set the equal probability to all possible $\mathrm{Pa}_j(i)$. In addition, by using the logistic function $\psi(x)$, we do not need to consider the threshold for the p-value because $\psi(x)$ is a continuous function. It should be noted that the p-value $p(\mathrm{g}_i, \mathrm{g}_j)$ is not available when g_j is not a transcription factor kept in the binding location data. In such a case, we use $\psi(\alpha - \beta \log_{10} p(\mathrm{g}_i, \mathrm{g}_j)) = \lambda |\rho_{\mathrm{R}}(\mathrm{g}_i, \mathrm{g}_j)|$, where $\rho_{\mathrm{R}}(\mathrm{g}_i, \mathrm{g}_j)$ is Pearson's correlation between $(x_i(2), ..., x_i(T))$ and $(x_j(1), ..., x_j(T-1))$ and λ $(0 \leq \lambda \leq 1)$ is a parameter. Here $x_i(t)$ is the original expression value of g_i at time t.

By using the score function $S_{\mathrm{BLD}}(\theta)$ defined above, we can describe an algorithm for selecting the optimal threshold θ and gene network G as follows:

Threshold Selection Step

Step 1 Discretize expression data by a threshold θ.

Step 2 Choose the parent sets $\mathrm{Pa}_j(i)$ for each g_i by the best-fit extension algorithm.

Step 3 Compute $S_{\mathrm{BLD}}(\theta)$ by (2) and (3).

Step 4 Repeat from Step 1 to Step 3 for each θ at a suitable interval.

Step 5 Choose the optimal threshold $\hat{\theta}$ that gives the largest $S_{\mathrm{BLD}}(\theta)$ in Step 4.

Network Construction Step

Step 6 Construct the parent sets $\mathrm{Pa}_j(i)$ of g_i for the binarized expression data against the optimal threshold $\hat{\theta}$ by the best-fit extension algorithm.

Step 7 Choose the optimal parent set $\widehat{\mathrm{Pa}(i)}$ from the parent sets $\mathrm{Pa}_j(i)$ in Step 6 by (4).

Step 8 Repeat Step 6 and Step 7 for all genes.

Step 9 The optimal gene network is constructed by connecting each optimal local structure $\hat{L}_i$ defined by g_i and $\widehat{\mathrm{Pa}(i)}$.

2.3 Computational Complexity

Lähdesmäki *et al.* [7] proved that the best-fit extension problem with p genes and k parents is solvable in

$$O\left(\binom{p}{k} \cdot p \cdot poly(k) \cdot N\right)$$

time, where N is the number of microarrays and $poly(k)$ is some polynomial of k. Although the threshold θ may be any real number, the number of candidates can be limited to pN, since N time course microarrays with p genes has only pN expression values.

Therefore solving the best-fit extension problem for all possible θ can be done in $O(\binom{p}{k} \cdot p^2 \cdot poly(k) \cdot N^2)$ time. Calculations for each gene and threshold are independent of each other, and the algorithm can easily be made to utilize multiple processors for parallel computation. In the experiment of Section 3 ($p = 52$, $k = 3$, $N = 32$), about 13 minutes were needed for a given threshold (Sun Fire 15K, 96CPUs, 1.2GHz).

3 Real Data Example

We analyze *Saccharomyces cerevisiae* cell cycle gene expression data collected by Spellman *et al.* [13]. In this analysis, we use 34 time course microarray data of α factor arrest and elutriation, because the other data contain many missing values. In the binding location data, 106 genes are considered as the transcription factors and the p-values are observed for each relationship between a transcription factor and the other gene. At first, we choose transcription factors from the above 106 genes that are also categorized as the cell cycle related genes defined by Spellman *et al.* [13]. This gives us 14 genes, ACE2, ASH1, CHA4, FKH1, GCR1,

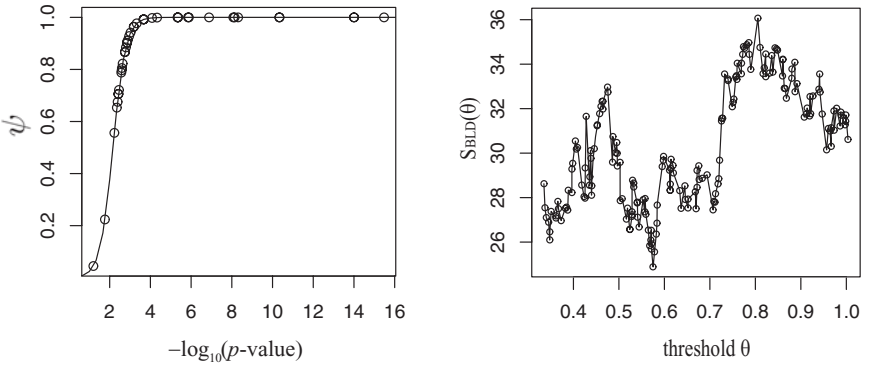


Fig. 1. (Left) Conversion of the p-values given in the binding location data into the scores by the logistic function. (Right) The result of the threshold selection step. Y-axis is the score of the network based on binding location data for each θ

PHD1, RAP1, RME1, SRD1, STB1, STP2, SWI4, SWI5 and YFL044C, as the transcription factors. Next, we collected three genes with the smallest p-value against each transcription factor listed above. As a result, 52 genes were selected to be analyzed. The purpose of this analysis is to construct a gene network among these 52 genes and to decide the threshold for expression data. The parameters α and β are set so that ψ takes 0.94 for $p = 10^{-3}$ and 0.001 for $p = 1.0$. The rationale for choosing these parameters are: (1) It is thought that 94% of gene pairs which have a p-value less than 0.001 interact with each other [8]. (2) We assume that two genes do not interact when the p-value equals to 1.00, and let ψ take a value close to 0 (0.001) when $p = 1.00$. Figure 1 (left) shows the logistic function in this parameter set. Each circle is the conversion of p-values between each transcription factor and its selected genes. The parameter λ is set to 0.867 which is the minimum value of $\psi(\alpha - \beta \log_{10}(g_{\mathrm{TF}}(j), g_{\mathrm{TOP}}(j)))$ for all j, where $g_{\mathrm{TF}}(j)$ is the jth transcription factor and $g_{\mathrm{TOP}}(j)$ is the gene which has the minimum p-value for $g_{\mathrm{TF}}(j)$. This is because regulations supported by the binding location data are more reliable than those supported by the correlation of expressions.

Figure 1 (Right) shows the values of score function $S_{\mathrm{BLD}}(\theta)$ for candidate values of the threshold θ. Since expression data are normalized for each gene, it does not make sense that almost all genes have significantly changed. Therefore, the minimum θ is set to 0.335 so that the number of 1's in the binarized expression data is less than 90%. We observe that the optimal threshold $\hat{\theta} = 0.802$ from Figure 1. Figure 2 represents the resulting gene network obtained by the proposed method. The 14 transcription factors are shown by red nodes. Note that the thick edges are from transcription factors to transcription factors, the thin solid edges are from transcription factors to non-transcription factor genes, and the dotted lines are from non-transcription factor genes to any gene. In the estimated network, there exist many interesting facts. First, most of the transcription factors

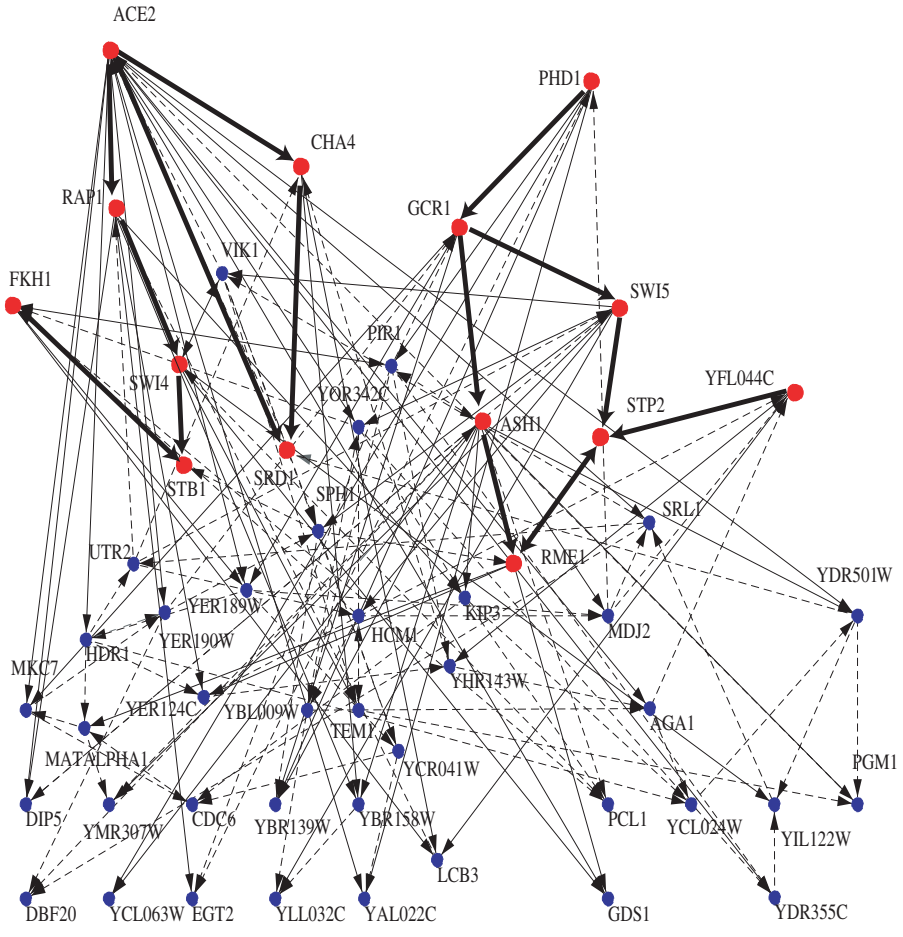


Fig. 2. The resulting network estimated by the proposed method

have many outgoing edges, and this fact agrees with real biological processes. Second, in the estimated network, SWI4 (one component of SBF) is a parent of STB1. This relationship is suggested by Costanzo *et al.* [4] and Ho *et al.* [6]. Third, it seems to be plausible that SWI5 is upstream of RME1 according to the previous work [5]. Fourth, there is a relationship "SWI5 → VIK1 → ACE2" in the estimated network. SWI5 and ACE2 are M/G1 phase specific genes [5], and VIK1 is known to have the microtubule motor activity [3]. Therefore this relationship may be correct. We expect that the estimated network contains other biologically meaningful relationships of true regulatory mechanisms.

Shmulevich and Zhang [12] proposed a method for binarizing expression data. An advantage of their method is that the threshold can be set for each gene. Their method is based on expression data only, and transcriptional information is not considered. For evaluating the proposed method, we obtained $S_{\mathrm{BLD}}(\theta) = 31.738$

for expression data binarized by the method of Shmulevich and Zhang [12]. On the other hand, our method obtained $S_{\mathrm{BLD}}(\theta) = 36.241$. Of course, this result cannot claim that our method is better than Shmulevich and Zhang [12]'s method because our method uses the binding location data for model learning. We can, however, conclude that the proposed method works well in practice.

4 Discussion

In this paper, we proposed a novel strategy to construct Boolean networks aimed at estimating gene networks based on expression data and binding location data. The choices of the threshold for discretization of expression data and the optimal parent set for each gene can be regarded as essential problems in the Boolean network modeling. By using the binding location data, we newly defined a score function so that we can choose the optimal threshold and the parent set for each gene. A real example showed that the effectiveness of the proposed method and we succeeded in estimating a reliable gene network.

We consider the following problems as our future works: (1) In our score function $S_{\mathrm{BLD}}(\theta)$, if a gene that is not a transcription factor is selected as a parent, we simply put the correlation between genes to the score function. However, we could establish a method to calculate the value of the score function in such a case by using other types of genomic data. (2) For selecting the optimal parent set for each gene, we used an exhaustive search method that calculates the score function for all possible parent sets. Therefore, when the number of genes we analyze is large, the computational time for constructing a gene network is a problem in practice. We would like to discuss these problems in our future paper.

References

1. Akutsu, T., Miyano, S., and Kuhara, S.: Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. Pac. Symp. Biocomput. **4** (1999) 17–28
2. Boros, E., Ibaraki, T., and Makino, K.: Error-free and best-fit extensions of partially defined Boolean functions. Information and Computation **140** (1998) 254–283
3. Cherry, J.M., Adler, C., Ball, C., Chervitz, S.A., Dwight, S.S., Hester, E.T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., Weng, S., Botstein, D.: SGD: Saccharomyces Genome Database. Nucleic Acids Research **26**(1) (1998) 73–79
4. Costanzo, M., Schub, O., Andrews, B.: G1 transcription factors are differentially regulated in *Saccharomyces cerevisiae* by the Swi6-binding protein Stb1. Mol Cell Biol. **14** (2003) 5064–5077
5. Frenz, L.M., Johnson, A.L., and Johnston, L.H.: Rme1, which controls CLN2 expression in *Saccharomyces cerevisiae*, is a nuclear protein that is cell cycle regulated. Mol. Genet. Genomics **266** (2001) 374–384
6. Ho, Y., Costanzo, M., Moore, L., Kobayashi, R., and Andrews, B.J.: Regulation of transcription at the *Saccharomyces cerevisiae* start transition by Stb1, a Swi6-binding protein. Mol. Cell. Biol. **19** (1999) 5267–5278

7. Lähdesmäki, H., Shmulevich, I., and Yli-Harja, O.: On learning gene regulatory networks under the Boolean network model. Machine Learning **52** (2003) 147–167
8. Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J-B., Volkert, T.L., Fraenkel, E., Gifford D.K., and Young, R.A.: Transcriptional regulatory networks in *Saccharomyces cerevisiae*, Science **298** (2002) 799–804
9. Liang, S., Fuhrman, S., and Somogyi, R.: REVEAL, a general reverse engineering algorithm for inference of genetic network architectures. Pac. Symp. Biocomput. **3** (1998) 18–29
10. Shmulevich, I., Dougherty, E.R., Kim, S., and Zhang, W.: Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. Bioinformatics **18** (2002) 261–274
11. Shmulevich, I., Dougherty, E.R., and Zhang, W.: From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. Proceedings of the IEEE **90** (2002) 1778–1792
12. Shmulevich, I., and Zhang, W.: Binary analysis and optimization-based normalization of gene expression data. Bioinformatics **18** (2002) 555–565
13. Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B.: Comprehensive identification of cell cycle regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol. Biol. Cell **9** (1998) 3273–3297

Efficient Matching and Retrieval of Gene Expression Time Series Data Based on Spectral Information

Hong Yan

Department of Computer Engineering, and Information Technology,
City University of Hong Kong, Kowloon, Hong Kong,
School of Electrical and Information Engineering,
University of Sydney, NSW 2006, Australia
h.yan@cityu.edu.hk

Abstract. In this paper, we propose an efficient method based on spectral analysis for matching and retrieval of gene expression time series data. In this technique, we decompose a gene expression time series into a set of spectral components. The spectral parameters can then be used to compute the correlation between the expression data for a pair of genes using a closed-form mathematical equation. This method provides a reliable similarity metric for the comparison of gene expression data and can be used for efficient data retrieval.

1 Introduction

Massive amount of gene expression data can be produced in DNA microarray experiments. Analysis of these data is a challenging problem [1-5]. In this paper, we focus on gene expression time series data. These data are commonly called discrete-time signals in signal processing. Spectral analysis has been used successfully in many applications, such as speech and music signal processing, wireless communications, ECG and EEG data analysis, and medical image reconstruction [6-9]. We demonstrate in this paper that parametric spectral estimation can also be used for gene expression time series data matching and retrieval effectively.

An important problem in microarray data analysis is how to compare two sets of gene expression data. For example, in data mining, we may have a given gene expression time series as a key and need to search for all expression data in a database that are similar to the key. Thus, we need a similarity metric to compare gene expression data. Commonly used similarity metrics include the Pearson correlation coefficient and the Euclidian distance. We show in Section 2 that the two metrics are in fact equivalent for normalized data. Due to noise and phase difference, the Pearson correlation coefficient and the Euclidian distance do not work reliably and we need to find better similarity metrics.

Several methods have been proposed to compare gene expression data more accurately. Aach and Church have developed an algorithm to compute the distance between two gene expression time series based on time warping [10]. The basic idea is to allow two waveforms to be expanded or contracted locally along the time axis to produce an optimal matching between them. This is often called flexible or elastic

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 357 – 373, 2005.

pattern matching and has been used in speech and handwriting recognition extensively. Filkov, Skiena and Zhi have introduced an edge detection method to filter the data and detect the edges, which are defined as piece-wise linear and monotonic segments, in a gene expression waveform [11]. The edge information is then used to match a pair of gene expression time series. Kwan, Hoos and Ng have developed an event based method to compare gene expression time series data based on the waveform slope information [12]. Their algorithm segments a gene expression waveform into rising (R), falling (F), and constant (C) regions. In this way, the problem of comparing two gene expression time series becomes one of alignment of event strings, which is similar to DNA or protein sequence alignment and can be solved efficiently based on dynamic programming. Ji and Tan have recently proposed a method to transform a gene expression time series data matrix to a slope matrix, which contains the information about temporal expression value changes, including increase, decrease or unchanged [13]. The slope information is then used to generate so-called q-clusters, each of which contains genes with similar expression patterns over q consecutive conditions. The q-clusters are then used to analyze the co-regulations between genes and gene clusters. They have also shown successful application of the slope information to mining gene expression data for positive and negative co-regulated gene clusters [14].

We have recently developed a spectral analysis based method for gene expression data comparison [15]. In our method, an expression time series is represented as the sum of spectral components and the correlation is measured component-wise. In our method, each spectral component is completely described by four parameters, its amplitude, phase, damping rate and frequency and we can calculate the correlation between two spectral components by setting the phase to zero. This effectively solves the phase shift problem for similar waveform shapes.

In this paper, we extend our spectral analysis based method and introduce a closed-form mathematical solution for the correlation between spectral components. This new solution makes the similarity computation very fast. In addition, the new solution has several advantages. It can extrapolate gene expression time series and can match data obtained with different sampling rates and data lengths. The method can be used to design efficient gene expression data matching and retrieval algorithms.

The remaining parts of the paper are organized as follows. In Section 2, we discuss the Pearson correlation coefficient and its limitation. In Section 3, we review the spectral component-wise correlation technique. Section 4 presents a fast method for computing the component-wise correlation and Section 5 develops an algorithm for data retrieval based on spectral parameters. Section 6 shows our experimental results on a database of expression data from genes with regulatory relations and Section 7 concludes the paper.

2 The Pearson Correlation Coefficient and Its Limitation

Assume that the expression time series data of two genes *a* and *b* are represented by two vectors:

$$\mathbf{x}_a = \begin{bmatrix} x_a[0] & x_a[1] & \cdots & x_a[N-1] \end{bmatrix}^T$$

and

$$\mathbf{x}_b = \begin{bmatrix} x_b[0] & x_b[1] & \cdots & x_b[N-1] \end{bmatrix}^T$$

where N is the number of sampled data points measured along the time axis. The Pearson correlation coefficient between these two vectors is defined as:

$$R(\mathbf{x}_a, \mathbf{x}_b) = \frac{\sum_{n=0}^{N-1} (x_a[n] - \bar{x}_a)(x_b[n] - \bar{x}_b)}{\sqrt{\sum_{n=0}^{N-1} (x_a[n] - \bar{x}_a)^2} \sqrt{\sum_{n=0}^{N-1} (x_b[n] - \bar{x}_b)^2}} \tag{1}$$

where

$$\bar{x}_a = \frac{1}{N} \sum_{n=0}^{N-1} x_a[n], \text{ and } \bar{x}_b = \frac{1}{N} \sum_{n=0}^{N-1} x_b[n], \tag{2}$$

are means values of the elements in $\mathbf{x}_a$ and $\mathbf{x}_b$ respectively.

Let

$$\sigma_a^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x_a[n] - \bar{x}_a)^2, \text{ and } \sigma_b^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x_b[n] - \bar{x}_b)^2, \tag{3}$$

which are the variance of the elements in $\mathbf{x}_a$ and $\mathbf{x}_b$ respectively. It can be proven that Equation (1) can be rewritten as

$$R(\mathbf{x}_a, \mathbf{x}_b) = 1 - \frac{1}{2} D^2(\mathbf{x}_a, \mathbf{x}_b) \tag{4}$$

where

$$D(\mathbf{x}_a, \mathbf{x}_b) = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} \left(\frac{x_a[n] - \bar{x}_a}{\sigma_a} - \frac{x_b[n] - \bar{x}_b}{\sigma_b} \right)^2} \tag{5}$$

is the Euclidian distance between the mean removed and variance normalized version of $\mathbf{x}_a$ and $\mathbf{x}_b$.

Equation (4) shows the relationship between the Pearson correlation coefficient and the Euclidian distance between two vectors. Geometrically, $R(\mathbf{x}_a, \mathbf{x}_b) = \cos\theta^{(a,b)}$, where $\theta^{(a,b)}$ is the angle between two vectors while $D(\mathbf{x}_a, \mathbf{x}_b)$ measures the distance between the end points of two vectors (Figure 1). Clearly, if the two vectors are normalized, when the angle is $\theta^{(a,b)}$ is small, the Pearson correlation coefficient $R(\mathbf{x}_a, \mathbf{x}_b)$ is large and the distance $D(\mathbf{x}_a, \mathbf{x}_b)$ is small, and vice versa. Because of the close relationship between $R(\mathbf{x}_a, \mathbf{x}_b)$ and $D(\mathbf{x}_a, \mathbf{x}_b)$, we will only consider $R(\mathbf{x}_a, \mathbf{x}_b)$ in the rest of the paper.

$R(\mathbf{x}_a, \mathbf{x}_b)$ measures the closeness of two vectors in terms of the angle between them. It works well only if the vector elements are aligned. As an example, let us consider $N = 6$, and the following five vectors:

$$\mathbf{x}_1 = [1 \quad -2 \quad 1 \quad 0 \quad 0 \quad 0]^T, \ \mathbf{x}_2 = [1 \quad -2 \quad 1 \quad 0 \quad 0 \quad 0]^T,$$
$$\mathbf{x}_3 = [0 \quad 1 \quad -2 \quad 1 \quad 0 \quad 0]^T, \ \mathbf{x}_4 = [0 \quad 0 \quad 1 \quad -2 \quad 1 \quad 0]^T,$$
$$\mathbf{x}_5 = [0 \quad 0 \quad 0 \quad 1 \quad -2 \quad 1]^T.$$

The five vectors represent the shifted versions of the same waveform. We have $R(\mathbf{x}_1, \mathbf{x}_2) = 1$, $R(\mathbf{x}_1, \mathbf{x}_3) = -0.667$, $R(\mathbf{x}_1, \mathbf{x}_4) = 0.167$, $R(\mathbf{x}_1, \mathbf{x}_5) = 0$. We can see that a shift of the waveform can dramatically change the value of the Pearson correlation coefficient.

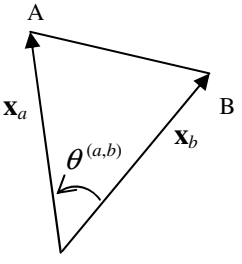


Fig. 1. Geometric meanings of the Pearson correlation coefficient and the Euclidian distance between two vectors.

In gene expression time series analysis, we consider two genes have similar functions or a regulatory relation when they have similar expression waveforms. The time delay between the waveforms may determine whether the two genes are an activation or inhibition pair, but whether two genes form a regulatory pair should depend on the similarity of the waveform shapes and is independent of the time delay between them. One may ask why we do not just shift a time series for all possible time units relative to the other time series, compute the correlation between the two time series and find the maximum correlation value among all shift distances. For example, we can shift the elements in $\mathbf{x}_5$ to the left by three units, ignore the non-overlapping samples, and then compute the Pearson correlation coefficient between $\mathbf{x}_1$ and $\mathbf{x}_5$. The correlation coefficient will now become 1. This approach has two disadvantages. Firstly, if we shift one time series relative to the other, the number of overlapping samples will be reduced and the computation becomes more noise prone and less reliable. Secondly, it may produce an erroneous result. For example, if we shift one arbitrary time series by $N - 1$ points against another arbitrary time series so that there is only one point overlapping in two time series, then the Pearson correlation coefficient is always 1 as long as the overlapping samples are nonzero and non-overlapping points are ignored.

As discussed above, the Pearson correlation coefficient is strongly affected by the alignment of waveforms along the time axis, and thus it does not produce reliable

results for gene expression time series analysis. Clearly an improved similarity metric is needed.

3 Spectral Component Correlation

To solve the alignment problem between two waveforms, we decompose a time series into a set of spectral components and set the phase of these components to zero before computing the correlations among them. Let

$$x[n] = \sum_{k=1}^{2K} c_k z_k^n = \sum_{k=1}^{2K} c_k e^{\sigma_k n + j\omega_k n} \quad (0 \le n \le N-1) \tag{6}$$

then it can be shown that the following equations hold [7-9,15]

$$\begin{bmatrix} x[0] & x[1] & \cdots & x[2K-1] \\ x[1] & x[2] & \cdots & x[2K] \\ \vdots & \vdots & \ddots & \vdots \\ x[N-2K-1] & x[N-2K] & \cdots & x[N-2] \end{bmatrix} \begin{bmatrix} p_{2K} \\ p_{2K-1} \\ \vdots \\ p_1 \end{bmatrix} = - \begin{bmatrix} x[2K] \\ x[2K+1] \\ \vdots \\ x[N-1] \end{bmatrix} \tag{7}$$

where p_k $(1 \le k \le 2K)$ are coefficients of the following polynomial

$$p(z) = \prod_{k=1}^{2K} (z - z_k) = \sum_{k=0}^{2K} p_k z^{2K-k} \quad (p_0 = 1) \tag{8}$$

Equation (7) represents the autoregressive (AR) model of the time series $x[n]$. The matrix equation consists of all linear prediction equations for the time series. The damping rates σ_k and frequencies ω_k, where $z_k = \sigma_k + \omega_k$, in Equation (6) can be determined from the roots of the polynomial in Equation (8) after we obtain p_k from Equation (7). Once z_k are known, Equation (6) can be rewritten as another set of linear equations for different n values and we can then find c_k from these equations.

Because $x[n]$ is a real-valued sequence, z_k and c_k must occur in complex conjugate pairs. Let $c_k = \alpha_k e^{j\varphi_k}$, then Equations (6) becomes

$$x[n] = \sum_{k=1}^{K} x^{(k)}[n;\varphi_k] = \sum_{k=1}^{K} \alpha_k e^{\sigma_k n} \cos(\omega_k n + \varphi_k) \quad (0 \le n \le N-1) \tag{9}$$

where α_k and φ_k are respectively the amplitude and phase of the k-th spectral component

$$x^{(k)}[n;\varphi_k] = \alpha_k e^{\sigma_k n} \cos(\omega_k n + \varphi_k) \tag{10}$$

Now we can compute the correlation between two gene expression time series in terms of spectral components. Let

$$\mathbf{x}_a^{(i)} = \left[x_a^{(i)}[0;0] \quad x_a^{(i)}[1;0] \quad \cdots \quad x_a^{(i)}[N-1;0]\right]^T \tag{11}$$

$$\mathbf{x}_b^{(k)} = \left[x_b^{(k)}[0;0] \quad x_b^{(k)}[1;0] \quad \cdots \quad x_b^{(k)}[N-1;0]\right]^T \tag{12}$$

which are component vectors of $x_a^{(i)}[n;\varphi_i]$ and $x_b^{(k)}[n;\varphi_k]$ respectively with their phases set to zero. We define the spectral component based similarity or simply spectral similarity between two gene expression time series $\mathbf{x}_a$ and $\mathbf{x}_b$ as

$$R_{SP}(\mathbf{x}_a, \mathbf{x}_b) = \arg \max_{\alpha_i, \alpha_k > \alpha_T} \left| R\left(\mathbf{x}_a^{(i)}, \mathbf{x}_b^{(k)}\right) \right| \tag{13}$$

where α_T stands for the spectral component amplitude threshold. The thresholding operation is used to remove noisy components, which will be further discussed in Section 5.

4 Fast Algorithm for Computing Spectral Component Correlation

The parametric representation of a gene expression time series in Equations (9) and (10) makes it possible to extract the phase information. We can solve the time alignment problem by setting the phases to zero before computing the correlation between two spectral components. Another advantage of this representation is that we can extrapolate the time series. Although a microarray can accommodate a lot of genes in an experiment, only a small number of data points are measured along the time axis because it is very costly to conduct an experiment. However, Equations (9) and (10) represent a time series using a parametric model and we can extrapolate the expression values for $n \geq N$ based on the equations. In fact, Equations (9) and (10) can be considered as discrete versions of continuous signals or functions, which will be analyzed below.

Now we consider continuous representations of Equations (9) and (10). In order to avoid too many notations, we use $x[n]$ to represent a discrete-time signal or a sequence, $x(t)$ to represent its corresponding continuous signal or function, and $R(\mathbf{x}_a, \mathbf{x}_b)$ to represent the Pearson correlation coefficient between either the discrete signals or the continuous signals, depending on the context. The continuous functions corresponding to Equations (9) and (1) can be written as

$$x(t) = \sum_{k=1}^{K} x^{(k)}[t;\varphi_k] = \sum_{k=1}^{K} \alpha_k e^{\sigma_k t} \cos(\omega_k t + \varphi_k) \quad (t \geq 0) \tag{14}$$

$$x^{(k)}[t;\varphi_k] = \alpha_k e^{\sigma_k t} \cos(\omega_k t + \varphi_k) \tag{15}$$

The Pearson correlation coefficient between two continuous functions $x_a(t)$ and $x_b(t)$ is defined as

$$R(\mathbf{x}_a, \mathbf{x}_b) = \frac{\int_0^\infty x_a(t)x_b(t)dt}{\sqrt{\int_0^\infty x_a^2(t)dt}\sqrt{\int_0^\infty x_b^2(t)dt}} \tag{16}$$

Strictly speaking, Equation (16) is different from Equation (1) in the sense that the mean values of $x_a(t)$ and $x_b(t)$ are not removed before computing the correlation. There are two reasons for neglecting the means values. Firstly, each of the three integrations in Equation (16) becomes infinity if we subtract mean values from $x_a(t)$ and $x_b(t)$ unless the means are equal to zero since the time variable t goes from 0 to ∞. Secondly, if we neglect the damping rate, each spectral component in Equations (14) and (15) indeed has zero mean.

Let us consider two continuous spectral components, corresponding to the discrete representations in Equations (11) and (12)

$$x_a^{(i)}[t;\varphi_i] = \alpha_i e^{\sigma_i t}\cos(\omega_i t + \varphi_i) \text{ and } x_b^{(k)}[t;\varphi_k] = \alpha_k e^{\sigma_k t}\cos(\omega_k t + \varphi_k).$$

The Pearson correlation coefficient between these components is

$$R(\mathbf{x}_a^{(i)}, \mathbf{x}_b^{(k)}) = \frac{\int_0^\infty e^{-(\sigma_i+\sigma_k)t}\cos(\omega_i t)\cos(\omega_k t)dt}{\sqrt{\int_0^\infty e^{-2\sigma_i t}\cos^2(\omega_i t)dt}\sqrt{\int_0^\infty e^{-2\sigma_k t}\cos^2(\omega_k t)dt}}$$

The above equation has a closed-form solution and it can be proven that

$$R(\mathbf{x}_a^{(i)}, \mathbf{x}_b^{(k)}) = \frac{2\sqrt{\sigma_i\sigma_k}}{\sigma_i+\sigma_k}\frac{\frac{1}{1+\left(\frac{\omega_i-\omega_k}{\sigma_i+\sigma_k}\right)^2}+\frac{1}{1+\left(\frac{\omega_i+\omega_k}{\sigma_i+\sigma_k}\right)^2}}{\sqrt{1+\frac{1}{1+\left(\frac{\omega_i}{\sigma_i}\right)^2}}\sqrt{1+\frac{1}{1+\left(\frac{\omega_k}{\sigma_k}\right)^2}}} \tag{17}$$

The Pearson correlation coefficient between the two spectral components can now be computed easily in terms of their damping rates and frequencies. Note that the correlation coefficient is independent of the amplitudes of the components because of the variance normalization.

We can see now that in addition to providing easy time alignment and data extrapolation of gene expression time series, spectral representation also makes the computation of correlation between two time series very efficient. These properties have useful applications to data matching and retrieval, which will be discussed in the next section.

5 Data Retrieval Based on Spectral Information

The spectral correlation technique and the fast algorithm discussed above have useful applications in gene expression time series analysis. Here we focus on the problem of data comparison and retrieval. An important scenario is that we are given a time series and need to compare it with those stored in a database and find all similar ones. To make the time series search efficient, we can decompose each time series in the database into spectral components and store their parameters. For the given time series, we compute its spectral parameters as well. Then to find similar time series in the database, we only need to compute the correlation coefficients based on Equation (17) and the similarity measure based on Equation (13).

The time series comparison and retrieval procedure based on spectral information discussed above is applicable to data obtained with different sampling rates and different sequence lengths. For example, assume that $x_a[n]$ and $x_b[n]$ have N_a and N_b samples and are obtained with sampling rates Δt_a and Δt_b respectively. We cannot compute the Pearson correlation coefficient between these two time series based on Equation (1) directly. However, different expression data sampling schema do not pose a problem for computing the spectral correlation. Firstly, although the time series lengths (N_a and N_b) affect the accuracy of spectral parameter estimation, they are irrelevant to $R\left(\mathbf{x}_a^{(i)}, \mathbf{x}_b^{(k)}\right)$ in Equation (17). That is, we only need the spectral parameters themselves, not the data lengths, to compute the correlation between two spectral components. Secondly, we can normalize the spectral parameters to take into account of different sampling rates. In this case, we modify Equation (17) as

$$R\left(\mathbf{x}_a^{(i)}, \mathbf{x}_b^{(k)}\right) = \frac{2\sqrt{\sigma_i \Delta t_a \sigma_k \Delta t_b}}{\sigma_i \Delta t_a + \sigma_k \Delta t_b} \frac{\dfrac{1}{1+\left(\dfrac{\omega_i \Delta t_a - \omega_k \Delta t_b}{\sigma_i \Delta t_a + \sigma_k \Delta t_b}\right)^2} + \dfrac{1}{1+\left(\dfrac{\omega_i \Delta t_a + \omega_k \Delta t_b}{\sigma_i \Delta t_a + \sigma_k \Delta t_b}\right)^2}}{\sqrt{1+\dfrac{1}{1+\left(\dfrac{\omega_i}{\sigma_i}\right)^2}}\sqrt{1+\dfrac{1}{1+\left(\dfrac{\omega_k}{\sigma_k}\right)^2}}} \tag{18}$$

We need to select two key parameters in implementing the spectral correlation algorithm. The first parameter is K, the number of spectral components in a time series. The second one is α_T, the threshold for spectral component magnitude. For K

spectral components, the order of linear prediction (the number of prediction coefficients or the number of variables) in Equation (7) must be $2K$. The number of linear prediction equations is then $N - 2K$. When the number of equations equals to the number of variables, that is, $N - 2K = 2K$, or

$$K = \frac{N}{4} \tag{19}$$

Equation (7) has a unique solution. In practice, because of noise, we often want the number of equations to be greater than the number of variables and the equations can be solved using the least squares error (LSE) algorithm. On the other hand, we want K to be as large as possible. A large prediction order will produce so-called extraneous poles and make the AR model more accurate [7-9]. In practice, we cannot choose a large K because N is often small. For example, $N = 18$ for the alpha dataset and $N = 17$ for the cdc28 dataset respectively [2,3,11]. We choose $K = 4$ for both datasets. For the amplitude threshold for each time series, we choose

$$\alpha_T = 0.05\alpha_{\max} \tag{20}$$

where $\alpha_{\max}$ is the maximum spectral component amplitude for the time series.

We summarize below the steps needed for data retrieval:

(1) Use Equation (7) to estimate σ_k and ω_k, and Equation (6) to estimate α_k and φ_k for each gene expression time series in the database and add these parameters to the database.

(2) For an input time series, estimate the same set of parameters, σ_i, ω_i, α_i and φ_i.

(3) Set the phases of all spectral components to zero.

(4) Remove all spectral components with amplitudes α_k and α_i less than α_T.

(5) For each spectral component i of the input gene expression time series, for each gene expression time series in the database, and for each spectral component k of the time series in the database, compute the correlation coefficient based on Equation (17) or (18).

(6) Compute the spectral similarity between the input gene expression time series and a time series from the database based on Equation (13). That is, find the maximum correlation value among all pairs of spectral components between the input time series and the time series from the database.

(7) Accept a time series from the database as a similar one to the input time series if the similarity measure in Equation (13) is greater than 0.5.

Repeat Steps (5) to (7) until the input time series is compared with all time series in the database.

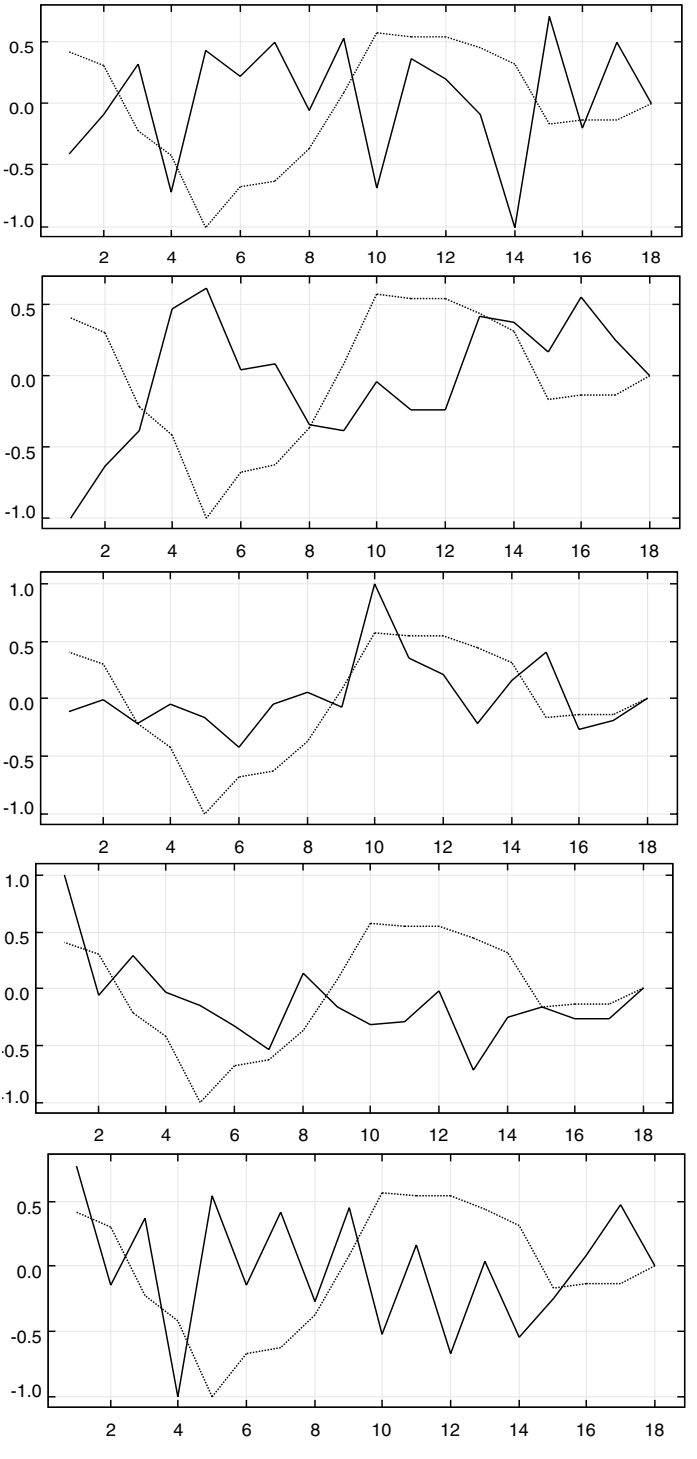
6 Experiment Results

We have used the regulatory gene pairs from [11] to carry out data comparison and retrieval experiments. The gene pairs are extracted by a biologist from the Cho and Spellman alpha and cdc28 datasets [2,3,11]. The purpose of our experiments is to test our spectral similarity metric for data retrieval. A retrieval works well if for a given input gene expression time series, expression time series from the genes that have regulatory relations with the input gene are retrieved. We consider two genes have a regulatory relation if the spectral similarity between their expression time series is greater than 0.5.

Two examples are shown in Table 1. In the first experiment, we use gene YGL055W in the alpha dataset as an input. All genes with regulatory relations with the input gene, which are reported in [11], are retrieved, that is, the spectral similarity measure defined in Equation (13) is greater than 0.5 for expression time series of all retrieved genes. In the second experiment, we use YFL026W in the cdc28 dataset as an input gene and can also retrieve all genes that have regulatory relations with the input gene. We can see clearly from Table 1 that none of the genes would be retrieved if we use the traditional Pearson correlation coefficient as the similarity metric for data comparison and retrieval. The gene pairs from the two examples are shown in Figures 2 and 3 respectively. To display the gene expression waveforms clearly, we have normalized the maximum data amplitude of each time series to 1.

Table 1. List of regulatory gene pairs from alpha and cdc28 datasets retrieved in two of our experiments

	Regulatory Gene Pair		Pearson Correlation Coefficient	Spectral Component Similarity
Alpha dataset	YGL055W	YBL021C	0.400982	0.984977
	YGL055W	YGL237C	0.352143	0.953599
	YGL055W	YGL035C	0.371153	0.952126
	YGL055W	YKL109W	0.264123	0.949755
	YGL055W	YLR256W	0.145242	0.943312
	YGL055W	YOR358W	0.156708	0.921669
	YGL055W	YPR065W	0.115811	0.878227
	YGL055W	YCR084C	0.495500	0.847377
	YGL055W	YNL052W	0.253532	0.527289
CDC28 dataset	YFL026W	YBL016W	0.035268	0.948141
	YFL026W	YHR084W	0.345209	0.918485
	YFL026W	YBR049C	0.379173	0.858287
	YFL026W	YCR084C	0.089666	0.845108
	YFL026W	YMR043W	0.114275	0.839648
	YFL026W	YOL004W	0.157451	0.736196


0.5
0.0
-0.5
-1.0
2 4 6 8 10 12 14 16 18
0.5
0.0
-0.5
-1.0
2 4 6 8 10 12 14 16 18
1.0
0.5
0.0
-0.5
-1.0
2 4 6 8 10 12 14 16 18
1.0
0.5
0.0
-0.5
-1.0
2 4 6 8 10 12 14 16 18
0.5
0.0
-0.5
-1.0
2 4 6 8 10 12 14 16 18

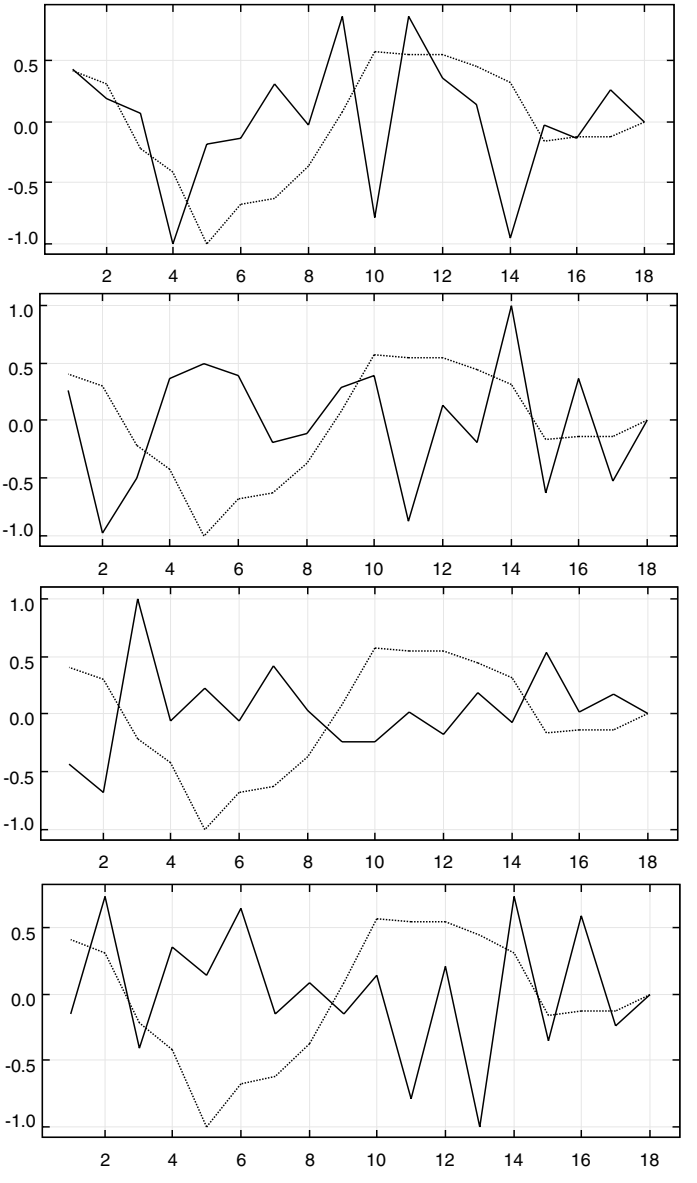
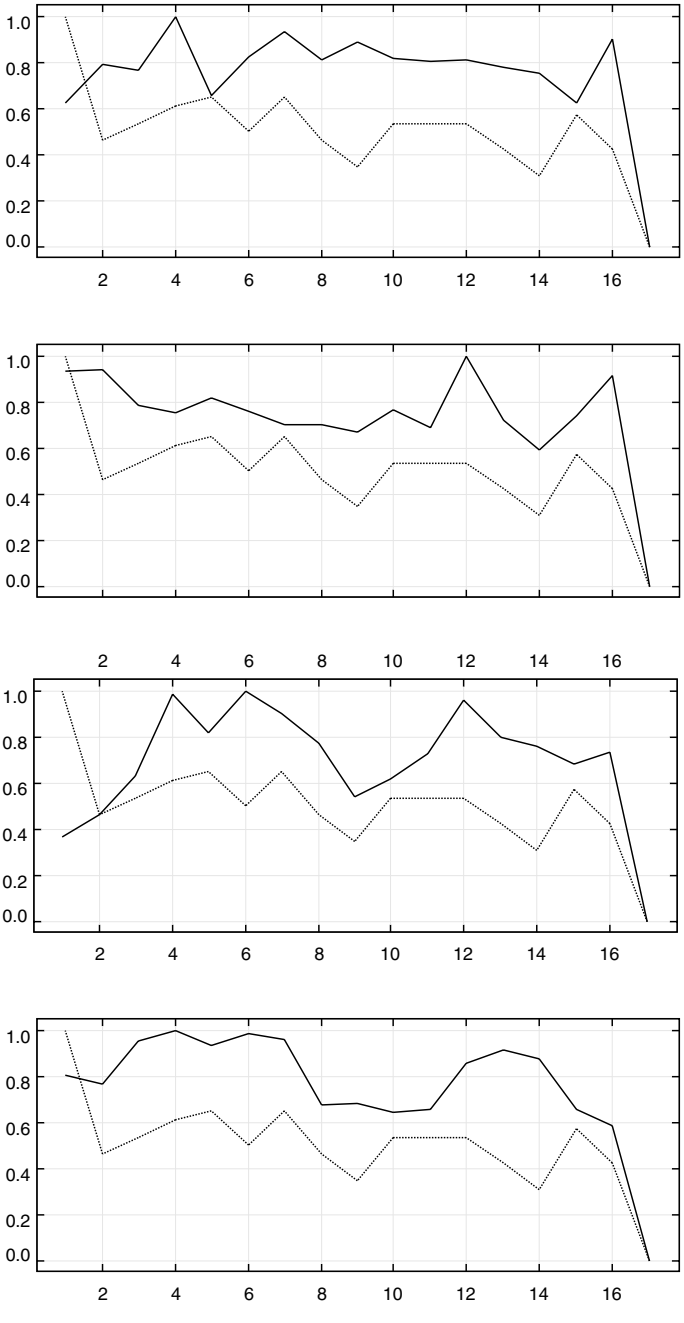


Fig. 2. Gene regulatory pairs from the alpha dataset in Experiment 1. The dashed line in each diagram represents the input expression time series from gene YGL055W and the solid line represents expression time series from a retrieved gene. From first to the ninth row, the retrieved genes are YBL021C, YGL237C, YGL035C, YKL109W, YLR256W, YOR358W, YPR065W, YCR084C, and YNL052W


1.0
0.8
0.6
0.4
0.2
0.0
2
4
6
8
10
12
14
16

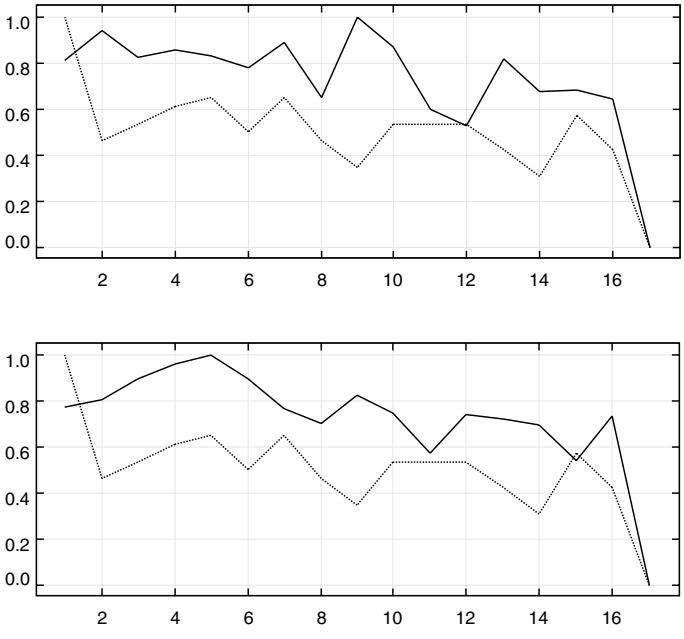


Fig. 3. Gene regulatory pairs from the cdc28 dataset in Experiment 2. The dashed line in each diagram represents the input expression time series from gene YFL026W and the solid line represents expression time series from a retrieved gene. From the first to the sixth row, the retrieved genes are YBL016W, YHR084W, YBR049C, YCR084C, YMR043W and YOL004W

Phase differences and noise fluctuations can be seen clearly in many gene expression time series in Figures 2 and 3 and it is not surprising why the Pearson correlation coefficient does not work well for these gene pairs. However, the spectral components do capture the shape similarity between the time series of these gene pairs. The expression time series of gene YLR256W (middle diagram in the second row of Figure 2) is very noisy, but it still has an overall waveform similarity with the time series of YGL055W. The spectral component similarity is robust enough to detect such similarity despite the noise.

We should also point out here that the spectral component similarity measure can also be greater than 0.5 for gene pairs not present in the regulatory pair list found in [11]. Three such examples are, genes YGL055W and YJL157C with a spectral component similarity measure of 0.977534, genes YGL055W and YJR086W with a similarity measure of 0.821296, and genes YGL055W and YBR279W with a similarity measure of 0.669940, whose expression time series are shown in Figure 4. It is not clear whether these gene pairs definitely have or do not have regulatory relations, or whether the large similarity values are false positives. The Pearson correlation coefficients for these gene pairs are 0.602023, 0.222435 and 0.068640 respectively. It is interesting to note YGL055W and YJL157C look very similar and even the Pearson correlation coefficient between them is greater than 0.5. Ignoring phase reversal in the second case (YFL026W and YJR086W) and slight phase difference in the third case (YFL026W and YBR279W), we can also observe the overall shape similarity between the time series of the gene pairs.

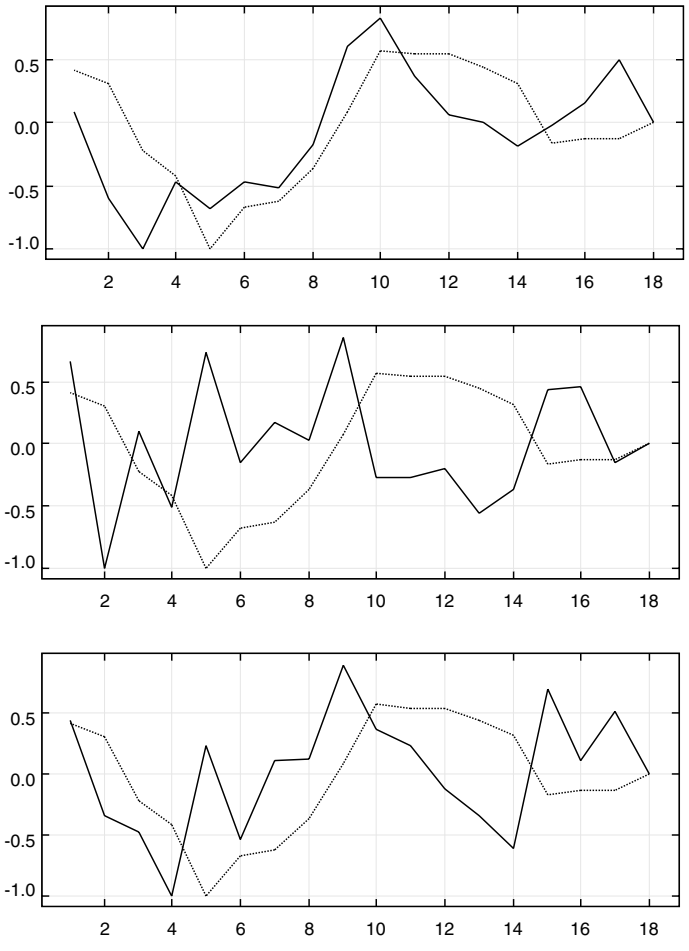


Fig. 4. Examples of gene pairs that are not in the regulatory pair list in [11] but with spectral component similarity measures greater than 0.5. The dashed line in each diagram represents the input expression time series from gene YFL026W and the solid lines from the first to the third row represent expression time series from YJL157C, YJR086W and YBR279W respectively

7 Conclusions

Designing a reliable similarity metric is a basic and crucial problem in gene expression time series analysis. In this paper, we focus on gene expression time series comparison and retrieval. We have shown that the commonly used Pearson correlation coefficient does not work well with gene expression time series due to noise and phase shift of the waveforms. In our spectral component correlation based method, we can estimate the parameters of all spectral components and can set the phase of each component to zero before computing the correlation. This effectively solves the phase shift problem. We have also demonstrated that the proposed method can reduce the effect of noise and detect the overall similarity between two gene expression time

series. The correlation between two spectral components can be computed based on a closed-form mathematical equation and the algorithms can be implemented very efficiently. The spectral correlation can also be easily computed between two time series which are obtained using different sampling rates and have different lengths. The spectral similarity measure can also be useful for other important gene expression analysis tasks, such as data clustering and genetic network inference [4,5,16-18].

Acknowledgement: This work is supported by an interdisciplinary grant (Project 9010003) and a strategic research grant (Project 7001706) from City University of Hong Kong.

References

1. Bar-Joseph Z.: Analyzing Time Series Gene Expression Data. Bioinformatics 20 (2004) 2493–2503
2. Cho, R. J., Campbell, M. J., Winzeler, E. A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T. G., Gabrielian, A. E., Landsman, D., Lockhart, D. J., Davis, R. W.: A Genome-Wide Transcriptional Analysis of the Mitotic Cell Cycle. Mol. Cell 2 (1998) 65-73
3. Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D., Futcher, B.: Comprehesive Identification of Cell Cycle-Regulated Genes of The Yeast Saccharomyces Cerevisiae by Microarray Hybridization. Mol. Biol. Cell 9 (1998) 3273-3297
4. D'haeseleer, P., Liang, S., Somogyi, R.: Genetic Network Inference: From C-Expression Clustering To Reverse Engineering. Bioinformatics, 16 (2000) 707-726
5. De Jong, H.: Modeling and Simulation ff Genetic Regulatory Systems: A Literature Review. J. Computational Biology 9 (2002) 67-103
6. Draghici, S.: Data Analysis Tools for DNA Microarrays. Chapman and Hall New York (2003)
7. Marple, S. L.: Digital Spectral Analysis, with Applications. Prentice Hall NJ (1982)
8. Stoica, P., Moses, R.: Introduction to Spectral Analysis. Prentice Hall NJ (1997)
9. Yan, H. (ed.): Signal Processing in Magnetic Resonance Imaging and Spectroscopy. Marcel Dekker New York (2002)
10. 7Aach J., Church, G. M.: Aligning Gene Expression Time Series with Time Warping Algorithms. Bioinformatics 17 (2001) 495-508
11. Filkov, V., Skiena, S., Zhi, J.: Analysis Techniques for Microarray Time-Series Data. J. Computational Biology 9 (2002) 317-330
12. Kwon, A. T., Hoos, H. H., Ng, R.: Inference of Transcriptional Regulation Relationships from Gene Expression Data. Bioinformatics 19 (2003) 905-912
13. Ji, L., Tan, K. L: Mining Gene Expression Data for Positive and Negative Co-Regulated Gene Clusters. Bioinformatics 20 (2004) 2711-2718
14. Ji, L., Tan, K. L: Identifying Time-Lagged Gene Clusters on Gene Expression Data. Bioinformatics 21 (2005) 509-516
15. Yeung, L. K., Szeto, L. K., Liew, A. W. C., Yan, H.: Dominant Spectral Component Analysis for Transcriptional Regulations Using Microarray Time-Series Data. Bioinformatics 20 (2004) 742-749

16. Szeto, L. K., Liew, A. W. C., Yan, H., Tang, S. S.: Gene Expression Data Clustering And Visualization Based on a Binary Hierarchical Clustering Framework. J. Vis. Lang. Computing 14 (2003) 341-362
17. Wu, S., Liew, A. W. C., Yan, H., Yang, M.: Cluster Analysis of Gene Expression Data Based on Self-Splitting and Merging Competitive Learning. IEEE Trans. Infor. Tech. Biomed. 8 (2004) 5-15
18. Liew, A. W. C., Yan, H., Yang, M., Chen Y. P.: Microarray Data Analysis. In: Chen, Y. P. (ed.): Bioinformatics Technologies. Springer-Verlag Berlin (2005) 353-388

SVM Classification to Predict Two Stranded Anti-parallel Coiled Coils Based on Protein Sequence Data

Zhong Huang[1], Yun Li[2], and Xiahohua Hu[2]

[1] Kimmel Cancer Center, Thomas Jefferson University,
Philadelphia, PA, USA
[2] College of Information Science and Technology, Drexel University,
3141 Chestnut Street, Philadelphia, PA, USA, 19104
thu@cis.drexel.edu

Abstract. Coiled coils is an important 3-D protein structure with two or more stranded alpha-helical motif wounded around to form a "knobs-into-holes" structure. In this paper we propose an SVM classification approach to predict the two stranded anti-parallel coiled coils structure based on the primary amino acid sequence. The training dataset for the machine learning are collected from SOCKET database which is a SOCKET algorithm predicted coiled coils database. Total 41 sequences of at least two heptad repeats of the two stranded anti-parallel coiled coils motif are extracted from 12 proteins as the positive datasets. Total 37 of non coiled coils sequences and two stranded parallel coiled coils motif are extracted from 5 proteins as negative datasets. The normalized positional weight matrix on each heptad register a, b, c, d, e, f and g is from SOCKET database and is used to generate the positional weight on each entry. We performed SVM classification using the cross-validated datasets as training and testing groups. Our result shows 73% accuracy on the prediction of two stranded anti-parallel coiled coils based on the cross-validated data. The result suggests a useful approach of using SVM to classify the two stranded anti-parallel coiled coils based on the primary amino acid sequence.

Keywords: coiled coil, SOCKET algorithm, SVM, protein sequence data.

1 Introduction

Coiled coils structure was first introduced by Crick in 1953 in which he postulated a hallmark structure of "knobs-into-holes" formed by wounded strands of alpha-helices [2]. The coiled coils structure is characterized by a heptad repeats of amino acids (*a-b-c-d-e-f-g*)n. Positions *a* and *d* in one chain are occupied by apolar hydrophobic amino acids to form the core packing structure with the same positions in partner chain. The coiled coils structure is further stabilized by side chain electrostatic interaction of *e-g* between two chains which generally occupied by polar charged amino acids. Recently it has been shown that intrachain interactions between heptad residues also contribute to the stability of the coiled coils structure. Figure 1 shows schematic representation of the two stranded coiled coils.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 374 – 380, 2005.

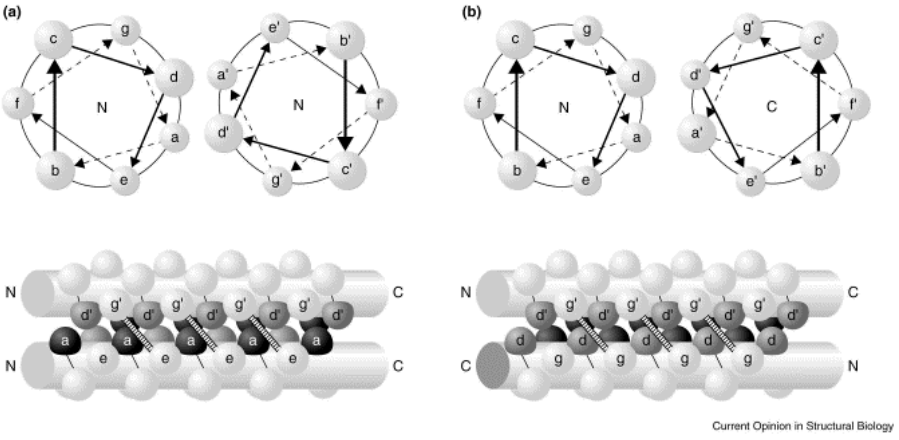


Fig. 1. Parallel (a) and anti-parallel (b) two stranded coiled coils. The heptad repeat are shown as a to g. (figure are from [6]) (Oakley and Hollenbeck 2001)

Due to its well characterized structure the coiled coils has long been a spotlight of the protein design and prediction study. However, the structure of coiled coils is of great diversity in terms of its interchain orientation and oligomer status. The coiled coils structure can be formed between two, three, four or even five chains and the orientation of each chain can be the same (parallel) or different (antiparallel). The core packing registers *a* and *d* are important for determining the number of strands while *e-g* interaction seems to be important in choosing the helices partners [3][4][7][11]. Therefore the primary sequence of the heptad repeats may be one of the determining factors on the specificity of the coiled coils but much is still poorly understood so far [12].

Two categories of algorithms have been proposed to predict the coiled coils structure based on either the primary amino acid sequence or the atomic 3-D coordinate information. The first category includes COILS [5], PAIRCOIL [1] and MULTICOIL [10]. These algorithms compare the sequence of the target protein with the amino acid sequence database of known two or three stranded parallel coiled coils and give the score of probability. However, none of them are suitable for predicting the anti-parallel coiled coils. The second category of prediction algorithm takes into consideration of 3-D coordinate information of the polypeptide and compares it with the database of known 3-D structure of coiled coils. SOCKET [9] and TWISTER [8] are two algorithms in this category. The SOCKET algorithm focuses on the core packing structure of the "knobs-into-holes" which is formed by interchain *a-d* interactions. The TWISTER algorithm is designed to identify not only the canonical coiled coils but also the special coiled coils with discontinuous heptad repeats interrupted by stutters and skips. Both algorithms take the 3-D atomic coordinate PDB file and DSSP file as input and are able to predict two or three stranded parallel and anti-parallel coiled coils. Comparing with the first category of algorithms, using 3-D coordinate as input may seem to be a better choice as the SOCKET algorithm attempts to identify the core packing structure of coiled coils based on the experimentally determined protein 3-D structure. However, it also impeded the application of the algorithm on

the majority of proteins due to the fact that only little greater than 6000 proteins have 3-D coordinate information determined by x-ray crystallization and NMR spectroscopy.

In this paper we used the machine learning SVM approach to discriminate the two stranded anti-parallel coiled coils structure based on the primary amino acid sequence using the normalized amino acid profile of heptad repeat generated by SOCKET [9]. We selected two stranded anti-parallel coiled coils proteins with at least two full set of heptad repeats from SOCKET database such that each vector has the same number of features for SVM training and testing. Our preliminary results suggest that SVM is a valuable tool to predict the two stranded anti-parallel coiled coils based on the amino acid profile originally determined by atomic 3-D coordinate.

2 Methods, Results and Discussion

We selected total 78 two stranded anti-parallel coiled coils from SOCKET database which currently lists total 134 entries [9] based on PDB release #89. The PDB files of 78 two stranded anti-parallel coiled coils and 8 two stranded parallel coiled coils were downloaded from PDB database using a perl script available from PDB ftp site. The PDB files from both the two stranded anti-parallel and parallel coiled coils were submitted to SOCKET server (http://www.biols.susx.ac.uk/Biochem/Woolfson/html/coiledcoils/socket/server.html) to identify the specific heptad repeats registers. Only the long coiled coils with full set of the heptad repeats larger than 2 were selected. This is mainly because we assume long coiled coils are more structurally stable and may include more positional information which may contributes to the stability and specificity of coiled coils. Considering the relatively small number of entries we obtained, we allow multiple contributions of two stranded coiled coils structure from the same protein. Because the SVM only accepts vectors with the same number of features, we chose 2 heptad repeats of total 14 amino acids from each entry. In the case of heptad repeats are more than 2, we allow partial overlap of the heptad repeats assuming each partially overlapped heptad repeats is an independent vector for SVM to avoid loss of any given heptad repeat. Total 41 sequences of at least two heptad repeats of the two stranded anti-parallel coiled coils motif are extracted from 12 proteins as the positive datasets. Total 37 of non coiled coils sequences and two stranded parallel coiled coils motif are extracted from 5 proteins as negative datasets.

We used the normalized amino acid profile for two stranded long anti-parallel coiled coils from Walshaw et al [9] to convert the amino acid into amino acid usage frequencies (shown in Table 1).

Two programs, namely score14.pl written in perl and CC2SMatrix.java written in java, were used to convert the list of 14 amino acid sequence into amino acid usage frequency at each position. Due to the small number of samples, we adopted cross-validation approach for SVM training and testing. The total 78 data sets are separated into two groups, with 77 data sets for training and 1 data set for testing in each cycle. The testing results are shown in Table 2.

Our results show that the average accuracy for the testing is 73% (summarized in Table 3).

Table 1. Normalized amino acid profile for two stranded long anti-parallel coiled coils from Walshaw et al [9]

Corner	A	B	C	D	E	F	G
A	1.42	1.48	1.28	1.55	0.69	1.04	1.98
C	0	0	0	0.59	0	1.19	0
D	0.09	1.06	1.71	0.56	1.37	1.75	0.25
E	0.54	1.32	1.75	1.31	1.45	1.76	1.23
F	0.84	0.51	0.85	0.96	0.8	0.96	0.64
G	0	0.72	0.81	0.07	0.48	0.38	0.19
H	1.09	1.56	0.31	0.44	0.88	1.47	3.79
I	2.87	0.48	0.24	1.6	1.59	0.79	1.12
K	0.33	2.24	1.52	0.41	1.22	1.55	0.66
L	2.96	1.04	1.33	3.64	1.31	0.35	1.11
M	1.45	1.18	0.59	1.65	0.83	0.83	1.93
N	0.88	1.58	0.94	0.22	0.44	1.78	0.59
P	0	0	0.28	0	0.27	0	0
Q	0.86	1.23	1.75	0.74	1.99	1.16	2.3
R	0.66	1.36	1.48	0.47	2.55	1.78	1.14
S	1.03	0.78	0.58	0.28	0.37	0.83	0.83
T	0.61	1.23	0.73	0.61	0.58	1.39	1.15
V	0.51	0.21	0.53	0.37	0.61	0.41	0.99
W	0	0.56	2.24	0.31	0.53	0.53	0.53
Y	1.23	0.66	0.44	1.69	0.82	0.41	0.2

Table 2. SVM Prediction Results Using Cross-Validation Approach

Corner	Class Label	Prediction by SVM
1a36-1	1	1
1a36-2	1	-1
1ab4-1	1	1
1ab4-2	1	-1
1aqt-1	1	-1
1cii-1	1	1
1cii-2	1	1
1cii-3	1	1
1cii-4	1	-1
1cii-5	1	1
1cii-6	1	-1
1cii-7	1	-1
1cii-8	1	-1

Table 2. (*Continued*)

1cnt-1	1	1
1cnt-2	1	1
1cnt-1	1	1
1cnt-2	1	1
1cnt-3	1	1
1cnt-4	1	1
1cnt-5	1	1
1ecm-1	1	1
1ecm-2	1	1
1ecm-3	1	1
1ecm-4	1	1
1ecr-1	1	1
1ecr-2	1	1
1ser-1	1	1
1ser-2	1	-1
2fha-1	1	1
2fha-2	1	-1
2fha-3	1	-1
2fha-4	1	1
2fha-5	1	1
2ktq	1	1
2spc-1	1	1
2spc-2	1	1
2spc-3	1	1
2spc-4	1	1
2spc-5	1	1
2spc-6	1	1
5eau	1	1
m6a-1	-1	-1
m6a-2	-1	-1
m6a-3	-1	-1
m6a-4	-1	-1
m6a-5	-1	-1
m6a-6	-1	1
m6a-7	-1	-1
m6a-8	-1	-1
m6a-9	-1	1
m6a-10	-1	-1
m6a-11	-1	1
m6a-12	-1	-1
m6a-13	-1	1

Table 2. (*Continued*)

m6a-14	-1	-1
m6a-15	-1	-1
1a93-1	-1	-1
1a93-2	-1	-1
1a93-3	-1	1
1a93-4	-1	1
1fos-1	-1	-1
1fos-2	-1	-1
1fos-3	-1	1
1fos-4	-1	-1
1fos-5	-1	-1
1fos-6	-1	-1
1fos-7	-1	1
1fos-8	-1	1
1fos-9	-1	-1
1fos-10	-1	-1
mbplike-1	-1	1
mbplike-2	-1	-1
mbplike-3	-1	-1
mbplike-4	-1	-1
mbplike-5	-1	-1
mbplike-6	-1	1
mbplike-7	-1	-1
MBP	-1	-1

Table 3. Average Accuracy Result

Total Dataset	78
Positive	41
Negative	37
Testing Result	
False Positive	11
False Negative	10
True Positive	31
True Negative	26
Average Accuracy	73%

3 Conclusion and Future Plan

In this paper we present a SVM approach for classifying two stranded antiparallel coiled coils structure based on primary sequences. The classification results indicate

that the support vector machine is a useful tool in classifying the antiparallel two stranded coiled coils structure, which by far has no direct algorithm to predict its structure based on its primary sequence. However, more data sets are needed to further validate the approach. With the rapid growing of the PDB database, the SOCKET database has been growing accordingly. We expect the future release of SOCKET database will be helpful in gathering more positive data sets and help to increase the SVM prediction accuracy.

References

1. Berger, B. (1995). "Algorithms for protein structural motif recognition." J Comput Biol 2(1): 125-38.
2. Crick, F. H. C. (1953). "The packing of alpha-helices: simple coiled-coils." Acta. Crystallog. 6: 689-697.
3. Harbury, P. B., T. Zhang, et al. (1993). "A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants." Science 262(5138): 1401-7.
4. Kohn, W. D., C. M. Kay, et al. (1998). "Orientation, positional, additivity, and oligomerization-state effects of interhelical ion pairs in alpha-helical coiled-coils." J Mol Biol 283(5): 993-1012.
5. Lupas, A., M. Van Dyke, et al. (1991). "Predicting coiled coils from protein sequences." Science 252(5010): 1162-4.
6. Oakley, M. G. and J. J. Hollenbeck (2001). "The design of antiparallel coiled coils." Curr Opin Struct Biol 11(4): 450-7.
7. O'Shea, E. K., R. Rutkowski, et al. (1992). "Mechanism of specificity in the Fos-Jun oncoprotein heterodimer." Cell 68(4): 699-708.
8. Strelkov, S. V. and P. Burkhard (2002). "Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation." J Struct Biol 137(1-2): 54-64.
9. Walshaw, J. and D. N. Woolfson (2001). "Socket: a program for identifying and analysing coiled-coil motifs within protein structures." J Mol Biol 307(5): 1427-50
10. Wolf, E., P. S. Kim, et al. (1997). "MultiCoil: a program for predicting two- and three-stranded coiled coils." Protein Sci 6(6): 1179-89
11. Woolfson, D. N. and T. Alber (1995). "Predicting oligomerization states of coiled coils." Protein Sci 4(8): 1596-607
12. Yu, Y. B. (2002). "Coiled-coils: stability, specificity, and drug delivery potential." Adv Drug Deliv Rev 54(8): 1113-29.

Estimating Gene Networks with cDNA Microarray Data Using State-Space Models

Rui Yamaguchi[1], Satoru Yamashita[2], and Tomoyuki Higuchi[2]

[1] Faculty of Mathematics, Kyushu University, 6-10-1 Hakozaki,
Higashi-ku, Fukuoka, 812-8581, Japan
ruiy@math.kyushu-u.ac.jp
[2] The Institute of Statistical Mathematics,
4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan
{ysatoru, higuchi}@ism.ac.jp

Abstract. State-space model is used in this paper to analyze dynamics of gene expression profile data. State-space models describe dynamics that the observed measurements depend on some hidden state variables. Hidden state variables can capture effects that cannot be measured in gene expression profiling experiment, for example, genes that have not been included in the observation variables, levels of regulatory proteins or the effect of mRNA. System identification is achieved by EM algorithm that is based on the maximum likelihood method. We apply this method to a published yeast cell-cycle gene expression time-series data under the assumption that state variables and observation variables are generated by Gaussian white noise process, that produces the simplest and most reasonable model to explain of behaviors of the gene expression profiles.

1 Introduction

It has become possible to measure gene expression levels on genomic scale. Although many details inside a cell are not precisely known, gene expression data on a genomic scale provides useful insights into a living cell. Data thus collected enhance fundamental understanding of life on the molecular level, and may prove useful also in medical diagnosis, treatment, and drug design. Using gene expression data, a wide variety of models, such as Boolean networks, differential equations and linear and non-linear auto-regression models have been introduced to model cellular systems. Many of these published models can be considered to be special cases of a general class of graphical models known as dynamic Bayesian networks. Dynamic Bayesian networks are suitable for modeling gene expression data, because they can handle noisy or missing data, or handle hidden variables such as protein levels hat may have an effect on mRNA expression levels. Although microarray technologies have made it possible to measure time series of the expression level of many genes simultaneously, we cannot hope to measure all possible factors contributing to genetic regulatory interactions. As well as this, gene expression data are known to include missing and outliers. Therefore the ability of Bayesian networks to handle such hidden variables would appear to be one of the main advantages as a modeling tool. The state-space models

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 381 – 388, 2005.

are in a class of dynamic Bayesian networks that assume that the observed measurements depend on some hidden state variables that evolve according to Markovian dynamics. In a state-space model, the observation variables typically depend on the hidden state variables, while the change in the internal state variables is completely determined by the current internal state variables.

Linear-Gaussian state-space models are known as linear dynamical systems or Kalman filter models, and have been used extensively in many areas of control and signal processing. In this paper we assumed that state variables and observation variables are generated by Gaussian white noise process, and a linear state-space model for gene regulatory networks, in which genes are viewed as the output observation variables and gene expression dynamics is governed by a group of the internal state variables, is employed. The dimensionality of the internal state variables will be determined by Akaike Information Criterion (AIC). System is identified by the EM algorithm on the basis of the maximum likelihood method. Observation variables are time series data of cDNA microarray dataset for yeast cell-cycle regulated genes obtained by Spellman et al. Wu et al. analyzed the same data set using factor analysis, but they obtained internal state variables and state transition matrix separately. Yukinaga et al. also analyzed the same data set. They used a linear-Gaussian state-space model and introduced prior distribution to state transition matrix and state-to-observation transition matrix. The dimensionality of the internal state variables was determined by Bayesian Information Criterion (BIC) in both of them works. The system is identified in this paper without any prior distributions and is identified self consistently through a state-space model. The difference between the result from AIC and that from BIC is clarified quantitatively.

2 Systems and Methods

2.1 Linear-Gaussian State-Space Models

In this paper, a sequence of *l*-dimensional observation vectors is modeled by assuming that at each time step, an observation vector y_n is generated from a *k*-dimensional hidden state vector x_n. A linear-Gaussian state-space model can be described as follows.

$$x_n = Fx_{n-1} + v_n ;\ n = 0,1,\Lambda\ ,N \tag{1}$$

$$y_n = Hx_n + w_n ;\ n = 1,2,\Lambda\ ,N \tag{2}$$

$$x_0 \sim N_k(x_0 \mid \mu, \Sigma) \tag{3}$$

$$v_n \sim N_k(v_n \mid 0_k, Q) \tag{4}$$

$$w_n \sim N_l(w_n \mid 0_l, R) \tag{5}$$

Here F is $k*k$ matrix and H is $l*k$ matrix. N is the length of the observed time series. The initial state vector x_0 is assumed to be a Gaussian random vector with mean vector μ and covariance matrix Σ. The error or noise term v_n is assumed to be a 0-mean uncorrelated Gaussian distributed noise vector with common covariance matrix Q. The observation noise term w_n is also 0-mean uncorrelated Gaussian distributed vector with common covariance matrix R. F, H, Q, R are assumed to be time independent in this paper. If one knows the values for the parameters μ, Σ, F, H, Q and R, the conventional Kalman smoothing estimator can be calculated as conditional expectations and will have minimum mean square error.

It was already mentioned in "Introduction'' section that state-space models have many attractive features when analyzing gene expression data. The first purpose of this paper is to introduce procedures to identify the dimensionality of a state vector ; k and to identify the values of μ, Σ, F, H, Q and R. The next one is to discuss the validity of the conventional Kalman smoothing estimator $x_{n|N}$ after k, μ, Σ, F, H, Q and R are fixed. The identification of the system is shown in next subsection.

2.2 Maximum Likelihood Estimation with the EM Algorithm

The parameters μ, Σ, F, H, Q and R are now estimated in this subsection. In order to develop a procedure for estimating these parameters in the given state-space model, the joint log likelihood of the complete data $x_0, x_1, \Lambda, x_N, y_1, y_2, \Lambda, y_N$ is the key and can be written in the form

$$\begin{aligned} \log L \equiv & -\frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_0-\mu)'\Sigma^{-1}(x_0-\mu) \\ & -\frac{N}{2}\log|Q| - \frac{1}{2}\sum_{n=1}^{N}(x_n - Fx_{n-1})'Q^{-1}(x_n - Fx_{n-1}) \\ & -\frac{N}{2}\log|R| - \frac{1}{2}\sum_{n=1}^{N}(y_n - Hx_n)'R^{-1}(y_n - Hx_n) \end{aligned} \tag{6}$$

where log L is to be maximized with respect to the parameters μ, Σ, F, H, Q and R. Since the log likelihood given above depends on the unobserved data series x_0, x_1, Λ, x_N, we consider applying the EM algorithm conditionally with respect to the observed series y_1, y_2, Λ, y_N. The estimated parameters at the $(r+1)$st iteration of the values μ, Σ, F, H, Q, R are defined as ones that maximize $E_r(\log L \mid y_1, y_2, \Lambda, y_N)$. Here E_r is the conditional expectation of the log likelihood calculated with the rth iterative estimated parameters of

μ, Σ, F, H, Q, R. It is well known that the log likelihood calculated with the (*r+1*)st iterative estimated parameters is larger than that with the *r*th iterative estimated parameters. Now the conditional expectation E_r is as follows.

$$\begin{aligned} E_r(\log L \mid y_1, y_2, \Lambda, y_N) = & -\frac{1}{2}\log|\Sigma| - \frac{1}{2}tr\{\Sigma^{-1}(V_{0|N} + (x_{0|N} - \mu)(x_{0|N} - \mu)')\} \\ & -\frac{1}{2}\log|Q| - \frac{1}{2}tr\{Q^{-1}(C - BF' - FB' + FAF')\} \\ & -\frac{1}{2}\log|R| - \frac{1}{2}tr\{R^{-1}\sum_{n=1}^{N}[(y_n - Hx_{n|N})(y_n - Hx_{n|N})' \\ & + HV_{n|N}H']\} \end{aligned} \tag{7}$$

$$x_{n|N} = E(x_n \mid y_1, y_2, \Lambda, y_N) \tag{8}$$

$$V_{n|N} = \mathrm{cov}(x_n \mid y_1, y_2, \Lambda, y_N) \tag{9}$$

$$V_{n,n-1|N} = \mathrm{cov}(x_n, x_{n-1} \mid y_1, y_2, \Lambda, y_N) \tag{10}$$

Here *tr* denotes trace and *A*, *B* and *C* denote following.

$$A = \sum_{n=1}^{N}(V_{n-1|N} + x_{n-1|N} \cdot x'_{n-1|N}) \tag{11}$$

$$B = \sum_{n=1}^{N}(V_{n,n-1|N} + x_{n|N} \cdot x'_{n-1|N}) \tag{12}$$

$$C = \sum_{n=1}^{N}(V_{n|N} + x_{n|N} \cdot x'_{n|N}) \tag{13}$$

The values $x_{n|N}, V_{n|N}, V_{n,n-1|N}$ can be computed through Kalman filter with the *r*th iterative estimated parameters of μ, Σ, F, H, Q, R. The (*r+1*)th iterative estimated parameters of μ, Σ, F, H, Q, R can be obtained by calculating following equations.

$$\frac{\partial E_r}{\partial \mu} = 0 \tag{14}$$

$$\frac{\partial E_r}{\partial F} = 0 \tag{15}$$

$$\frac{\partial E_r}{\partial H} = 0 \tag{16}$$

$$\frac{\partial E_r}{\partial Q} = 0 \tag{17}$$

$$\frac{\partial E_r}{\partial R} = 0 \tag{18}$$

We can get equations (19)-(22) as (*r+1*)th iterative estimated parameters except for Σ by solving equations (14)-(18). Σ can be taken as a fixed reasonable value. It is easily confirmed that the estimated parameters maximize the conditional expectation of the log likelihood : E_r.

$$\mu = x_{0|N} \tag{19}$$

$$F = BA^{-1} \tag{20}$$

$$H = (\sum_{n=1}^{N} y_n \cdot x_{n|N}^{'}) \cdot C^{-1} \tag{21}$$

$$Q = \frac{1}{N}(C - BA^{-1}B^{'}) \tag{22}$$

$$R = \frac{1}{N} \sum_{n=1}^{N} [(y_n - Hx_{n|N})(y_n - Hx_{n|N})^{'} + HV_{n|N}H^{'}] \tag{23}$$

The parameters of the state-space equations (1) and (2) can be identified as follows.

1. Select initial values of μ, F, H, Q, R and some reasonable baseline level of Σ. The conventional Kalman smoothing estimators $x_{n|N}, V_{n|N}, V_{n,n-1|N}$ can be recursively calculated with the upper initial parameters.
2. Calculate the conditional expectation of the log likelihood with equations (6) and (7). (E-step)
3. Calculate the equations (19)-(22) and obtain the next iterative estimated parameters that maximize conditional expectation of the log likelihood. (M-step)
4. Insert estimated parameters of μ, F, H, Q, R to the state space equations (1) and (2), and calculate the conventional Kalman smoothing estimators.
5. Repeat the upper procedure until the log likelihood is stable.

The EM algorithm for the maximum likelihood estimation may fall into a local maximum. Therefore the global maximum must be chosen by comparing results between several sets of the initial values of μ, F, H, Q, R.

2.3 Identification of Dimensionality of State Vectors Through AIC

The dimensionality of the state vectors is not yet identified. Akaike Information Criterion (AIC) is introduced in this section in order to solve this problem.

$$AIC = -2 \cdot \log L + 2 \cdot NP \tag{24}$$

Log L is the maximum log likelihood and NP is the number of the parameters estimated in the equations (1) and (2) including the dimensionality of the state vectors; k. The maximum log likelihood increases if the dimensionality of state vectors increases, which may causes over fitting to the observation. The dimensionality of the state vectors that has minimum AIC is accepted in this paper, which avoids over fitting to the observation.

3 Application Results

The proposed methodology was applied to a publicly available cDNA microarray dataset, Spellman et al's experiment for studying yeast cell-cycle regulated genes (CDC-15). The dataset is available at http://www.cellcycle-www.stanford.edu. The whole data set describes 6177 gene expression profiles over 290 minutes with 10 minutes time resolution. A log transformation were applied to all original intensity rations (Cy5/Cy3) of gene expression, and the expression profile for each gene was normalized to have a median 0 and a standard deviation of 1. Spellman et al. identified 800 genes, so we use the 800 genes for the 19 equally spaced time points in this paper. The analysis is now continuing and the application results will be reported at the conference.

4 Discussion

The linear-Gaussian state space model denoted as equations (1) and (2) was in this paper introduced in order to analyze the yeast cell-cycle gene expression time-series data. The system identification was achieved by the EM algorithm and the maximum likelihood method. AIC was used in order to estimate the number of internal hidden state variables. We will show, in this section, the characteristics of these methods thorough the comparison with the past similar studies.

Wu et al. made factor analysis using the same data set and the same equations as equations (1) and (2). From the equation (2), the matrix H and x_n were identified at first by factor analysis through the maximum likelihood method and the EM algorithm. The dimensionality of x_n was obtained by Bayesian Information Criterion (BIC).

$$BIC = -2 \cdot \log L + m \cdot NP \tag{25}$$

Here log L is the log likelihood, m is sample size (number of genes), and NP is the number of estimated parameters. They emphasized that BIC avoids over fitting to the observation more than AIC because BIC takes sample size into account. The matrix F was identified from the equation (1) by least square method after these upper procedures. They treated equations (1) and (2) separately and neglected the Gaussian distribution of v_n and w_n. On the other hand, our study treated equations (1) and (2)

simultaneously as state-space models, therefore our study might have more self consistent solutions.

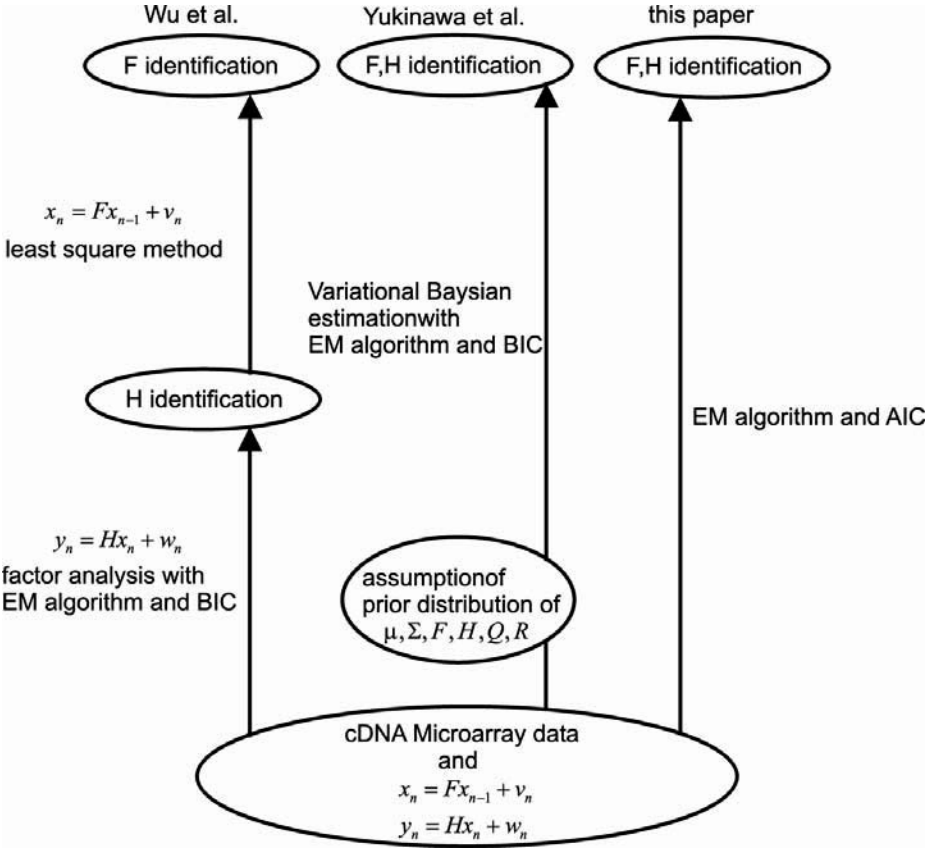


Fig. 1. Schematic drawing of system identification used in Wu et al., Yukinawa et al. and this paper

Yukinawa et al. also made system identification with the same data set and the same equations as (1) and (2) with the variational Bayesian estimation. They also used BIC in order to estimate the dimensionality of the state vector. The variational Bayesian estimation assumes explicitly prior distribution of μ, Σ, F, H, Q, R in order to estimate these parameters. The introduction of the prior distributions avoids instability during estimation of these parameters, although it might cause some problems.

Both of the upper studies may underestimate gene expression dynamics because of the use of BIC. It is interesting to compare the result by AIC with that by BIC quantitatively and to discuss the effect of prior distributions. Schematic drawings of system identification using in each study are seen in Figure 1. The result of our analysis and comparison will be discussed in the conference.

References

1. Rangel, C., Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., Gaiba, A., Wild, D. L. and Falciani, F., Modeling T-cell activation using gene expression profiling and state-space models, *Bioinformatics*, Vol. 20, pp. 1361-1372. (2004)
2. Shumway, R. H. and Stoffer, D. S. : Time series smoothing and forecasting using the EM algorithm, *Technical Report*, Vol. 27, Division of Statistics, University of California, Davis. (1981)
3. Shumway, R. H. and Stoffer, D. S. : An approach to time series smoothing and forecasting using EM algorithm, *Journal of Time Series Analysis*, Vol. 3, pp. 253-264. (1982)
4. Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. and Futcher, B. Comprehensive identification of cell cycle regulated genes of the yeast *Sacchromyces cerevisiae* by microarray hybridization, *Molcular Biology of the Cell*, Vol. 9, pp. 3273-3297. (1998)
5. Wu, F. X., Zhang, W. J. and Kusalik, A. J. : Modeling gene expression from microarray expression data with state-space equations, *Pacific Symposium on Biocomputing*, Vol. 9, pp. 581-592. (2004)
6. Yukinawa, N., J. Yoshimoto, S. Oba and S. Ishii : System identification of gene expression time-series based on a linear dynamical system model with variational Bayesian estimation, (in Japanese). 2003年日本神経回路学会全国大会(2003)

A Penalized Likelihood Estimation on Transcriptional Module-Based Clustering

Ryo Yoshida[1], Seiya Imoto[2], and Tomoyuki Higuchi[1]

[1] Institute of Statistical Mathematics,
4-6-7 Minami-Azabu, Minato-ku, Tokyo, Japan
[2] Human Genome Center, Institute of Medical Science,
University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, Japan

Abstract. In this paper, we propose a new clustering procedure for high dimensional microarray data. Major difficulty in cluster analysis of microarray data is that the number of samples to be clustered is much smaller than the dimension of data which is equal to the number of genes used in an analysis. In such a case, the applicability of conventional model-based clustering is limited by the occurence of overlearning. A key idea of the proposed method is to seek a linear mapping of data onto the low-dimensional subspace before proceeding to cluster analysis. The linear mapping is constructed such that the transformed data successfully reveal clusters existed in the original data space. A clustering rule is applied to the transformed data rather than the original data. We also establish a link between this method and a probabilistic framework, that is, a penalized likelihood estimation of the mixed factors model. The effectiveness of the proposed method is demonstrated through the real application.

1 Introduction

Microarray dataset is a collection of microarray experiments, $\mathbf{x}_j \in \mathbb{R}^d$, $j \in \{1, \cdots, N\}$ in which each experiment represents the expression levels of d genes corresponding to the jth sample. Usually, microarray dataset has a fairly small sample size N, typically less than one hundred, whereas the number of genes involved is more than several thousands. Cluster analysis of microarray has been considered as a challenge to the automated search for molecular subtypes of desease. In view of statistics, major difficulty in this problem is that the number of samples to be clustered is much smaller than that of genes, i.e. $N << d$. This fact limits the applicability of conventional model-based (or distance-based) clustering by the occurence of overlearning. For instance, clustering based on the Gaussian mixture model, which also includes the K-means clustering as a special case, usually leads to the overfitting during the density estimation process with $N << d$. In this article, a new procedure is proposed to overcome such intractability inherent in microarray studies.

The goal of cluster analysis is to partition a set of N samples $\{\mathbf{x}_j\}_{j=1}^N$ into G-nonoverlapping clusters $\{\mathcal{P}_g\}_{g=1}^G$, such that those in a particular cluster are

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 389–401, 2005.

cohesive and separated from those in other clusters. This problem amounts to estimating the vector of G-unknown class labels $\mathbf{c}(\mathbf{x}_j)^T = (c_1(\mathbf{x}_j), \cdots, c_G(\mathbf{x}_j))$, $j \in \{1, \cdots, N\}$:

$$c_g(\mathbf{x}_j) = \begin{cases} 1 & \text{if } \mathbf{x}_j \in \mathcal{P}_g \\ 0 & \text{otherwise.} \end{cases}$$

The estimation of $\{\mathbf{c}(\mathbf{x}_j)\}_{j=1}^N$ is achieved by constructing a suitable classifier $\hat{\mathbf{c}}(\mathbf{x}_j) = \{\hat{c}_g(\mathbf{x}_j)\}_g$ which declares the assignment of the jth sample to the gth cluster by $\hat{c}_g(\mathbf{x}_j) = 1$ and $\hat{c}_h(\mathbf{x}_j) = 0$ for $h \neq g$.

Unfortunately, constructing the clustering rule as defined over $\mathbb{R}^d$ is very hard with $N << d$ as the finite mixture model leads to the overfiting during the density estimation. Reducing the dimension of data, that is construction of a mapping of data onto the low-dimensional subspace, has been considered as a key issue in microarray study. In this article we consider to seek a linear mapping of data onto the low-dimensional subspace, $\mathbf{P}^T\mathbf{x}_j \in \mathbb{R}^q$ as with $q << d$ before proceeding to cluster analysis:

$$\mathbf{P}^T\mathbf{x}_j = \{\mathbf{p}_k^T\mathbf{x}_j\}_k.$$

Here, the $\mathbf{p}_k$ stands for the kth column of $\mathbf{P}$. Then, the corresponding classifier is defined over $\mathbb{R}^q$ rather than $\mathbb{R}^d$:

$$\hat{\mathbf{c}}(\mathbf{x}_j) \equiv \hat{\mathbf{c}}(\mathbf{P}^T\mathbf{x}_j), \quad \mathbf{x}_j \in \mathbb{R}^d.$$

Hereafter, we implicitly assume a correspondence between the q-mappings of data and the transcriptional module genes as each direction $\mathbf{p}_k$ plays a role to correct up the gene expression patterns in a transcriptional module. In this sense, we call the clustering system based on a linear mapping the transcriptional module-based clustering.

In clustering context, the q-directions $\{\mathbf{p}_k\}_{k=1}^q$ should be chosen such that the transformed data $\{\mathbf{P}^T\mathbf{x}_j\}_{j=1}^N$ successfully reveal the clusters existed in $\mathbb{R}^d$. Then we can identify the clusters based on the lower-dimensional dataset. These two tasks are formulated as the statistical estimation for $\{\mathbf{c}, \mathbf{P}\}$. This problem amounts to an optimization problem that minimizes a loss function $Q(\mathbf{c}, \mathbf{P})$ with respect to the unknown encoders function $\mathbf{c}(\mathbf{x})$ and the q-directions $\{\mathbf{p}_k\}_{k=1}^q$. One of the key results in this study is to establish a link these two processes, i.e. the dimension reduction of data and the clustering algorithm, and a probabilistic framework. In this context, the optimization for $\min_{c,P} Q(\mathbf{c}, \mathbf{P})$ is converted into a penalized likelihood estimation of a probability model of which we call the mixed factors model. Such formulation gives us a great deal of utilities, in either the computation for finding $\min_{c,P} Q(\mathbf{c}, \mathbf{P})$ and the determination of the number of clusters and the appropriate dimension of projected data space, $\{G, q\}$, respectively.

The rest of this article is organized as follows. In section 2 we introduce two criteria to be minimized in the construction of linear mapping. In section 3, we will define a generalized loss function that links the two criteria introduced in section

2. Section 4 presents a probabilistic formulation of this approach. Section 5 present an optimization algorithm for minimizing the proposed generalized loss function. Section 6 contains the determination of the number of clusters and some another parameters. In section 7, the effectiveness of our method will be demonstrated thorough the application to a well-known microarray data, the small round cell tumors of childhood. Finally, the concluding remarks are give in Section 8.

2 Clustering Based on Linear Mapping

2.1 Principal Component Analysis

Principal component analysis (PCA, [1]) is one of the most commonly used techniques for constructing a linear mapping of data in statistical data analysis including bioinformatics ([4],[5]). PCA determines the q-directions $\{\mathbf{p}_k\}_{k=1}^q$ to minimize the negative variance of $\{\mathbf{P}^T\mathbf{x}_j\}_{j=1}^N$ with taking account $||\mathbf{p}_k||^2 = 1$, $k \in \{1, \cdots, q\}$. Thus, the objective function to be minimized is

$$Q_A(\mathbf{P}) := -\frac{1}{N}\sum_{k=1}^{q}\sum_{j=1}^{N}\mathbf{p}_k^T\mathbf{x}_j\mathbf{x}_j^T\mathbf{p}_k + \sum_{k=1}^{q}\lambda_k(||\mathbf{p}_k||^2 - 1).$$

where the $\{\lambda_k\}_{k=1}^q$ denote the Lagrange multipliers to impose $||\mathbf{p}_k||^2 = 1$, $k \in \{1, \cdots, q\}$. Here, we assume that the origin of $\{\mathbf{x}_j\}_{j=1}^N$ has been shifted to zero by subtracting the sample mean from all samples.

The optimal q-directions $\{\hat{\mathbf{p}}\}_{k=1}^q$ are equal to the q-principal axes of the sample covariance matrix corresponding to the dominant eigenvalues $\{\hat{\lambda_k}\}_{k=1}^q$. However, as was remarked by some literatures, PCA sometimes fails to reveal the presence of clusters shown by the original data [2, 7]. For instance, when the within-cluster variance on a particular cluster largely dominates the between-clusters variance, a direction tends to the principal axis corresponding to one clusters [2, 7]. Most such limitation are related to the fact that PCA only takes into consideration the second order characteristic of data.

2.2 Within-Cluster Variances

Alternatively, consider to seek a linear mapping to minimize the overlaps of clusters revealed onto $\mathbb{R}^q$. Let us define a loss function to be the Euclid distance between $\mathbf{P}^T\mathbf{x}$ and the unknown centroids $\{\mu_g\}_{g=1}^G$ of G-clusters:

$$L(\mathbf{P}, \mu; \mathbf{x}) := \sum_{g=1}^{G} c_g(\mathbf{x})||\mathbf{P}^T\mathbf{x} - \mu_g||^2, \quad \mathbf{x} \in \mathbb{R}^d. \tag{1}$$

Hereafter we stand for the true distribution of data by $f(\mathbf{x})$. Besides, we also denote the conditional distribution of $\mathbf{c}(\mathbf{x})$ by

$$f(\mathbf{c}(\mathbf{x})|\mathbf{x}) := \prod_{g=1}^{G} w_g(\mathbf{x})^{c_g(\mathbf{x})},$$

where the unknown functionals $\mathbf{w}(\mathbf{x}) = \{w_g(\mathbf{x})\}_g$ satisfy

$$\{w_g(\mathbf{x}) \geq 0\}_{g=1}^{G}, \quad \sum_{g=1}^{G} w_g(\mathbf{x}) = 1, \quad \mathbf{x} \in \mathbb{R}^d.$$

Taking the expectation of (1) with respect to $f(\mathbf{x}, \mathbf{c}) = f(\mathbf{x})f(\mathbf{c}(\mathbf{x})|\mathbf{x})$ defines a risk function to be minimized in the construction of estimators for $\{\mathbf{c}(\mathbf{x}), \mathbf{P}\}$ although the true distribution of $\{\mathbf{x}, \mathbf{c}(\mathbf{x})\}$ is unknown. Instead, replacing $f(\mathbf{x})$ by the empirical distribution $\hat{f}(\mathbf{x})$, we can obtaine an empirical loss function

$$\begin{aligned} Q_B(\mathbf{w}, \mathbf{P}, \mu) :&= E_f L(\mathbf{P}, \mu; \mathbf{x}) \\ &= \frac{1}{N} \sum_{g=1}^{G} \sum_{j=1}^{N} w_g(\mathbf{x}_j) ||\mathbf{P}^T \mathbf{x}_j - \mu_g||^2 - \sum_{k=1}^{q} \lambda_k (||\mathbf{p}_k||^2 - 1). \quad (2) \end{aligned}$$

Here the $\{\lambda_k\}_{k=1}^{q}$ denote the Lagrange multiplier. The first term in (2) presents just the within-cluster variances of $\{\mathbf{p}_k^T \mathbf{x}_j\}_{j=1}^{N}$, $k \in \{1, \cdots, q\}$. An optimal $\hat{\mathbf{P}}$ minimizes the overlap of G-clusters revealed onto the $\mathbb{R}^q$ although the conditional distributions $\{\mathbf{w}(\mathbf{x}_j)\}_{j=1}^{N}$ and the G-centroids $\{\mu_g\}_{g=1}^{G}$ remain to be unknown. The optimization method will be described in later under more general setting.

3 Generalized Criterion

While the minimum within-cluster variances are a suitable criterion in the construction of linear mapping to reflect the group structure of original dataset, its applicability might be limited due to the dimensionality of the data. Most limitations are related to the occurence of overlearning. Such unsuitableness occurs due to the fact that the N data points are sparsely distributed on $\mathbb{R}^d$. Then, the degree of freedom in the determination of $\{\mathbf{p}_k\}_{k=1}^{q}$ is extreamly large. Accordingly, the compressed samples $\{\mathbf{P}^T \mathbf{x}_j\}_{j=1}^{N}$ might improperly exhibit the clusters despite no clusters on $\mathbb{R}^d$.

To overcome such limitation, we propose a criterion for estimating parameters by combining the score functions $Q_A(\mathbf{P})$ and $Q_B(\mathbf{w}, \mathbf{P}, \mu)$ of the form

$$\begin{aligned} Q_\alpha(\mathbf{w}, \mathbf{P}, \mu) = &-\frac{1}{N} \sum_{j=1}^{N} ||\mathbf{P}^T \mathbf{x}_j||^2 + \frac{\alpha}{N} \sum_{g=1}^{G} \sum_{j=1}^{N} w_g(\mathbf{x}_j) ||\mathbf{P}^T \mathbf{x}_j - \mu_g||^2 \\ &+ \sum_{k=1}^{q} \lambda_k (||\mathbf{p}_k||^2 - 1), \quad (3) \end{aligned}$$

where $\alpha \in [0, 1]$ is a mixing rate that controls the trade-off between the total variance and the between-clusters variance. Here the $\{\lambda_k\}_{k=1}^{q}$ denote the Lagrange multipliers for taking account $\{||\mathbf{p}_k||^2 = 1\}_{k=1}^{q}$. Notice that for any $\{\mathbf{w}(\mathbf{x}_j), \mathbf{P}\}$,

the minimization of (3) with respect to the G-centroids is accomplished by the weighted average of $\{\mathbf{P}^T\mathbf{x}_j\}_{j=1}^N$:

$$\hat{\mu}_g = \frac{1}{N\bar{w}_g}\sum_{j=1}^{N} w_g(\mathbf{x}_j)\mathbf{P}^T\mathbf{x}_j, \tag{4}$$

where $\bar{w}_g = (1/N)\sum_{j=1}^N w_g(\mathbf{x}_j)$. Equating $\mu_g = \hat{\mu}_g$ for $g \in \{1, \cdots, G\}$, the first two terms in (3) can be rewritten as

$$-\frac{1}{N}\sum_{j=1}^{N}||\mathbf{P}^T\mathbf{x}_j||^2 + \frac{\alpha}{N}\sum_{g=1}^{G}\sum_{j=1}^{N} w_g(\mathbf{x}_j)||\mathbf{P}^T(\mathbf{x}_j - \bar{\mathbf{x}}_g)||^2. \tag{5}$$

where the $\{\bar{\mathbf{x}}_g\}_{g=1}^G$ denote the group means corresponding to $\{\mathbf{x}_j\}$,

$$\bar{\mathbf{x}}_g = \frac{1}{N\bar{w}_g}\sum_{j=1}^{N} w_g(\mathbf{x}_j)\mathbf{x}_j, \quad g \in \{1, \cdots, G\}. \tag{6}$$

As $\alpha \to 0$, the quantity (5) tends to the variance of $\{\mathbf{P}^T\mathbf{x}_j\}_{j=1}^N$, and then, the optimal $\{\hat{\mathbf{p}}_k\}_{k=1}^q$ tends to the principal axes. To the contrary, as $\alpha \to 1$, the (5) tends to the negative between-clusters variance of $\{\mathbf{P}^T\mathbf{x}_j\}_{j=1}^N$:

$$-\text{ trace}\left(\mathbf{P}^T\sum_{g=1}^{G}\bar{\mathbf{x}}_g\bar{\mathbf{x}}_g^T\mathbf{P}\right) = -\text{trace }\sum_{g=1}^{G}\hat{\mu}_g\hat{\mu}_g^T.$$

Then, a linear mapping of data with the optimal q-directions tends to separate the G-centroids.

Next, consider a computational aspect in the construction of the optimal q-directions. Differentiating (3) with respect to $\mathbf{p}_k$ with equating $\mu_g = \hat{\mu}_g$ leads to an equation to be solved,

$$\left[\frac{1}{N}\sum_{j=1}^{N}\mathbf{x}_j\mathbf{x}_j^T - \frac{\alpha}{N}\sum_{g=1}^{G}\sum_{j=1}^{N} w_g(\mathbf{x}_j)(\mathbf{x}_j - \bar{\mathbf{x}}_g)(\mathbf{x}_j - \bar{\mathbf{x}}_g)^T - \lambda_k\mathbf{I}\right]\mathbf{p}_k = \mathbf{0}. \tag{7}$$

Obviously, the solutions can be given by the corresponding eigenvalues which satisfy

$$\hat{\lambda}_k = \mathbf{p}_k^T\left[\frac{1}{N}\sum_{j=1}^{N}\mathbf{x}_j\mathbf{x}_j^T - \frac{\alpha}{N}\sum_{g=1}^{G}\sum_{j=1}^{N} w_g(\mathbf{x}_j)(\mathbf{x}_j - \bar{\mathbf{x}}_g)(\mathbf{x}_j - \bar{\mathbf{x}}_g)^T\right]\mathbf{p}_k. \tag{8}$$

Thus, the $\min Q_\alpha(\mathbf{w}, \mathbf{P}, \mu)$ with any fixed third arguments $\{\mathbf{w}\}$ can be attained at $Q^* = -\sum_{k=1}^q \hat{\lambda}_k$ with a series of the dominant eigenvalues, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq, \cdots, \geq \hat{\lambda}_q$: the optimal $\hat{\mathbf{p}}_1$ corresponds to the largest $\hat{\lambda}_1$, and the rest of directions $\{\hat{\mathbf{p}}_k\}_{k=2}^q$ are orthogonal to all of the preceding ones.

Given $\{\mathbf{P}, \mu\}$, the estimation of $\{\mathbf{w}(\mathbf{x}_j)\}_{j=1}^N$ can be accomplished by the K-means-like rule: the solution will put unit value on $w_g(\mathbf{x}_j)$ with a smallest distance between $\hat{\mathbf{P}}^T\mathbf{x}_j$ and $\{\hat{\mu}_g\}_{g=1}^G$. In this article, we generalized this type of clustering, i.e. the hard clustering, to the soft clustering. This goal can be accomplished by imposing a smoothness on the $\{w_g(\mathbf{x})\}_g$. One such penalization is the negative entropy of $\{w_g(\mathbf{x})\}_g$:

$$H(\mathbf{w}(\mathbf{x})) := \sum_{g=1}^{G} w_g(\mathbf{x}) \log w_g(\mathbf{x}).$$

This function achieves the minimum value for equal values $w_g(\mathbf{x}) = 1/G$, $g \in \{1, \cdots, G\}$, and is correspondingly larger as the $\{w_g(\mathbf{x})\}_{g=1}^G$ tends to more unequal. Consequently, the modified criterion becomes

$$Q_\alpha(\mathbf{w}, \mathbf{P}, \mu) - \beta \sum_{j=1}^{N} H(\mathbf{w}(\mathbf{x}_j)). \tag{9}$$

The quantity $\beta \geq 0$ controls the strength of penalty that tunes the trade-off between soft and hard clustering. The optimal $\{\hat{w}_g(\mathbf{x}_j)\}_{g=1}^G$ for this objective function turns to

$$\hat{w}_g(\mathbf{x}_j) \propto \exp\Big(-\frac{\alpha}{\beta} ||\mathbf{P}^T\mathbf{x}_j - \mu_g||^2 \Big). \tag{10}$$

for all $j \in \{1, \cdots, N\}$. This solution puts the increased weight on a particular group to be the smallest $||\mathbf{P}^T\mathbf{x}_j - \mu_g||^2$ as β tends to small, and setting $\beta \to \infty$ places the equal weights on all groups. Correspondingly, our approach alternates between two steps, solving (7) and the grouping (10) with an initial starting value until a series of the corresponding Q_α is in convergence.

We will revist the computational aspect of this method in section 5. The proposed optimization algorithm can be implemented without solving the eigenvalues equation that might be computationaly very demanding in microarray study. The remined tasks are the determination of smoothness $\{\alpha, \beta\}$ and a suitable q on which data are mapped. Moreover, the number of clusters G must often be deduced from data. As will be shown in section 6, these tasks can be converted into the statistical model selection throught the probabilistic formulation of the method.

4 Probabilistic Formulation

We now discuss the proposed clustering method within a probabilistic framework. Let $\mathbf{f}_j \in \mathbb{R}^q$ be a latent random variable corresponding to the jth sample where q is much smaller than d. Then, suppose that a set $\{\mathbf{x}_j, \mathbf{f}_j\}_{j=1}^N$ is independently distributed according to

$$\mathbf{x}_j = \mathbf{P}\mathbf{f}_j + \epsilon_j, \tag{11}$$

$$\mathbf{f}_j|c_g(\mathbf{x}_j) = 1 \sim N(\mu_g, \sigma\mathbf{I}), \quad g \in \{1, \cdots, G\}, \tag{12}$$

where the ϵ_j is assumed to be Gaussian noise with $N(0, \gamma\mathbf{I})$ and to be independent to $\mathbf{f}_j$. The observational equation (11) states that for a given $\mathbf{f}_j$, the $\mathbf{x}_j$ is distributed to be $N(\mathbf{P}\mathbf{f}_j, \gamma\mathbf{I})$. Accordingly, this generative model also states the distribution of data conditional on the class label by

$$\mathbf{x}_j|c_g(\mathbf{x}_j) = 1 \sim N(\mathbf{P}\mu_g, \sigma\mathbf{P}\mathbf{P}^T + \gamma\mathbf{I}), \quad g \in \{1, \cdots, G\}.$$

Thus, the distributional aspect of data is characterized by G-clusters centered at $\{\mathbf{P}\mu_g\}_{g=1}^G$. This model, called the mixed factors model, was originally proposed by Yoshida et al. [8] to intend a parsimonious parameterization of the Gaussian mixture.

As the preceding method imposes the orthogonality on the q-directions, we now assume the orthogonality of q-columns in the loading matrix $\mathbf{P} = \{\mathbf{p}_k\}_k$. Then the logged-density of $\mathbf{x}$ can be written as

$$\log P(\mathbf{x}|\mathbf{c}(\mathbf{x})) = const. - \frac{1}{\gamma}(||\mathbf{x}||^2 - ||\mathbf{P}^T\mathbf{x}||^2) - \frac{1}{\gamma+\sigma}\sum_{g=1}^{G} c_g(\mathbf{x})||\mathbf{P}^T\mathbf{x} - \mu_g||^2 \tag{13}$$

Taking the expectation of (13) with respect to the empirical distribution $\hat{f}(\mathbf{x})$ and the conditional distribution of unknown class labels $f(\mathbf{c}(\mathbf{x})|\mathbf{x})$ leads to the log-likelihood function of unknown parameters $\{\mathbf{w}, \mathbf{P}, \mu\}$ after multiplying (13) by γ:

$$L(\mathbf{w}, \mathbf{P}, \mu) = const. + \frac{1}{N}\sum_{j=1}^{N}||\mathbf{P}^T\mathbf{x}_j||^2 - \frac{\alpha}{N}\sum_{g=1}^{G}\sum_{j=1}^{N} w_g(\mathbf{x}_j)||\mathbf{P}^T\mathbf{x}_j - \mu_g)||^2,$$

where $\alpha = \gamma/(\gamma+\sigma)$. Adding both the regularization term $\beta\sum_{j=1}^{N} H(\mathbf{x}_j)$ and the Lagrange terms $-\sum_{k=1}^{q}\lambda_k(||\mathbf{p}_k|| - 1)$ to this function gives a criterion equivalent to the negative of (9). This implies that the problem to be solved in our method turns to a penalized likelihood estimation of the mixed factors model.

5 Maximization-Maximization Algorithm

Here, we present an optimization algorithm to maximize the penalized likelihood of the mixed factors model, that is equivalent to find the minimizer of (9). This can be achieved by the EM algorithm (Dempster et al.[3]). The EM algorithm takes $\{\mathbf{x}_j, \mathbf{f}_j\}_{j=1}^N$ as a complete data set and then alternates between the two steps: the expectation of the complete data likelihood with respect to the posterior distribution of the unknown factors $\{\mathbf{f}_j\}_{j=1}^N$,

$$P(\mathbf{f}_j|c_g(\mathbf{x}_j), \mathbf{x}_j) = \phi(\mathbf{f}_j; \alpha\mu_g + (1-\alpha)\mathbf{P}^T\mathbf{x}_j, \lambda\alpha\mathbf{I}),$$

and the maximization of the expected complete data likelihood. Hereafter, we let $\phi(\ ; \mathbf{a}, \mathbf{B})$ be the Gaussian density with the mean $\mathbf{a}$ and the covariance matrix $\mathbf{B}$.

Consider now to update the hth direction $\mathbf{p}_h$ whereas the another parameters are fixed at the current values $\{\hat{\mathbf{p}}_k\}_{k\neq h}$, $\{\hat{\mu}_g\}_g$ and $\{\hat{\mathbf{w}}(\mathbf{x}_j)\}_j$. By the definition, the complete data log-likelihood of the mixed factors model, $L_c = (1/N)\sum_j \log P(\mathbf{x}_j, \mathbf{f}_j|\mathbf{c}(\mathbf{x}_j))$, can be explicitly represented by

$$L_c(\mathbf{w}, \mathbf{P}, \mu) := \frac{1}{N}\sum_{j=1}^{N} \phi(\mathbf{x}_j; f_{hj}\mathbf{p}_h + \sum_{h\neq k} f_{kj}\hat{\mathbf{p}}_k, \gamma\mathbf{I}) + \frac{1}{N}\sum_{g=1}^{G}\sum_{j=1}^{N} \hat{w}_g(\mathbf{x}_j)\phi(\mathbf{f}_j; \hat{\mu}_g, \sigma\mathbf{I}), \quad (14)$$

where f_{hj} is the hth element of $\mathbf{f}_j$. Note that the hth direction depends only the first term in (14) which corresponds to the observational equation (11).

Let $\langle L_c \rangle$ be the conditional expectation of (14) where the expectation is taken with respect to the $P(\mathbf{f}_j|c_g(\mathbf{x}_j), \mathbf{x}_j)$ evaluated with the current parameters $\{\hat{\mathbf{w}}, \hat{\mathbf{P}}, \hat{\mu}\}$. Then the objective function to be maximized at this step is

$$\langle L_c \rangle + \eta_h(||\mathbf{p}_h||^2 - 1) + \sum_{k\neq h} \eta_k \hat{\mathbf{p}}_k^T \mathbf{p}_h.$$

Here, the $\{\eta_k\}_{k=1}^q$ are the Lagrange multipliers to impose the orthogonality on the q-directions. Solving this gives the optimal $\hat{\mathbf{p}}_h$ as

$$\hat{\mathbf{p}}_h = \frac{1}{S}\Big[\sum_{j=1}^{N}\langle f_{hj}\rangle \mathbf{x}_j - \sum_{k\neq h} \hat{\mathbf{p}}_k^T \sum_{j=1}^{N}\langle f_{hj}\rangle \mathbf{x}_j \hat{\mathbf{p}}_k\Big],$$

where the S denotes the normalizing constant to satisfy $||\hat{\mathbf{p}}_h||^2 = 1$, and the conditional expectation of the latent variables, $\langle f_{hj}\rangle$, is equal to the hth element of

$$\langle \mathbf{f}_j \rangle = \alpha \sum_{g=1}^{G} \mu_g + (1-\alpha)\mathbf{P}^T \mathbf{x}_j.$$

Repeating this process for $h \in \{1, \cdots, q\}$, we would have a series of q-directions $\{\hat{\mathbf{p}}_k\}_{k=1}^q$.

Given an estimate of q-directions, the G-centroids of clusters and the conditional distribution of class labels, $\{w_g(\mathbf{x}_j)\}_g$, for $j \in \{1, \cdots, N\}$ are estimated by (6) and (10), respectively. Thus, we just compute the simple recursive formulas until the sequence of estimates and the corresponding penalized likelihood are judged to be converged. Such sequence of parameters yields a non-decreasing sequence of the penalized likelihood of the mixed factors model. To sum up, we summarize this algorithm in below:

1. Set the initial values for $\{\hat{\mathbf{w}}, \hat{\mathbf{P}}, \hat{\mu}\}$ and $\{G, q, \alpha, \beta\}$. Then repeat the step 2 to 4 until the sequence of either parameters and the corresponding penalized likelihood will be converged:

2. (q-directions)
   for $h = 1$ to q, update $\hat{\mathbf{p}}_h$ by

$$\hat{\mathbf{p}}_h = \frac{1}{S}\Big[\sum_{j=1}^{N}\langle f_{hj}\rangle \mathbf{x}_j - \sum_{k\neq h}\hat{\mathbf{p}}_k{}^T \sum_{j=1}^{N}\langle f_{hj}\rangle \mathbf{x}_j\hat{\mathbf{p}}_k\Big].$$

3. (G-centroids) for $g = 1$ to G, update $\hat{\mu}_g$ by

$$\hat{\mu}_g = \frac{1}{N\bar{w}_g}\sum_{j=1}^{N}\hat{w}_g(\mathbf{x}_j)\hat{\mathbf{P}}^T\mathbf{x}_j.$$

4. (Grouping function) for $g = 1$ to G and $j = 1$ to M, update $\hat{w}_g(\mathbf{x}_j)$ by

$$\hat{w}_g(\mathbf{x}_j) \propto \exp\Big(-\frac{\alpha}{\beta}||\hat{\mathbf{P}}^T\mathbf{x}_j - \hat{\mu}_g||^2\Big).$$

Notice that we have no need to evaluate the noise variances $\{\gamma, \sigma\}$ of the mixed factors model during this procudure. However, the model selection method described in next section requires the evaluation of these parameters. It follows from $\alpha = \gamma/(\gamma + \sigma)$ that the σ can be estimated by $\hat{\sigma} = (1 - \alpha)\hat{\gamma}/\alpha$ with a given estimate $\hat{\gamma}$. By the simple calculation, it can be seen that an optimal $\hat{\gamma}$ necessarily satisfies

$$\hat{\gamma} = \frac{1}{(d-q)N}\sum_{j=1}^{N}\Big(||\mathbf{x}_j||^2 - ||\hat{\mathbf{P}}^T\mathbf{x}_j||^2\Big).$$

6 Penalized Mixed Factors Analysis

A basic issue arising in this method is the determination of the number of clusters G, the suitable dimension of the linear mapping q and the strength of penalties $\{\alpha, \beta\}$. Within statistical framework, this issue can be converted into the model selection problem that chooses a suitable set $\{G^*, q^*, \alpha^*, \beta^*\}$ among the possible combinations. In this article, we address this problem by selecting a particular combination to show the best predictability.

Consider to split $\{\mathbf{x}_j\}_{j=1}^N$ into the two disjoint subsets, a training sample set $\{\mathbf{x}_j^e\}_{j=1}^{N_e}$ used in the estimation of parameters and a set of the blinded test sample $\{\mathbf{x}_j^b\}_{j=1}^{N_b}$. Let $\{\hat{\mathbf{w}}^e, \hat{\mathbf{P}}^e, \hat{\mu}^e\}$ be a set of parameters estimated by the training samples with a particular $\{G, q, \alpha, \beta\}$. One possible approach is to select a combination $\{G^*, q^*, \alpha^*, \beta^*\}$ to minimize the prediction error

$$C(\{G, q, \alpha, \beta\}) := -\frac{1}{N_b}\sum_{j=1}^{N_b}\log P(\mathbf{x}_j^b; \{\hat{\mathbf{w}}^e, \hat{\mathbf{P}}^e, \hat{\mu}^e\}),$$

where $P(\mathbf{x}; \{\mathbf{w}, \mathbf{P}, \mu\})$ is the unconditional density of data to be the Gaussian mixture as

$$P(\mathbf{x}; \{\hat{\mathbf{w}}, \hat{\mathbf{P}}, \hat{\mu}\}) = \sum_{g=1}^{G} \frac{1}{G} \phi(\mathbf{x}; \hat{\mathbf{P}}\hat{\mu}_g, \hat{\sigma}\hat{\mathbf{P}}\hat{\mathbf{P}}^T + \hat{\gamma}\mathbf{I}). \tag{15}$$

Given a set $\{G^*, q^*, \alpha^*, \beta^*\}$, our method calibrates G-clusters based on the estimated conditional distribution of class labels $\{w_g(\mathbf{x}_j)\}_g$. The most common classifier is to assign $\mathbf{x}_j$ to a cluster with the highest posterior probability of belonging:

$$\hat{c}_g(\hat{\mathbf{P}}^T \mathbf{x}_j) = \begin{cases} 1 \text{ if } w_g(\mathbf{x}_j) = \max\limits_{h \in \{1,\cdots,G\}} w_h(\mathbf{x}_j), \\ 0 \qquad\qquad \text{otherwise.} \end{cases}$$

Biological interpretation of q-coordinates corresponding to $\{\hat{\mathbf{P}}^T \mathbf{x}_j\}_{j=1}^N$ is important for real data analysis. This can be achieved by investigating the values in q-directions $\{\hat{\mathbf{p}}_k\}_k$. Obviously, a particular element of $\{\mathbf{x}_j\}_{j=1}^N$ had a large contribution in the calibration of clusters if the corresponding element of $|\hat{\mathbf{p}}_k|$ takes a large value. To the contrary, if an element of $|\hat{\mathbf{p}}_k|$ takes a value close to zero, the kth coordinate is not affected by the corresponding gene. In this way, by investigating all values in $\hat{\mathbf{P}}$, each of q-directions can be understood. In practice, it will be helpful to list the top L of genes to give the highest positive values in $\hat{\mathbf{p}}_k$ at Ω_+^k and to give the highest negative values in $\hat{\mathbf{p}}_k$ at Ω_-^k for each $k \in \{1, \cdots, q\}$. As will be demonstrated in next section, for the gene expression analysis, these $2q$ sets can be useful either to find the biologically meaningful groups of genes and to elucidate a causal link from the calibrated clusters to the biological knowledge.

7 Real Application

Khan et al. [6] classified the small round blue cell tumors (SRBCT's) of childhood into the four diagnostic categories, neuroblastoma (NB), rhabdomyosarcoma (RMS), non-Hodgkin's lymphoma (NHL) and the Ewing family of tumors (EWS) using cDNA gene expression profiles. The dataset is available at the website `http://www.nhgri.nih.gov/DIR/Microarray/Supplement/`. For each of the 83 SRBCT samples, the expression levels of $2,308$ genes were measured. Khan et al. [6] splited the data into two parts; the training set comprising 63 cases (NB, 12; RMS, 20; BL, 8; EWS, 23) and the test set, 20 cases (NB, 6; RMS, 5; BL, 3; EWS, 6) where Burkitt's lymphoma (BL) is a subset of NHL. All samples are summarized in Figure 1. Note that the name of sample specifies the cancer type suffixed with -T for a tumor biopsy material and -C for a cell line. Khan et al. [6] successfully classified the tumor types into the four categories using artificial neural networks. Unlike this, the purpose of our study is to identify the clusters of these SRBCT's in the unsupervaised manner, and then, to look at the association between the caribrated clusters and some medical outcome, that is, the unsupervised learning.

For the preprocessing, we removed genes whose range of expression values across 83 samples is less than 3.0, 680 genes then remain to be analysed. We then adjusted the columns of 680×88 data matrix to have mean zero after centering the rows. To find an optimal $\{\hat{\mathbf{w}}, \hat{\mathbf{P}}, \hat{\mu}\}$, we used the 63 training samples including tumors and cell lines, 13 EWS-T, 10 EWS-C, 12 NB, 8 BL, 10 RWS-T and 10 RWS-C. The 20 blinded samples were used to select $\{G^*, q^*, \alpha^*, \beta^*\}$ and also to assess the predictability of the resulting clusters.

We candidated a set of the number of clusters ranging from $G = 4, 5, 6$ and the dimension onto which the data are mapped varying $q = 1$ to 11. We also candidated a set of possible combinations of the regularization parameters as $\{\alpha, \beta\} \in \{0.1, 0, 2, \cdots, 0.9\} \times \{0.8, 0.9, 1.0, 1.1, 1.2\}$. The smallest local minimum of the generalized criterion corresponding to $q = 8$ gave the minimum scores of $C(\{G, q, \alpha, \beta\})$ for all G in which the most suitable smoothing parameters $\{\alpha^*, \beta^*\}$ were given. Figure 1 shows the groupings given by $G = 4, 5, 6$ fixed at $q = 8$. The calibrated model with $G = 4$ correctly grouped all samples into the diagnostic categories, i.e. EWS, NB, BL and RWS. It also could be seen from $G = 5$ in Figure 1 that the RWS samples were divided into the two subgroups as corresponding to the heterogeneity between RWS-Ts and RWS-Cs. Moreover, the clustering given by $G = 6$ yielded a partition as reflecting the molecular dissimilarity between the tumor samples and cell lines on the EWSs. Indeed, the model selection based on $C(\{G, q, \alpha, \beta\})$ showed the evidence of molecular subtypes on either EMS and RWS as the model of $G = 6$ was judged to be optimal. We also tested the capability of the caribrated clustering rule using the 20 blinded samples (see Figure 1). When these samples were assigned into a particular cluster using the resulting classifiers for each G, we obtained the plausible grouping as likely to reflect the diagnostic categories of cancer types, for all G. For instance, of the 20 blinded samples, TEST20-EWS-T were misclassified into the RMS related category for $G = 4$. In addition, for $G = 6$, 19 of the 20 test samples were correctly grouped into the related diagnostic categories in which TEST19-EWS-T was misclassified into the EWS-C related cluster. From this analysis, the predictability confirm us the effectiveness of the estimated grouping.

A causal link from the clusters to the biological knowledge can be elucidated thorough the inspection of relevant genes. Figure 1 illustrates the expression patterns of 16 set of relevant genes selected by $G = 6$ and $q = 8$. For instance, the genes in Ω_+^1 are good discriminators on the basis of the lack of expression in BL and RMS, and the high expression in EWS and NB. Note also that the genes in Ω_-^1 showed the opposite expression patterns to Ω_+^1. The relevant genes in the some sets were expressed in one or two of the six molecular categories as the Ω_+^5, Ω_-^7 and Ω_+^8 are specific to NB, EWS and RMS, respectively. Of interest is that the genes in Ω_+^2 are specifically expressed in the tumor samples as EWS-T, RWS-T and and not expressed in the cell lines. The genes in Ω_-^2 shows the opposite patterns to Ω_+^2 as the lack of expression in the tumor samples and the high-expression in the cell lines. This fact validates the presence of heterogeneity corresponding to the molecular types within a cancer type.

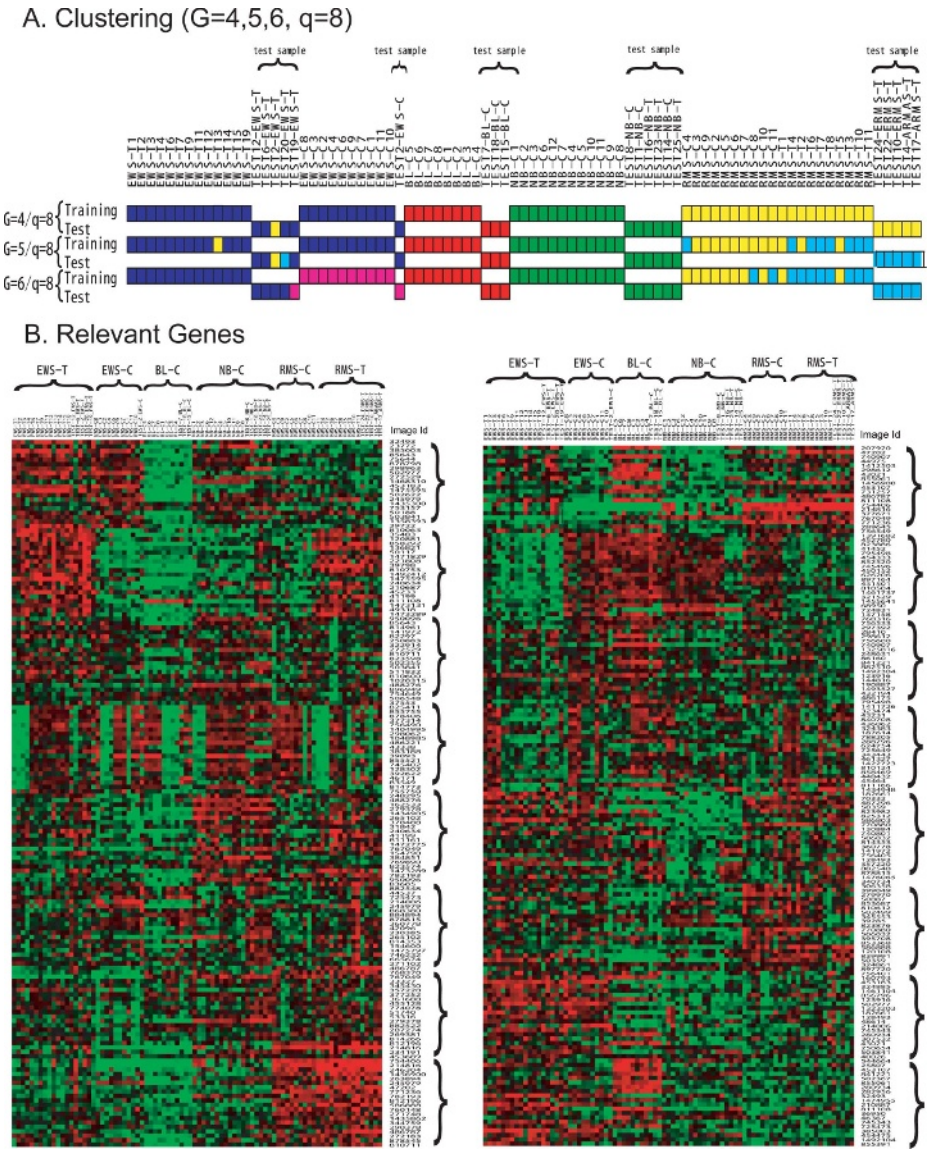


Fig. 1. Caribrated Clusters and relevant genes. **A**. Clustering result. The 63 SRBCTs samples (training samples) were used for finding $\{\hat{\mathbf{w}}, \hat{\mathbf{P}}, \hat{\mu}\}$, and then, the training set and the 20 test samples were grouped into clusters base on the calibrated classifiers for each combination $\{G, q, \alpha, \beta\}$. Shown here are clusters caribrated by the smallest local minima of the generalized criterion corresponding to $q = 8$ for $G = 4, 5, 6$ where the smoothness paramters were induced from the test samples. The resulting groups are depicted by the colors. **B.** Relevant genes selected by the optimal model, $G = 6$, $q = 8$. Shown in the left panel are the expression patterns of the 8 sets of 20 genes listed at Ω_+^k, $k \in \{1, \cdots, 8\}$. The expression patterns of genes listed at Ω_-^k, $k \in \{1, \cdots, 8\}$, are also shown in the right panel

8 Concluding Remarks

In this study, we proposed a method of clustering for the high-dimensional microarray dataset. A distinction of our method is that the clustering rule is applied to the linear mapping of data onto the low-dimensional subspace, rather than the original dataset. In the construction of linear mapping, the directions are chosen as to minimize a criterion that links the variance and the within-cluster variances of the compressed data. We also established an optimization algorithm to find such directions and a suitable clustering rule.

The effectiveness of the proposed method was demonstrated through the application to a well-known gene expression data, the small round blue cell tumors of childhood (SRBCTs). The clustering system could find the biologically meaningful groups of SRBCTs as we confirmed a plausible correspondence between the calibrated clusters and the diagnostic categories. Besides, the method identified sets of relevant genes associated with the calibrated clusters. These sets might be helpful to elucidate the causal link between the obtained grouping and the existing knowledge on biology.

References

1. Anderson, T.W.: An Introduction to multivariate statistical analysis. Wiley, New York, (1984)
2. Chang, W.C.: On using principal components before separating a mixture of two multivariate normal distributions. Applied Statistics **32** (1983) 267–275
3. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm (with discussion), J. Royal Stat. Soc. B. **39** (1977) 1–38
4. Ghosh, D., Chinnaiyan, A.M.: Mixture modeling of gene expression data from microarray experiments. Bioinformatics **18(2)** (2002) 275–286
5. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gassenbeck, M., Mersirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S.: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science **286** (1999) 531–537
6. Khan, J., Wei, J.S., Ringner, M., Saal, L.H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Atonescu, C.R., Peterson, C., Meltzer, P.S.: Classification and Diagnostic Prediction of Cancers using Gene Expression Profiling and Artificial Neural Networks Nature Medicine **7** (2001) 673–679
7. McLachlan, G.J., Peel, D.: Finite mixture models. Wiley New York (1997)
8. Yoshida, R., Higuchi, T., Imoto, S.: A mixed factors model for dimension reduction and extraction of a group structure in gene expression data. Proc. 3rd Computational Systems Bioinformatics (2004) 161-172

Conceptual Modeling of Genetic Studies and Pharmacogenetics

Xiaohua Zhou and Il-Yeol Song

College of Information Science and Technology, Drexel University
3141 Chestnut Street, Philadelphia, PA 19104, USA
{xiaohua.zhou, song}@drexel.edu

Abstract. Genetic Studies examine relationships between genetic variation and disease development. Pharmacogenetics studies the responses to drugs against genetic variation. These two lines of research evaluate relationships among genotype, phenotype, and environment regarding subjects. These studies demand a variety of other information; such as clinical observations, disease development history, demographics, life style, and living environment. Correct and informative modeling of these data is critical for bioinformaticians; the model affects the capacity of data manipulation and the types of queries they can ask as well as performance of the implemented system. In this paper, we present a conceptual model on genetic studies and Pharmacogenetics using Unified Modeling Language (UML). Our model provides a comprehensive view of integrated data for genetic studies and Pharmacogenetics by incorporating genomics, experimental data, domain knowledge, research approaches, and interface data for other publicly available resources into one cohesive model. Our model can support diverse biomedical research activities that use both clinical and biomedical data to improve patient care through incorporation of the roles of environment, life style and genetics. Our model consists of a set of class diagrams organized into a hierarchy of packages diagrams to clearly and intuitively show inter-object relationships at different levels of complexity.

1 Introduction

Completion of a high quality comprehensive sequence of the human genome marked the arrival of the Genomic Era. Genomics heavily impacts research in the fields of biology, health, and society. Clinical opportunities for gene-based pre-symptomatic prediction of illness and adverse drug responses are emerging at a rapid pace [2].

Genetic studies examine relationships between genetic variation and disease susceptibility. Pharmacogenetics is the study of response to drugs against genetic variation. These two lines of research address the triad of relationships among genotype, phenotype, and environment. Apart from genomic information, a variety of other information is required such as clinical observations, disease development history, demographics, life style, and living environment. Comprehensive and correct modeling of these data is necessary for bioinformaticians; the model affects the

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 402 – 415, 2005.

capacity of data manipulation and the types of queries they can ask as well as performance of the implemented system.

There have been a lot of cutting-edge works on conceptual modeling of biological data. However, most focus on a single type of data such as genome sequences, protein structures, protein interactions, and metabolic pathways (see Section 2 for details). Few works address the problem of conceptual modeling on comprehensive applications like genetic and Pharmacogenetics research. Though a large number of works address implementation of complex biologic database systems, especially distributed ones through recent technology like CORBA [8] [9], they address the difficulty of integrating heterogeneous data sources rather than the complexity of one comprehensive data model.

In this paper, we present a conceptual model on genetic studies and Pharmacogenetics using Unified Modeling Language (UML); the standard for object-oriented analysis and design [14]. Our conceptual model supports biomedical research activities that use both clinical and biomedical data to improve patient care by incorporating the roles of environment, life style and genetics. We developed class and package diagrams for these domains to clearly and intuitively show inter-object relationships at different levels of complexity. We used UML to represent conceptual models. Conceptual models can be the basis for diverse research activities for bioinformaticians.

One distinction between our research and previous studies is that our conceptual model covers more comprehensive and integrated data to support various genetic studies and pharmacogenetic research. Thus, the model is more useful for the requirements of complex and comprehensive real world projects. The rest of the paper is organized as follows: Section 2 reviews the related work. Section 3 presents our conceptual models. Section 4 concludes our paper.

2 Related Work

In the literature, many publications address the topic of conceptual modeling of bioinformatic data and processes. Papers range from representations of single type of data like genome sequences [16] [17], protein structures [6] [18], protein interactions [3] [16], and metabolic pathways [4] [10] [21], to complex applications such as genetic study and Pharmacogenetics that often cover experimental data and knowledge management system.

Recently, conceptual modeling of genome sequence has received a lot of attention. Approaches include object-oriented models [16] [22], semi-structured models [5], relational [1], and extended ER models [17, 18]. Ram and Wei [18] create their own notations such as Sequences, Sequential Aggregate, Subsequences and Fragment to express semantics of biologic sequences. The work by Paton *et al.* [16] is most similar to ours in terms of approach. Both use UML notation to address representation of biologic data and process. Their conceptual model covers genomic data, transcriptome data, and protein interactions. Their work, however, does not include Pharmacogenetics. Thus, our work is complementary to the work by Paton *et al.*

Metabolic pathways are an important source to discover the gene functionality [10]. Usually it is helpful for provision of gene candidates for genetic studies. How to represent, store, compare and mine metabolic pathway data is also a hot topic for the research community. Schreiber [21] represents pathways as a directed bipartite graph and develops an algorithm to compute a better visual comparison of metabolic pathways. Like Schreiber, Heymans and Singh [7] use graphs to represent enzymes involved in metabolic pathways and, further, offer an iterative approach to score the similarity of metabolic pathways among multiple organisms.

With many fundamental biologic databases available, researchers in bioinformatics turn to apply established databases to address complex real world problems such as cancer prevention, diagnosis and treatment. Usually those research activities involve complex experimental data and multiple-source knowledge systems. One interesting work is the ontology development for a Pharmacogenetics knowledge base (PharmGKB) [24]. The research group of PharmGKB implemented and compared the ontological and relational approach to genomic sequences for Pharmacogenetics [20]. Further, they used a frame-based system to represent ontology for experimental data and domain knowledge [15]. Their work on modeling of pharmacogenetic data is similar to ours. We borrow some ideas from them and use a different approach, UML, to model data for Pharmacogenetics as well as genetic studies.

3 Systems and Methods

This section presents the information models using the package diagram and class diagram notation of the Unified Modeling Language (UML). Due to the high complexity of Information models for genetic studies and Pharmacogenetics, all classes are not shown in one big diagram, but are split into subsystems each of which is represented by a package. A package in UML is a mechanism that groups semantically related items. A subsystem, an independent part of the system, is usually comprised of classes loosely coupled and very cohesive.

In UML class diagrams, classes are drawn in rectangles with the class name at the top and optionally with attributes and operations listed below. In this paper, we only list attributes important to understanding the model for the sake of space saving. Besides, inter-class relationships including generalization, aggregation, association and realization and its cardinality are shown in the diagram. The remainder of this section is organized as follows: Section 3.1 shows the overview of the whole system as a hierarchical package diagram; Sections 3.1-3.6 show the class diagram of each package.

3.1 A System Model

The system model shown in Figure 1 presents all subsystems in the form of a package diagram. A package is rendered by a tabbed folder and inter-relationships between two packages are represented by a dotted arrow. Because each subsystem depends on others, only heavy dependencies are shown in the diagram.

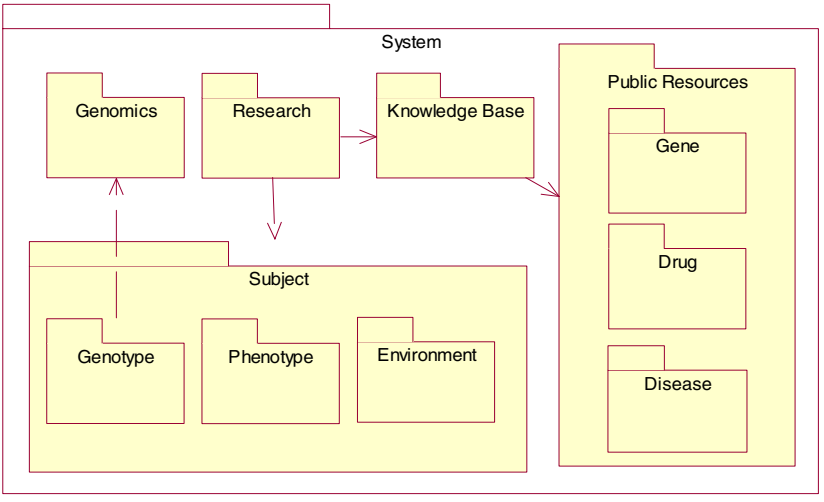


Fig. 1. A System Model of Data for Genetic Studies and Pharmacogenetics

Subject package contains classes related to experimental data of subjects including phenotype, genotype and environment. Models for phenotype, genotype and environment are still quite complex; thus, they are separated into *Phenotype* subsystem, *Genotype* subsystem, and *Environment* subsystem, respectively.

Research package includes research dataset preparation, research approaches, and results. Dataset preparation heavily depends on the subject package. A dotted arrow from *Research* package to *Subject* package shows the dependency.

Knowledge base serves as an important supporting tool for genetic studies and pharmacogenetic research. First, it provides controlled vocabulary and alternative names of objects in the system, which allows us to query, retrieve, merge and analyze data more efficiently and effectively. Second, annotated knowledge about relationships among genes, diseases and drugs allows us to interpret evidence of intermediate or final research results.

Genomics and *Public Resources* are two other important packages. However, the details of these two packages are not shown in the paper due to the limit of space.

3.2 A Subject Model

The genotype is defined as all or part of the genetic constitution of an individual or a group [25]. Environment includes subjects' lifestyle, living environment, and demographic information. The phenotype represents visible properties of an organism produced by interaction of the genotype and environment [25].

The genotype here is denoted by a subset of polymorphisms of interest. SNP is a site in the DNA where different chromosomes differ in their base. Haplotype [27] is a set of

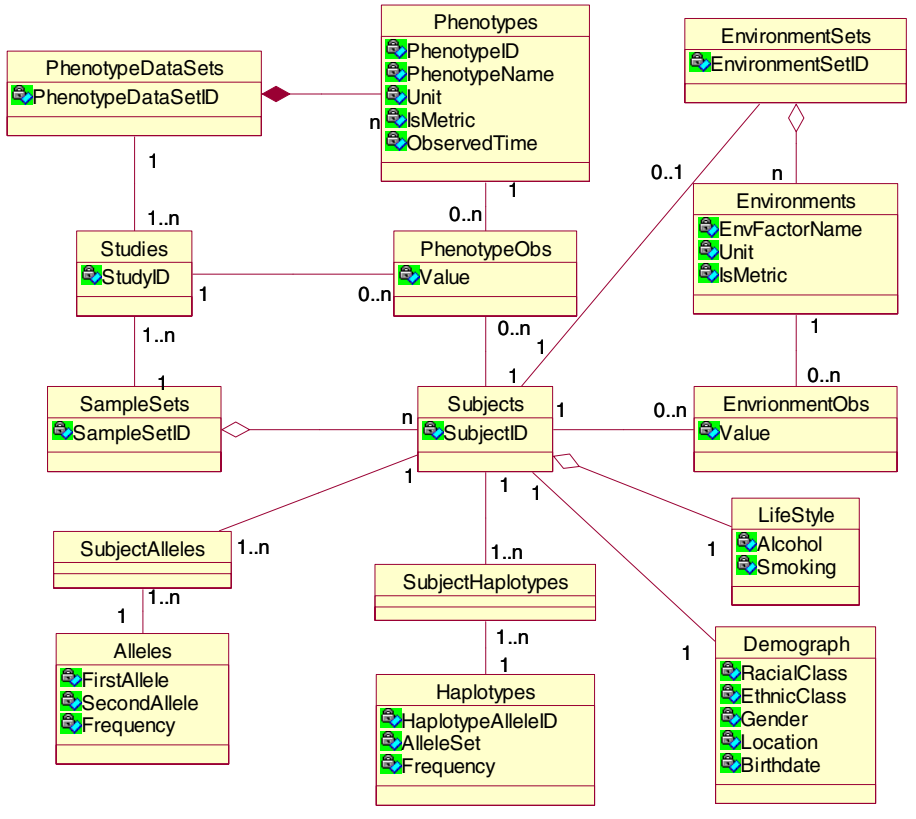


Fig. 2. A Subject Information Model shows the triad of relationship among genotype, phenotype and environment

Single Nucleotide Polymorphisms (SNPs) of interest in a DNA sequence block. The motivation of the Haplotype for research is explained in the section on the Genotype Model. A subject is associated with multiple alleles of polymorphisms and with multiple haplotype alleles.

Phenotypes of interest depend on the specific clinic trial or research experiment. For the sake of flexibility, each clinical trial or research is associated with an aggregation of phenotypes. Since observed phenotypes could vary across research experiments, an object called Trial Observation is created to store the observation of the subject phenotype in a specific clinical trial or research.

Environment not only describes the general sense of living environment factors, but also includes lifestyle and demographic information of the subject. Similar to phenotypes, living environment factors of interest are specific to a clinic trial or a research experiment. Distinct from phenotypes, although living environment factors of interest may be different for subjects belonging to different samples, they are fixed for a subject across trials or research. Therefore, an aggregation of environment factors is

created to associate with a subject rather than a trial. Attributes of demographics of interest, such as ethnic or racial class, almost reach consensus. Lifestyle attributes, such as smoking or not, are in the same case. Each subject is simply associated with one demographic and one lifestyle object.

3.3 A Research Model

Availability of map, both physical and genetic, provided the infrastructure to boost studies linking phenotype to the interaction of the genotype and environment [11] [12]. Two particular study areas of interest are genetic studies and Pharmacogenetics.

Genetic studies use genetic markers including SNPs and Short Tandem Repeats (STRs) as tools to identify the genes responsible for disease susceptibility [28]. Researchers typically utilize one of two approaches, linkage and association studies. Linkage studies typically use families with multiple affected individuals, ideally including three or more generations to identify genetic regions more likely inherited with a disease or biologic response than expected by random chance [28].

Linkage studies are useful to indicate single-gene disorders, but not suitable to identify genes involved in polygenetic disorders. Polygenetic disorders are more challenging because the multiple disease genes tend to diminish statistical significance of a linkage to any one gene. Conversely, association studies that compare genetic differences in case and control samples provide more statistical power to identify disease susceptibility genes [28]. Modern association studies can be categorized into three groups: candidate-gene-based association studies, candidate-region-based association studies, and whole genome association studies.

Pharmacogenetics is a discipline that seeks to describe how inherited differences in genetic sequences among people influence their response to drugs [13, 19]. Pharmacogenetics helps determine why some medicines work better for some people than others and why some people are more likely to experience serious side effects. Knowledge that scientists gain from this research results in the delivery of safer, more effective medicines.

For any genetic study or pharmacogenetic research, a research data set associated with a sample and a data structure is required. For both association studies and pharmacogenetic research, samples consist of both case and control subjects. For linkage studies, family-based samples include only individuals with diseases.

A data structure is defined as a variable set necessary for the research. The variable set includes three types of variables; genotypes (polymorphisms or haplotypes), environment, and phenotypes. Usually phenotypes are treated as dependent variables, while genotypes and environmental factors are as independent variables. Polymorphisms here usually refer to SNPs as they are the most powerful genetic markers thus far. A haplotype is a set of ordered SNP alleles in a region of a chromosome. The motivation for introduction of haplotypes will be explained in Section 3.4. Environment factor here has a broad sense; it refers to subjects' demographic information, lifestyle and living environment factors. A data structure is associated with either one genetic disease or one or multiple drugs. In some cases, researchers test the response of combined drugs on subjects.

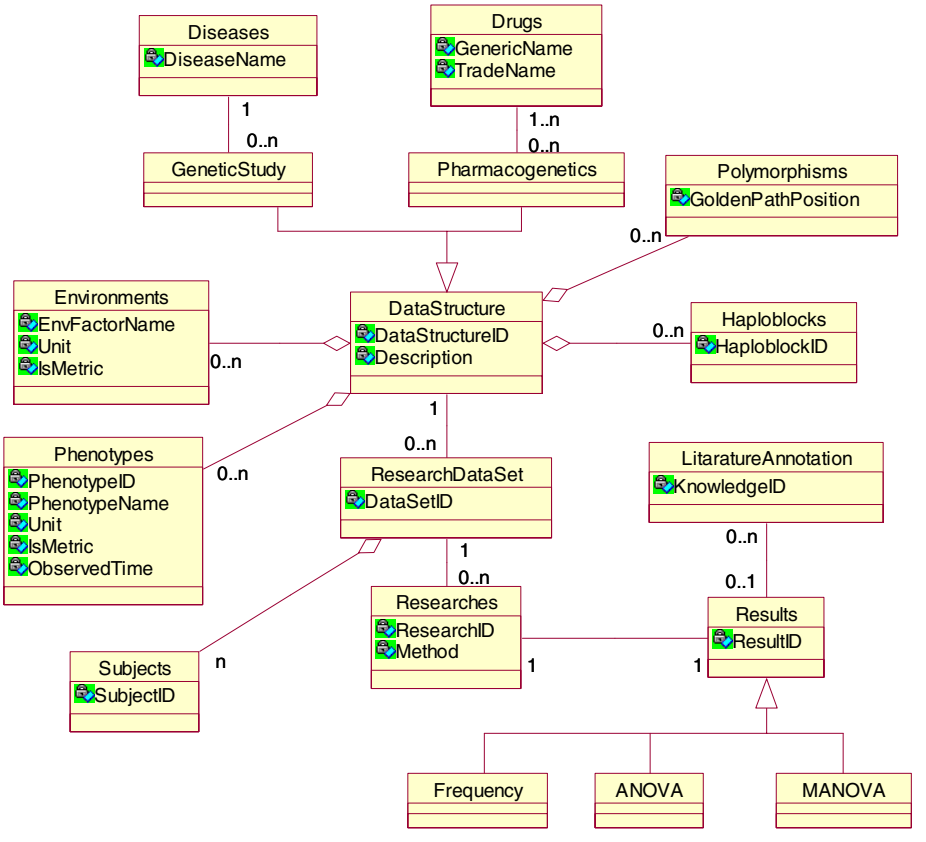


Fig. 3. A Research Model presents the data items for genetic studies and Pharmacogenetics and the intermediate or final research results

A research data set can be used by many research experiments which apply a specific method such as frequency, ANOVA (Analysis of Variance), and MANOVA (Multivariate Analysis of Variance). A research will yield a result; format of research results depends on the statistical approaches applied. Some significant findings might be published and stored in the knowledge base. Significant findings can be roughly categorized into two groups. One is binary relationships between phenotypes and genotypes. For instance, a gene is associated with or without a disease; a set of SNPs is associated with or without a drug response. The other is to the degree to which the genotype affects the phenotype. For instance, different alleles might influence disease.

3.4 A Genotype Model

Herein, an individual's genotype is defined as a set of polymorphisms of interest. A polymorphism is a locus in a reference sequence with a variable length, where the sequence content might be different against individuals and each of possible sequence

content is called an allele. Thus, a polymorphism has at least two alleles. The main types of polymorphisms include SNPs, STRs, Insertions and Deletions. SNPs and STRs have played significant roles in genetic research as powerful genetic markers due to their high densities in a genomic sequence.

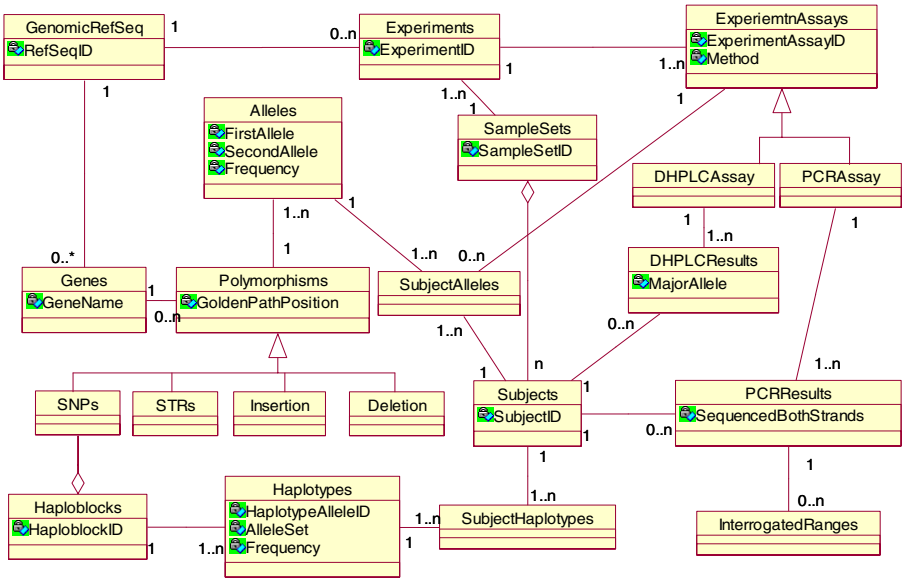


Fig. 4. A Model of Genotype and Experiment Essays shows the genotype of each subject that is observed from a variety of experiment assays

The number of SNPs in humans is known to be between 10 and 30 million SNPs. To find the regions with genes that contribute to a disease, the frequencies of many SNP alleles are compared in individuals with and without a disease. When a particular region has an SNP allele that is more frequent in individuals with a disease than those without the disease, those SNPs and their alleles are hypothesized to be associated with the disease.

A haplotype is the set of SNP alleles along a region of a chromosome. Theoretically, there could be many haplotypes in a chromosome region. But some recent studies found that a haplotype occurs in a block pattern: the chromosome region of a block has just a few common haplotypes, followed by another block region also with just a few common haplotypes [27]. The recent studies also show that the common haplotypes are found in all populations studied, and that the population-specific haplotypes are generally rare [27].

The cost of genotyping is currently too high for whole-genome association studies. If a region has only a few haplotypes, then only a few SNPs need to be typed to determine which haplotype a chromosome has and whether the region is associated

with a disease. That is the reason haplotypes are introduced for genetic studies and other genotype related studies.

To detect the alleles of a set of polymorphisms on a certain subject, we need experiments. Usually a sample set is assigned to a group of experiments, each of which is to query a set of polymorphisms within a reference sequence. Each experiment contains a group of assays. The purpose of an assay is to detect a particular polymorphism. Different types of assays are designed to detect different types of polymorphisms. For instance, a genotyping assay can detect the allele of an SNP, and a Polymerase Chain Reaction (PCR) sizing assay can detect the allele of an STR. Though there are at least eight types of assays available, most of them just determine the allele of a polymorphism for a subject and do not yield additional information. For this reason, only an attribute named method indicates the type of the assay. Denaturing High Performance Liquid Chromatography (DHPLC) Assay and PCR Assay produce additional information so that child classes are created for them. DHPLC Assay is associated with DPHLC result that additionally analyzes whether the allele is major or not.

3.5 A Phenotype Model

The phenotype represents visible properties of an organism produced by interaction of the genotype and environment. In this paper, phenotypes refer to disease symptoms and drug responses. Phenotypes of subjects are mainly collected either from clinical observations or laboratory tests with the permission of patients. In some cases, patients voluntarily submit their clinical profile for a research purpose.

Standardization of phenotypes across studies is important to the success of systems. The standardization brings at least two advantages. First, standardized measures will reduce learning cost and misunderstanding among researchers. As soon as a submission is accepted by the system, the data set will be used by various research groups in the world. Second, it makes the merge of separately small size sample sets from different research groups possible. In order to obtain stable statistical findings, most research experiments have limitations on sample sizes. For example, a candidate-gene-based association study might involve more than 1000 samples including cases and controls [28]. However, it is extremely costly to collect samples for various reasons and most separate sample sets contain much less than 1000 samples. If all the measures in different sample sets are standardized, the problem can be solved easily by ad hoc merging for a specific study.

In general, the standardization of phenotype measures involves the following three problems. The first problem is the unit inconsistency. Most phenotypes measures are metric. So an attribute called Unit is to identify the unit of a measure.

The second problem is the term inconsistency. It is very common that a term has multiple alternative names in the areas like medicines, biology and bioinformatics. In our models, three methods are used to relieve this problem. A normalized phenotype and a hierarchically organized attribute are employed to modify phenotypes. Besides, in the model of knowledge base systems, which we will explain in Section 3.6, all synonyms for a term are listed.

The third problem is the dimensional limitation. Now the system can only accept two-dimensional data. One dimension is a subject or a case. The other is a phenotype measure (or column). However, one phenotype is often measured several times on a subject at different time points for a real world clinical study. In our model, a phenotype name might include observing time information, "Blood Pressure Day 1", "Blood Pressure Day 10", for example. And a phenotype is associated with a normalized phenotype which does not include time dimensional information. We also set an attribute to store time dimensional information for each phenotype measure, which enables at least semi-automatic phenotype measure comparisons across sample sets. For the flexibility, we assign a new phenotype ID to a phenotype submitted by any research groups because we are not sure it is identical to any extant phenotypes in the system the moment the data is submitted.

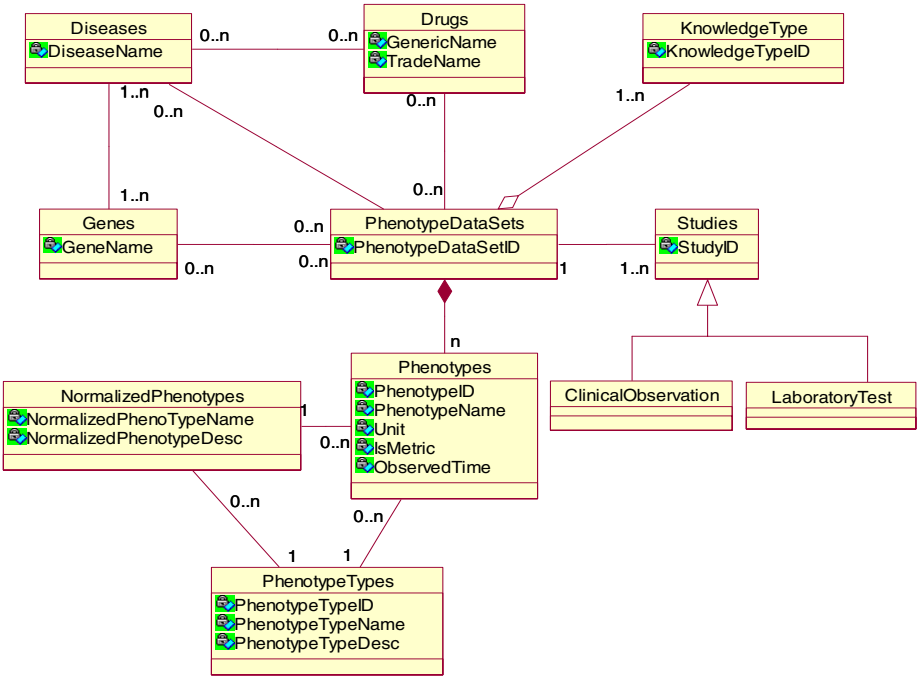


Fig. 5. A Phenotype Model addresses the representations of phenotypes in clinical trial or other experiments

For any clinical studies or research experiments, more than one phenotype will be observed or tested. Thus each study is associated with one phenotype dataset. And the same phenotype dataset might be used for different studies conducted by the same research group. For the reason stated in the previous paragraph, a phenotype can belong to only one phenotype dataset. A phenotype dataset may involve multiple genes, diseases, and drugs (for Pharmacogenetics only).

A phenotype dataset is always designed on purposes. Like PharmGKB [24], we represent the purpose by a set of knowledge types. In PharmGKB project, knowledge is categorized into five groups: clinical outcome, pharmacodynamics and drug responses, pharmacokinetics, molecular & cellular functional assays, and genotypes.

3.6 Knowledge Representation Model

The knowledge base contains only domain knowledge. The experimental data, for example genetic assays or tests, medicine response tests, and clinic studies, are not represented in the model. The knowledge system provides the following four types of knowledge [15]: controlled vocabulary terms, alternative names, accession numbers, and literature annotation.

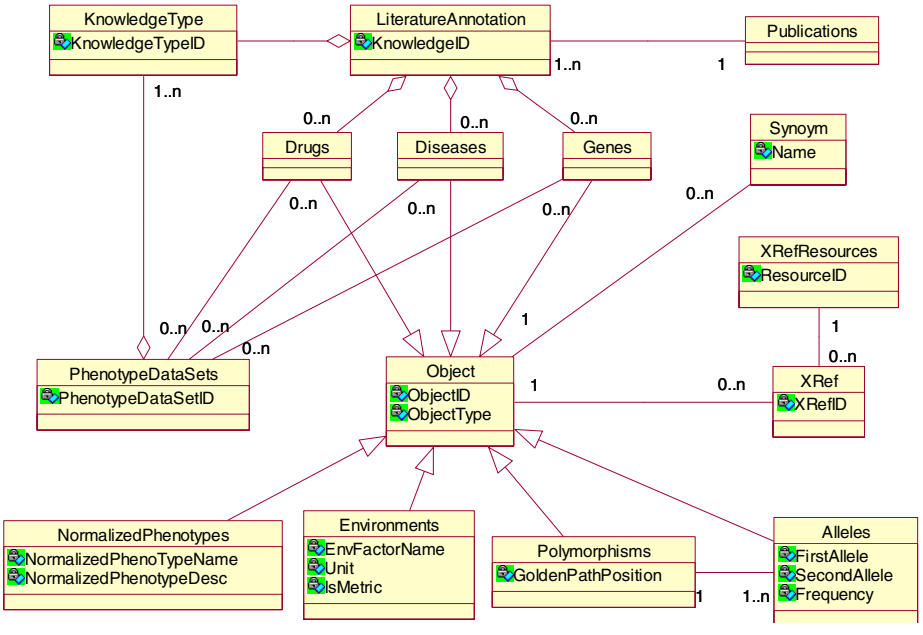


Fig. 6. A Model for Knowledge Representation provides the ontological information of drugs, diseases and genes that is evidenced by scientific literature

Controlled Vocabulary Terms. The submitted experimental data must use the name of an object defined in the knowledge base. This convention ensures the consistency of concepts through the system. The controlled vocabulary terms include drugs, diseases, genes, normalized phenotypes, environment, polymorphisms, and alleles. They are all inherited from the common superclass *Object*.

Alternative Names. Objects may have many synonyms. Maintenance of alternative names in the system assists a user in searching for a concept.

Accession Numbers. Accession numbers are unique identifiers for entities in external databases. Accession numbers are stored in object XRef (cross references) to facilitate communications with those databases. An object can have more than one cross reference which is associated with one standard resource such as dbSNP [33], dbSTS [34], GenBank [29], MedLine [35], or OMIM [32].

Literature Annotation. Genetic studies and Pharmacogenetics explore relationships among genes, diseases, and drugs. One relationship can involve multiple genes, diseases, and drugs. The instance of such a relationship is first of all normalized by an aggregation of knowledge types, and then annotated by a publication which contains the findings with respect to the instance of the relationship. From the predefined types the knowledge is associated with, a user may have ideas like what the knowledge is about. By further reading the abstract or full text of the publication, a user knows the details.

Though the domain knowledge base system has no direct contributions to genetic studies and pharmacogenetic research, it is a very important supporting tool. First, the literature annotations serve as interpretations or evidences for intermediate or final findings produced by statistical approaches. Second, it is important to the reparation of research data.

4 Conclusions and Discussions

In this paper we presented a conceptual model for genetic studies and Pharmacogenetics. Comparing previous studies in conceptual modeling of bioinformatics data and process, our model provides a comprehensive view of integrated data for genetic studies and Pharmacogenetics by incorporating genomics, experimental data, domain knowledge, research approaches, and interface data for other publicly available resources into one model.

We used the class diagram notation of UML to represent the conceptual model. Observing the notation of inter-class relationships such as generalization, realization, aggregation, association, and their cardinalities, a user can intuitively understand the schema of the model. In order to show different levels of complexity of the model, we organized the whole system as a hierarchy of packages. By aggregating a set of low coupled and highly cohesive classes into a package, the whole system can be split to a set of subsystems. The hierarchical package diagram is helpful for users to easily understand the complex system. We believe that our conceptual models are comprehensive and easy to understand. They can be a basis in supporting various aspects of Pharmacogenetics and genetic studies.

Bioinformatics is a new field. Development of genome-based approaches to disease susceptibility and drug response is still in its infant stage. We can not come up with in advance all queries users will use and dataset researchers expect. Thus, it is inevitable to iteratively improve our model in practice after the system is implemented.

In the future, we plan to enhance our model as follows. First, we will enhance the model to capture recording of experimental approaches and procedures as well as recording experimental results. Second, we also plan to support longitudinal studies as

well as cross-sectional studies. Experimental data like genotype, phenotype and living environment factors are organized around subjects. Though longitudinal data of a subject can be stored and accessed in the current model, a more extensive support may be desirable.

References

1. Berghozlz, A., Heymann, S., Schenk, J. and Freytag, J.: Sequence comparison using a relational database approach, *Proceedings of the International Database Engineering and Application Symposium (IDEAS)*, Montreal, August 1997.
2. Collins, F., Green, E., Guttmacher, A., Guyer, S.: A vision for the future of genomics research, *Nature*, Vol. 422, 24 April 2003.
3. Eilbeck, K., Brass, A., Paton, N., and Hodgman, C.: INTERACT: an object oriented protein-protein interaction database, Proceedings of Intelligent Systems in Molecular Biology (1999) 87-94.
4. Ellis, L.B., Speedie, S.M., and McLeish, R.: Representing metabolic pathway information: an object-oriented approach, *Bioinformatics*, Vol. 14, (1998) 803-806.
5. Garcia-Molina, H., Papakonstantinou, Y., Quass, D., Rajaraman, A., Sagiv, Y., Ullman, J., and Wildom, J.: The TSIMMIS approach to mediation: Data models and languages. *Proceedings of Second International Workshop on Next Generation Information Technologies and System*, June, (1995) 185-193.
6. Gray, P.M.D., Paton, N.W., Kemp, G.J.L., and Fothergill, J.E: An object-oriented database for protein structure analysis. *Protein Engineering*, 3, (1990) 235-243.
7. Heymans, M. and Singh, A.: Deriving phylogenetic trees from the similarity analysis of metabolic pathways, *Bioinformatics*, Vol. 19, Suppl. 1 (2003) i138-i146.
8. Hu, J., Mungall, C., Nicholson, D., and Archibald, A.: Design and implementation of a CORBA-based genome mapping system prototype, *Bioinformatics*, Vol.14 No.2 (1998) 112-120.
9. Jungfer, K., Rodriguez-Tome, and Mapplet, P.: a CORBA-based genome map viewer, *Bioinformatics*, Vol. 14, No. 8 (1998) 734-738.
10. Karp, P.: Representing, Analyzing, and Synthesizing Biochemical Pathways, *IEEE Expert*, April 1994.
11. Kell, D: Genotype—Phenotype mapping: genes as computer programs, *Trends in Genetics*, Vol. 18, No. 11 (2002).
12. Kim, J.: Computers are from Mars, Organisms are from Venus. *Computer*, July 2002, 25-32.
13. Krynetski, E.Y. and Evans, W.E.: Pharmacogenetics as a molecular basis for individualized drug therapy: the thiopurine S-methyltransferase paradigm, *Pharm. Res.*, 16(1999) 342-349.
14. Larman, C.: *Applying UML and Patterns*, 2nd edition, Prentice Hall (2001).
15. Oliver, D.E., Rubin, D.L., Stuart J.M., Hewett, M., Klein, T.E., and Altman, R.B.: Ontology Development for a Pharmacogenetics Knowledge Base, *Proceedings of Pacific Symposium on Biocomputing* (2002) 65-76.
16. Paton, N. et al: Conceptual Modeling of Genomic Information. Bioinformatics, Vol. 16 No. 6 (2000) 548-557.
17. Ram, S. and Wei, W.: Semantic Modeling of Biological Sequences, *Proceedings of the 13th Workshop on Information Technologies and Systems*, Seattle, (2003) 183-188.
18. Ram, S. and Wei, W.: Modeling the Semantics of Protein Structures, *Proceedings of the 23rd International Conference on Conceptual Modeling (ER 2004)*, Shanghai, China (2004).

19. Roses, A.D.: Pharmacogenetics and the practice of medicine, *Nature*, 2000 Jun 15, 405(6788):857-865.
20. Rubin, D. et al.: Representing genetic sequence data for Pharmacogenetics: an evolutionary approach using ontological and relational models, *Bioinformatics*, Vol. 18 Suppl. (2002) S207-S215.
21. Schreiber, F.: Comparison of Metabolic Pathways using Constraint Graph Drawing, *First Asia-Pacific Bioinformatics Conference (APBC2003)*, Adelaide, Australia (2003).
22. Wong, L.: Kleisli: a functional query system, *J. Functional Programming*, Vol.10, No.1 (2000) 19-56.
23. Wang, Z. and Moult, J.: SNPs, Proteins Structures and Diseases, *Human Mutation*, Vol. 17 (2001) 263-270.
24. [PharmGKB] Pharmacogenetics Knowledge Base. http://www.pharmgkb.org/
25. Primer on Molecular Genetics, taken from the June 1992 DOE Human Genome 1991-92 Program Report.
26. Reference Sequences, http://www.ncbi.nlm.nih.gov/RefSeq/
27. Developing a Haplotype Map of the Human Genome for Finding Genes Related to Health and Disease. http://www.genome.gov/10001665.
28. White Paper: SNPs—Powerful Tools for Association Studies. August 2003, Applied Biosystems.
29. GenBank, http://www.ncbi.nih.gov/Genbank/.
30. EMBL Nucleotide Sequence Database, http://www.ebi.ac.uk/embl/
31. DNA Database of Japan (DDBJ), http://www.ddbj.nig.ac.jp/
32. Online Mendelian Inheritance in Man (OMIM), http://www.nslij-genetics.org/search_omim.html
33. dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
34. dbSTS, http://www.ncbi.nlm.nih.gov/dbSTS/
35. MedLine, http://www.medline.com/

A Dynamic Parallel Volume Rendering Computation Mode Based on Cluster

Weifang Nie, Jizhou Sun, Jing Jin, Xiaotu Li,
Jie Yang, and Jiawan Zhang

IBM New Technology Center, Tianjin University, Tianjin 300072, P.R.China
Luna_pure@hotmail.com

Abstract. Based on cluster, a dynamic parallel mode–the MSGD (Master-Slave-Gleaner-Dynamic) mode is presented in this paper. Combining the load balancing with the MSG (Master-Slave-Gleane) mode, this mode brings in a load balancing strategy–task pool, which can coordinate the workload of each node effectively. A volume rendering algorithm called Splatting is used to test the performance of this mode on IBM Cluster1350. The results prove that this mode can effectively increase the total calculation speed and improve the global load balance. In addition this mode has good scalability and is suitable for various parallel volume rendering algorithms with large-scale data volume.

Keywords: volume rendering, parallel computing, cluster, load balancing.

1 Introduction

As the key technology of analyzing and processing three-dimensional or multi-dimensional data in scientific visualization, the volume rendering develops very fast and is widely used in scientific research and the engineering area[1]. However, with the improvement of the measure technique and calculation ability, the data scale is getting larger and larger, which puts forward the new challenges to the speed and quality of the volume rendering. It is urgent to parallelize the related volume rendering algorithm to adapt to the demands of the processing.

In the existing parallel computing modes, the Master-Slave mode is widely used. However, the task of Master node becomes too heavy with the size increment of the volume data when this kind of mode is used in the parallel volume rendering algorithm. Because of that the waiting time of the Slave nodes rises and the global algorithm performance declines. On the basis of MS mode, the MSG mode[2] was presented to reduce the workload of the Master node. But usually only common static load balancing is carried out in this mode, there is no further consideration of the load equilibrium problem. As a result the Slave nodes with the high calculation ability idle, and each node cannot exert its all capacity.

To address this difficult problem, the authors combine the features of the volume rendering with the cluster to bring forward the MSGD mode–the MSG

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 416–425, 2005.

dynamic parallel mode for the volume rendering. Dynamic load balancing is brought into the MSG mode to improve the performance of parallel algorithm. The Splatting algorithm is used to test the performance on IBM Cluster1350 and the relevant data are obtained from the experiment to evaluate this mode. It is indicated that MSGD mode can effectively improve the global load balance and raise the speed of parallel procedure.

The rest of this paper is organized as follow. In Section 2, the authors give a brief review of volume rendering algorithm and the problems it has. In Section 3, the authors introduce the parallel computing and cluster. The MSGD mode based on IBM Cluster1350 is presented in Section 4. Experimental results are reported in Section 5. Finally, the authors give our conclusion in Section 6.

2 Volume Rendering

Volume rendering technique can effectively render the three-dimensional data from CT(Computed Tomography), MRI (Magnetic Resonance Imaging), CFD (Computational Fluid Dynamics). And it is able to display the object from any sides in high quality[3]. The volume rendering algorithm is paid more and more attention to in the medical science, geology science, calculation flow mechanics areas and finds more application.

However the volume rendering technique sometimes has to deal with large-scale data sets. To be interactive, merely depending on graphics workstation with single processor is not enough. Besides, the memory capacity needed in volume rendering is gigantic. Usually it is impossible to put all the data in the memory. In fact the calculation speed and memory problems become two foremost obstacles of volume rendering technique development. It is impractical to carry out the volume rendering with large-scale data in a common PC. Therefore, parallel computing becomes the most practicable method to solve the problem.

3 Parallel Computing and Cluster

With development of science and technology, the requirements of the high computation speed in numerical modeling and simulation in science and engineering are increased as well. In such areas, it is usually necessary to carry out a lot of repetitious calculations to obtain results in a reasonable time interval. A method of accelerating the speed is to use several processors together. The whole task is cut into small parts and every part is computed synchronously on an independent processor. This is the main idea of parallel computing. It has many advantages in calculating and processing huge numbers, so it is widely employed in the various scientific areas, including the graphics area. Bringing the parallel computing into the volume rendering algorithm is able to fully speed up the operation rate and improve the displaying quality.

Cluster is one of the most important platforms for parallel computing. It is made up of high performance workstation or the PC machines that are linked

by the high-speed networks. It has merits of great scalability, high usability, the high performance and high price to performance ratio. A typical cluster system must have a low delay communication protocol and its coupling is lower than SMP, especially its network is faster and more expert than normal LAN.

The Cluster used in this paper is IBM Cluster1350, which is made up of several independent computers connected by the network. Each computer has its own processor and local memory. None of the processors can access to the remote memory. Data is shared across a communications network using message passing.

4 The Research and the Implementation of the MSGD Mode

4.1 The MS Mode and MSG Mode

In parallel volume rendering algorithm, MS mode is widely used. This mode consists of a Master node and some Slave nodes. The Master node controls the procedure and is responsible for starting up other subprograms, collecting results, and showing the calculation results. In the volume rendering algorithm, its main duty is to start up other subprograms, initialize the procedure, divide the data, collect the result data and display picture. Subprogram is a computing node. It carries on the calculation of the actual problem. The information transmission uses message passing.

The MS mode is very flexible and has good scalability. It supports parallel in task and algorithm level and the stability of it is very strong. However, as shown in last paragraph, the assignment of Master node is too heavy when it is used in the volume rendering algorithm. Furthermore the characteristics of low bandwidth and high communication delay also exist in cluster system. As a result, the communications traffic increases greatly, which raises the waiting time of the Slave node and reduces the running efficiency. A new kind of mode–the MSG mode was proposed to improve the efficiency of parallel volume rendering algorithm.

The MSG mode adds a new kind of node: Gleaner node. Fig.1 shows the architecture of the Master-Slave-Gleaner mode. This mode decomposes the work of the Master node in MS mode and gives out a part of the Master node's work to the Gleaner node. The Master node is responsible for starting up the subprograms, collecting the result data and showing the picture. The Gleaner is responsible for dividing the task and sending out the data. The Slave node is only responsible for the calculation.

As shown in Fig.1, dividing tasks and collecting results are independent in the mode, and then a pipeline operation is performed. This reduces the work of Master node consumedly, improves the load balancing, decreases the waiting time of Slave nodes and raises the global executive performance. It has all advantages of the MS mode and it is simple. In addition, the MSG model has good practicability and scalability.

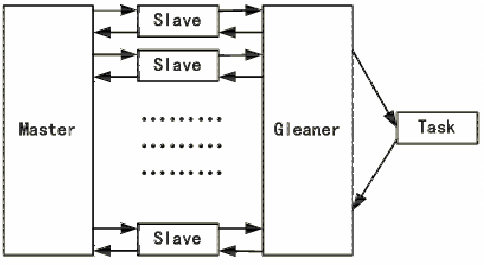


Fig. 1. The architecture of the MSG mode

However, the existing MSG mode usually achieves the static load balancing. It evaluates the capability of processors, which according to computation tasks are arranged, during compilation before the program running. The strategy is simple and its communication expense is low. But it is hard to assign task according to the actual work ability of the processor and impossible to make full use of each processor. Therefore, computation time usually depends on the slowest processor and the load balancing is not easily attained when using the static task assignment. In addition, the static load balancing could not take the influence of the extraneous factors into account, such as the communication delays in some systems. In static load balancing it is very difficult to show the variety of the communication delays. Considering the characteristics of parallel volume rendering algorithms and combining the dynamic load balancing with the MSG mode, the MSGD mode has been designed and implemented.

4.2 The MSGD Mode

According to the feature of volume rendering and cluster system, the MSDG mode combines two key techniques–the MSG mode and the dynamic load balancing, and adopts task pool to improve the load balancing.

In parallel computing, generally a problem is divided into several parallel processes. And each process does certain work. However, since the work is not divided on average or some processors run faster than other ones, there are always some processors that complete their task more quickly than others and become idle. The ideal circumstance is that all processors are in the busy situation all the time and continually do their tasks. In this case, the running time is the shortest[4]. Load balancing refers to achieving the best efficiency of parallel execution by the distributing tasks to the processors.

When designing the load balancing, it is necessary to keep in mind of the network structure, parallel cluster architecture, utilization and reliability of the system. There are two kinds of methods in common use: the static load balancing and the dynamic load balancing. The dynamic load balancing is adopted because of the defects of the static load balancing mentioned above.

In the dynamic load balancing, the whole work is completed according to the actual ability of each processing node. The tasks are transferred dynamically to

reach the load balancing, which need investigate into the processing capacity of each node. In dynamic load balancing the task is arranged to the processors in the running time. The authors adopt task pool for the dynamic load balancing.

Fig.2 is a sketch map of task pool. As shown in it, firstly the whole problem is divided into several small tasks, which forms a big task pool. After each processor completes its own existing task, they ask for new task and execute the new task continuously until the task pool is empty. For example, if there are n Slave nodes the problem could be divided into m tasks (usually m is the integral multiple of n) to make up a task pool. Firstly the Master node starts up the Gleaner and Slave nodes, and then the Gleaner node divides the problem to small ones, creates a task pool and assigns the first tasks to the Slave nodes. When each Slave node completes the task independently, the results are sent to the Master node and the Slave nodes apply for next task toward Gleaner node at the same time. If the pool is not empty, the Gleaner node assigns next task to it and cancels this task simultaneously, otherwise the Master node closes the Slave node. As a result the Slave nodes with high calculation capability can calculate more data than the ones without.

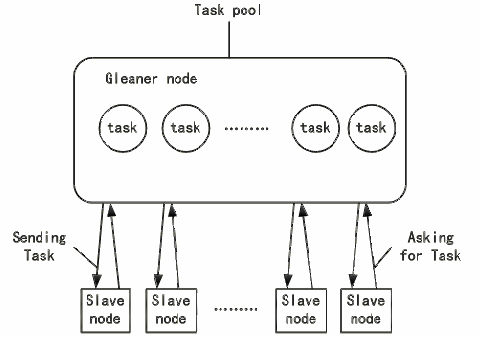


Fig. 2. The working map of task pool

In parallel volume rendering algorithm, the work of the three nodes in the MSGD mode is: the Master node is responsible for starting up other nodes, collecting the result data and displaying the picture; the Gleaner node is responsible for dividing the task, creating task pool and sending data; the Slave node is responsible for calculation.

In this mode, the task can be sent out to each Slave node according to the actual workload of the processors in any running time. The node with strong computing capacity is assigned with more work. Each processor is busy with the work and no processor is idle. Although it results in additional communication expense, it is much better efficient than the static load balancing. And with the increment of the volume data, the ratio of the cost time in connection to the total time will descend accordingly. As a result the computing efficiency increases continuously. It is suitable for the large-scale parallel volume rendering.

In section 5, the Splatting algorithm is used to test the performance of this mode based on IBM Cluster1350 system. Analyzing the data, the features and advantages of this mode will be shown in the following.

5 Experiment Analysis

IBM Cluster1350 is a new Linux cluster that is used in the paper. It is a consolidation and follow-on of the IBM Cluster1300 and the IBM xSeries "custom-order" Linux cluster offering delivered by IGS[5]. The system provides an array of computational and storage devices connected through high-speed networks. Cluster1350 is an engineered system, not just a collection of components, and therefore provides the following advantages: interoperability, increased reliability, ease of installation and service, electromagnetic interference (EMI) compliance, safety and ease of expansion[6]. Also, it is based on the operate system of Linux and is ideally suitable for large-scale and medium-scale volume rendering.

There are 7 sub-nodes and a management node with 10/100/1000Mbps Ethernet in our cluster system. The operation system on each node is Linux (RedHat7.3). Computing nodes connect one another by two high speed Ethernet switches (one is for computing (1000M), and the other is for system management (100M)).

The parallel programming tool adopted in this paper is PVM (Parallel Virtual Machine) software. It provides a unified framework within which parallel programs can be developed in an efficient and straightforward manner using existing hardware. PVM enables a collection of heterogeneous computer systems to be viewed as a single parallel virtual machine. PVM transparently handles all message routing data conversion and task scheduling across a network of incompatible computer architectures[7].

A Splatting program, a kind of volume rendering algorithms with large complex volume datasets, is used to compare the performance of the MSGD with that of the MSG mode.

The Splatting algorithm is an international popular volume rendering algorithm. The main idea of this algorithm is to calculate the contribution of the volume data to the screen pixel layer by layer, line by line, one by one, then synthesize all the values, finally generate the picture. Because the footprint table can be computed in the course of pretreatment, the Splatting algorithm is quicker than other common volume rendering algorithms and can produce a picture of high quality. Moreover, because of its characteristic, it is very suitable for the large-scale data parallel visualization computation and can be used in the design of accelerative hardware.

Firstly the Splatting program is paralleled using Master-Slave-Gleaner mode and tested on IBM Cluster1350. Then the same program is paralleled using Master-Slave-Gleaner-dynamic mode and tested on the same system. The size of the volume data used in this paper is 16 megabyte. And the parallelization method used is data parallel.

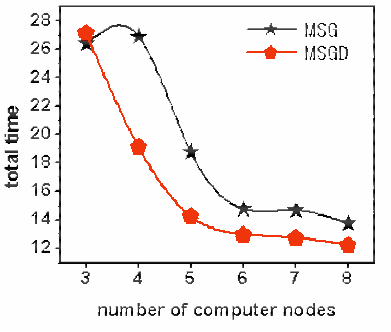


Fig. 3. The total time of Parallel algorithm using different mode

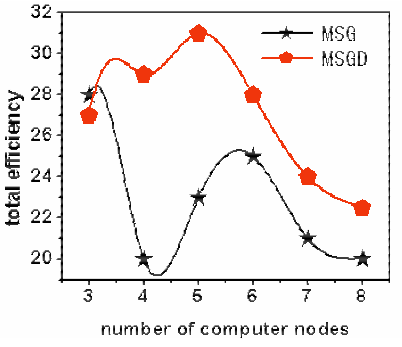


Fig. 4. The total efficiency of parallel algorithm using different mode

In the implementation of the program, the management node of the cluster acts as the Master node, one of the sub-nodes as the Gleaner node and the rest nodes as Slave nodes.

In the experiment, there are several important parameters: (1) total runtime of the parallel program, (2) total efficiency of the system, (3) the absolute velocity. In this paper the absolute velocity of program is defined as the ratio of data size to computing time. The unit of time is second and the data size is measured in megabyte. The absolute velocity is expressed by numerical part of the ratio directly. Each experiment has been done for three times. The final result is the average value.

According to the parameters and the amount of datasets, the authors get comparative figure of parallel program: total runtime figure, total efficiency figure the absolute velocity figure.

Fig.3 compares the total time of parallel program using the MSG and MSGD modes from 3 to 8 computing nodes. The horizontal axis measures the number of working nodes and the vertical axis measures the total time. Because there are three kinds of nodes totally in the MSGD mode, at least three processors must be used (suppose one process is mapped to one processor). It has been

observed from the graph that the total runtime of the MSGD mode is shorter than the MSG mode form 4 to 8 working nodes. For example, when there are 5 nodes, the total runtime of the MSG mode is 18.78 seconds and that of the MSGD mode is 14.26 seconds. Compared with the MSG model, the total time of the MSGD mode parallel program is reduced by 24.1%. However when there are 3 nodes, the costing time of the MSGD mode is a little more. This is due to the additional communication expense caused by the dynamic load balancing. Except that, the MSGD mode is quicker than the MSG mode by several seconds. This is because of the dynamic load balancing too. As a result of it the Slave nodes themselves can make full use of computing power. Therefore, the total time diminishes correspondingly.

Fig.4 compares the total efficiency of parallel program. The horizontal axis of these measures the number of working nodes, while the vertical axis measures the total efficiency. The curves of the MSGD model parallel program are much higher than that of the MSG one from 4 to 8 nodes. From the figure we know that the MSGD mode can raise the efficiency of the parallel volume rendering algorithm and has better rationality.

Fig.5 compares the absolute velocity of the parallel programs. The horizontal axis measures the number of working nodes, while the vertical axis measures the absolute velocity. It is observed that there is a continual increase in the absolute velocity from 3 to 8 working nodes in these two curves. Furthermore the absolute velocity of the MSGD model parallel program is much higher than that of the MSG one from 4 to 8 nodes. Taking example for 7 nodes, the absolute velocity of the MSG mode is 1.09 M/S, while that of the MSGD mode is 1.26 M/S and is quicker than the MSG mode by 15.3%. That is owing to that the MSGD mode raises the total efficiency of parallel algorithm, so it can reduce the total runtime and also the absolute velocity.

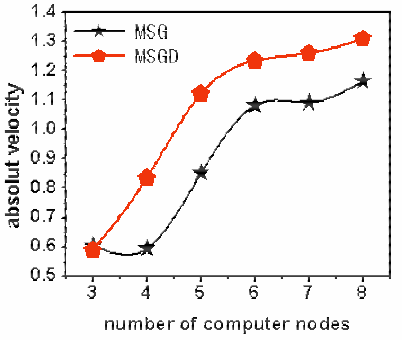


Fig. 5. The absolute velocity of parallel algorithm using different mode

The MSGD mode has all advantages over the MSG mode, and it supports the parallel volume rendering better in cluster system. Besides, this kind of load balancing has good scalability and is simple to implement. It is not limited by

the number and the power of processors and can be used in the heterogeneous network. According to practical requirement, it is possible to bring other different strategies of dynamic load balancing into the MSGD mode to attain the best efficiency.

6 Conclusion

To support the processing and visualizing of the large-scale volume data, a MSGD mode has been designed and implemented in this paper based on IBM Cluster1350. The Splatting algorithm is used to test the performance of this mode. This mode links the dynamic load balancing and the MSG mode together, so it can improve the load balancing. The result of experiment proves that this mode can raise the total performance of the parallel volume rendering algorithm and surpasses the static MSG mode in the computing efficiency, total runtime and the running speed. In addition, calculation efficiencies of this mode would increase with the increment of data scale and it is suitable for the parallel volume rendering algorithm and dealing with large-scale data.

The merits of the MSGD mode are: (1) it can improve the performance of the parallel volume rendering algorithm; (2) it can coordinate the amount of calculation in each processor and make them incline toward the load balance; (3) its calculation efficiency will increase correspondingly with the increase of the volume data scale; (4) it has good scalability and is not limited by the number of the processor.

The disadvantage of this mode is that its dynamic load balancing strategy of is not universally used. That is to say, it should adopt the different strategy to load the equilibrium according to actual circumstance, such as the hardware platform and the network condition. Task pool, a popular strategy of dynamic load balancing, is used in this paper because of its cluster platform. But it has not been proved that this strategy is suitable for all circumstances. It should be found out that the best dynamic strategy according to the specific algorithm and hardware platform.

Acknowledgements

This work is supported by Natural Science Foundation of China under grant No. 60373061. The authors gratefully acknowledge the valuable comments and suggestions of their associates in Parallel Programming Research in CADC Group. Special thanks are expressed to Ce Yu, Qianqian Han, Linlin Guo, Weiliang Meng for their interesting and rewarding comments.

References

[1] Xiaomei Li,Zhaohui Huang,Xun Cai etc.: Parallel and Distributed Visualization:Techniques and Applications, Defence industry publisher.

[2] Xiaotu Li, Jizhou Sun, Zhanhui Qi, Jiawan Zhang: A Modified Parallel Computation Model Based on Cluster. COMPUTATIONAL SCIENCE AND ITS APPLICATIONS - ICCSA 2004, PT 2 LECTURE NOTES IN COMPUTER SCIENCE 3044: 252-261 2004.
[3] Zhigang Sun, Jiawan Zhang, Jizhou Sun: A improved Splatting algorithm for volume rendering, Journal of Tianjin University, Vol.36 No.5 Sep.2003.
[4] Barry Wilkinson, Michael Allen: Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers, Prentice Hall, 1999.
[5] Ibm.com/redbooks, Building a Linux HPC Cluster with xCAT.
[6] Ibm.com/redbooks, Preinstallation Planning.
[7] G.A. Geist, A. Beguelin, J. Dongorra, W.Jiang, R.Manchek, V.S.Sunderam: PVM-Parallel Virtual Machine -A Users' Guide and Tutorial for Networked Parallel Computing, MIT Press, 1994.

Dynamic Replication of Web Servers Using Rent-a-Servers*

Young-Chul Shim, Jun-Won Lee, and Hyun-Ah Kim

Dept. of Computer Engineering, Hongik University, Seoul, Korea
{shim, junlee, hakim}@cs.hongik.ac.kr

Abstract. Some popular web sites are attracting large number of users and becoming hot spots in the Internet. Not only their servers are overloaded but also they create excessive traffic on the Internet. Many of these hot spots are appearing and disappearing rapidly. Traditional approaches such as web caching and static replication of web servers are not proper solutions to the hot spot problem. We present a dynamic replication method of web servers using a concept called rent-a-servers. The proposed method not only distributes web requests among replicated web servers evenly but also significantly reduces traffic on the Internet. We compare the proposed method with other approaches and show the effectiveness of the proposed method through simulation.

1 Introduction

With the success of the WWW technology, the number of Internet users/hosts and Internet traffic are increasing rapidly. With such a rapid spread of WWW, the traffic overload is becoming a serious problem in the Internet. Moreover, clients experience excessive delay in accessing very popular web sites. We observe frequently that some popular web sites are attracting an enormous number of users and, therefore, becoming hot spots of network traffic in the Internet. These hot spots are changing continuously. Recent studies reveal that the cases of abrupt increase in the number of users accessing popular web sites are quite common and these popular web sites are changing continuously. It is observed that hot spots causing performance bottlenecks in the Internet are appearing and disappearing rapidly. We have to consider the following issues so that we can provide services of a satisfiable level to the web clients.

- The delay for bringing web contents to the client should be minimized. This delay consists of the time to deliver a client request to the web server, handle it at the web server, and return the response to the client.
- Availability of web contents should be increased. Although one web server fails, the contents stored in this failed web server should be accessible to clients through other replicated web servers.

* This research was supported by University IT Research Center Project, Korea.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 426–434, 2005.

– The amount of network traffic carrying web requests and responses should be minimized and the phenomenon of network traffic concentrating in some areas of the Internet should be avoided.

There has been much research work to satisfy the above requirements and this work can be classified into two categories[1]: Web caching and the replication of web servers. But as will be explained in the next section, these approaches are not suitable to handle the problem of hot spots. Recently there emerged the concept of rent-a-servers in which fast computers with a big storage space are installed in the Internet and are rented to users who need fast computation and/or a large storage space[2]. This concept lends itself to handling hot sports in which client requests increase abruptly at unpredictable time and locations. If a sharp increase in client requests is observed at a certain web site and, therefore, this web site is decided to be a new hot spot, we borrow a rent-a-server at a proper location and replicate a web server at this rent-a-server to distribute loads among web servers. If client requests dwindle after some time, we uninstall the replicated web server from that rent-a-server. With this approach we do not have to purchase and install a computer to replicate a web server but need to pay the fee for the rental time of a rent-a-server. The decision of both when/where to replicate a web server and when to uninstall the replicated server can be automatically made based upon the real-time analysis of the client request traffic pattern. So rent-a-servers can be a cost-effective solution to the hot spot web sites that appear and disappear rapidly.

In this paper we describe our approach to automatically replicating web servers for hot spot web sites using rent-a-servers to reduce both the response time and the network traffic. We also show the effectiveness of our approach through simulation. The rest of the paper is organized as follows. Section 2 presents related works. Section 3 describes the overall approach and explains algorithms for replicating and uninstalling replicated web servers. Section 4 presents simulation results and is followed by the conclusion in Section 5.

2 Related Works

Web caching stores recently accessed web documents in specialized web caching servers such as proxy servers. If the same document is requested, it can be retrieved from the cache without going to the web server. Therefore, web caching can reduce the time to handle web requests and the network traffic. There has been much research on web caching including Harvest[3] and WebWave[4]. Although web caching can improve performance in case of retrieving static documents, it cannot be used in case of retrieving results by sending parameters and executing a dynamic document on the web server. When a document is retrieved the first time and the validity of a cached document is checked, the web server should be accessed and the load on the web server increases and the network traffic concentrates on the web server. Moreover web caching cannot solve the problem of web server failure completely.

These problems in web caching can be solved by replicating and distributing web servers. An important issue in a replicated and distributed web servers is to determine to which web server among replicated servers a client request should be directed so that the load of web servers is well balanced and the response time to the client request is minimized. A distributed web server system consists of many web server hosts distributed in LAN or WAN and has the mechanism to spread client requests evenly among web severs. In this system a client request is delivered to either a lightly loaded web server or a geographically close server and, therefore, the load of client requests is spread among replicated servers resulting in the reduction of the response time. Depending upon who distributes requests among servers, a distributed web server system can be classified into four categories as follows: client-based systems[2], DNS-based systems[5], dispatcher-based systems[6], and server-based systems[7]. Another important issue in a distributed web server system is the determination of the locations where web servers are deployed. If the request pattern is stable in terms of time and location, the traffic analysis of requests can determine the location of web servers. Altavista and Yahoo are such examples. To replicate web servers, locations of servers should be determined, hosts should be deployed at the selected locations, and DNS servers or dispatchers also should be installed and properly configured. But this process takes a lot of time. Therefore, a distributed web server system cannot cope with hot spots that are created in a very short time. If a web site loses its popularity and client requests to that site drop significantly, then all the investment of deployed web servers for that site becomes wasted.

There has been some work on dynamic replication of web servers[8] but their work was focused on the replication within a single Internet Service Provider(ISP) and, therefore, the replication benefit is somewhat limited. Our approach lift that limitation and try to maximize the replication benefit.

3 Dynamic Replication of Web Servers

ISPs enable users to connect to the Internet and utilize the ISP's high-performance computers. Currently these computers are mainly used for web hosting but we can extend this concept further so that users can use these computers for any purposes and be charged for their usage. We call computers of such an extended concept to be rent-a-servers. We intend to use rent-a-servers for replicating web servers of hot spot web sites. Rent-a-servers are connected to ISP routers directly or through LAN but in this paper we treat a rent-a-server and its router as one node for simplicity. We assume that using the active network technology[9], a router can analyze a request being routed to an upstream web server and redirect it to its rent-a-server if that rent-a-server can service that request. If a web server is replicated on a rent-a-server as in Figure 1, the router and its rent-a-server can catch the client request being delivered to the primary web server, process it, and return the result to the client.

When a web site is opened, it starts with one web server. This web server is called a primary web server of that site. This primary web server, all of its

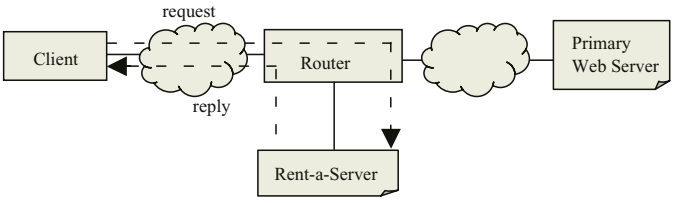


Fig. 1. User request being processed by a rent-a-server

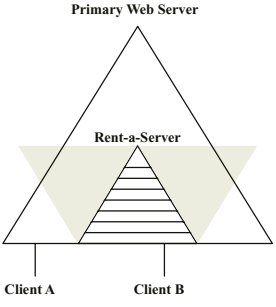


Fig. 2. Routing tree of web requests

clients, the shortest paths connecting these clients to the primary web server, and rent-a-servers on the shortest paths form one big routing tree where the primary web server is the root, clients are leaves, and rent-a-servers are inner nodes as in Figure 2. In the beginning all the requests are sent to the primary web server. But when its load becomes too high, a web server is replicated on the rent-a-server at a proper location. Thereafter, requests of all the clients in the subtree rooted by that rent-a-server is processed by the replicated server. In the figure, client A's requests are processed by the primary web server but those of B are processed by the rent-a-server. So request loads are distributed among web servers. Because B's request heading to the primary web server is caught by the rent-a-server, the transmission delay of that request is reduced and the request delivery from the rent-a-server to the primary server is saved resulting in the reduction in the overall network traffic. Now we explain algorithms for replicating and uninstalling web servers.

3.1 Replication Algorithm

To explain the algorithm we use the following terminologies.

- PS : Primary web server of a certain web site.
- TS : Rent-a-server.
- RS : Web server replicated on a rent-a-server.

At a server S that is either PS, TS, or RS, its load, $L(S)$, due to requests to the web site consists of the following 3 components.

- $L_L(S)$: Load due to requests from clients directly connected to S.
- $L_C(S)$: Load coming from all the child TS. If there is RS on a child TS, the load from that child node becomes 0.
- $L_D(S)$: Load redirected from PS or other RSs. This load exists only on RSs.

If S is TS with RS installed on it, its total load becomes $L_L + L_C + L_D$ and is processed by that RS. If S is TS without RS, its total load, $L_L + L_C$, is forwarded to its parent server. We assume that each server S records the values of $L_L(S)$, $L_D(S)$, and $L_C(S)$ for each child server.

We explain the algorithm for replicating RS on TS. PS or RS becomes overloaded when its load goes over L_{OVER}. An overloaded server tries to find TS satisfying following conditions among its descendant TSs to install RS there.

- TS does not have RS and its load is not less than L_{REP}.
- TS does not have any descendant TS satisfying the above condition.

The first condition tries to find TS with enough load (at least L_{REP}) so that the overloaded server can be offloaded sufficiently when RS is installed on TS and the cost of installing RS can be justified. The second condition tries to find TS that is closest to the client group from which excessive requests are coming. Only if RS is installed on such TS, not only the load on web servers are balanced but also the traffic due to excessive requests can be reduced. If there is no dominant client group causing the overload, and the overloaded PS or RS cannot find any TS with its load greater than L_{REP}, the overloaded web server selects TS with the biggest load among its child TSs without RS and installs RS on that TS. Then among its child TSs without RS it chooses a set of TSs and redirects requests from those TSs to the newly installed RS. The set of TSs are chosen in such a way that the load of the originally overloaded server falls below L_{REP} after the redirection. Now we present the algorithm.

```
ReplicateWS(S) { /* called at an overloaded PS or RS */
    Among child TSs without RS, select TS with the biggest load;
    If (TS has load > L_REP)
        FindTS(TS) /* TS has enough load */
    Else
        Replicate&Redirect(S) }

FindTS(S) {
    Among child TSs without RS, select TS with the biggest load;
    If (TS has load > L_REP)
        FindTS(TS) /* there is TS with sufficient load among descendant */
    Else
        Install RS on this rent-a-server }

Redirect&Replicate(S) {
    Sort child TSs without RS in the non-increasing order of load;
    Install RS on the child TS with the highest load;
```

While (load of overloaded server, S, is not lower than L_{REP}) {
Pick the next TS and arrange
to redirect requests from this TS to the newly installed RS;
Subtract the redirected load from S's total load }}

3.2 Uninstallation Algorithm

When the load of RS becomes significantly low (below L_{UNDER}), we consider unistalling this RS. If RS is unistalled, all of its load will be forwarded to its parent web server, S. If the resulting load in S, $L(RS) + L(S)$ becomes too high (above L_{ALERT}), there is the possibility that the parent server will soon get overloaded and will try to replicate a web server again. To avoid this undesirable situation, RS is uninstalled only if the sum of its load and its parent web server's load is not too high (not above L_{ALERT}).

If more than two RSs try to uninstall themselves simultaneously, the situation can become complex. So we allow RS to be uninstalled only if there is no descendant RS trying to retreat simultaneously. Now we present the algorithm.

```
UninstallWS(RS) {
    Check if there is any descendant RS trying to retreat;
    If (there is such a descendant RS)
        Wait some time and exit; /* descendant RS is trying to retreat. */
        /* so wait some time before trying to uninstall again. */
    Calculate L to be sum of its load and its parent load;
    If (L(RS) + L(ParentWebServer)> L_ALERT)
        Wait some time and exit
        /* if sum of loads is still too high at the parent server, */
        /* wait some time before trying to uninstall again */
    Else
        Uninstall the replicated server }
```

4 Simulation Results

We experimented the proposed algorithms by simulating with NS-2[10]. In the simulation we made the following assumptions.

- Load on servers are measured in terms of the queue length of client requests and the maximum queue length is set to 1000
- $L_{OVER} = 800$, $L_{ALERT} = 700$, $L_{REP} = 600$, $L_{UNDER} = 300$
- Average service time of a client request = 10 ms
- Bandwidth: 45Mbps between ISP and 100Mbps within ISP

The network topology used in the simulation is depicted in Figure 3. A network consists of ISPs. In an ISP, there can be many routers, but we show only those routers with rent-a-servers (TSs) and assume that clients are connected to TSs directly. We also show only those paths connecting clients to the primary

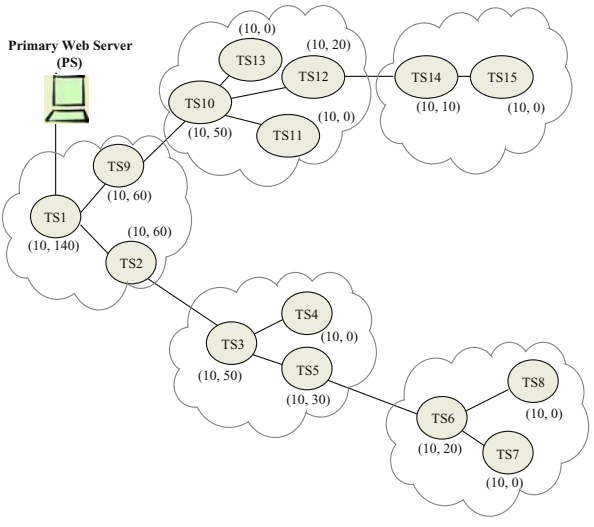
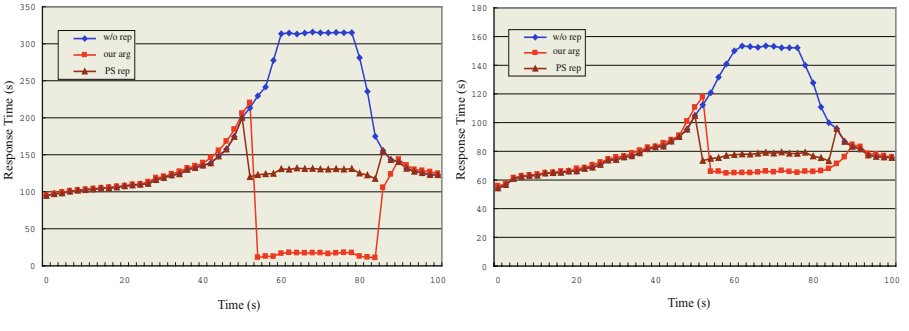


Fig. 3. Initial state of the network for simulation

web server (PS) because other paths are irrelevant in our simulation study. In the figure we define the load on PS and TSs with a pair of numbers. The first number is the load from local clients and the second is the load from child TSs.

Now we increase L_L at TS7 and TS8 to 300 and 400, respectively, resulting in overload at PS ($L(PS) = 820$). We study the performance of three cases. In the first case there is no web server replication. In the second case our algorithm is run to install RS at TS6 that has load over L_{REP} and is the closest to the client group causing the overload. In the third case (PS replication), dynamic replication is allowed only at PS. So PS is replicated when the load at PS increases over L_{OVER}. We measured the response time of requests from clients as in Figure 4. As soon as the overload is detected at PS, RS is installed at TS6



(a) Response time of client at TS6

(b) Response time of client at TS3

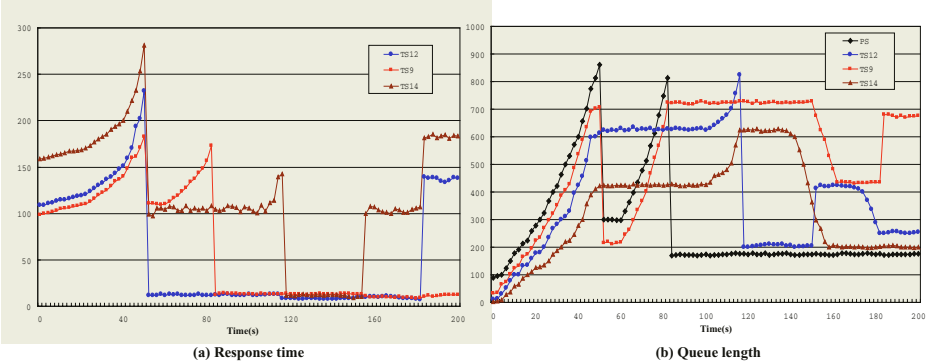
Fig. 4. Comparison with other approaches

in our algorithm and at PS in the PS replication approach. Figure 4 (a) shows the response time of the client which is located in the area causing the overload (a client connected to TS6). As soon as the overload is detected at PS, RS is installed and the client's response time is reduced in our approach and the PS replication approach. But our approach shows better response time because the client request is processed at RS which is closer than PS while the client request in the PS replication approach is serviced by replicated PSs and experiences longer request/result transmission delay. Figure 4 (b) shows the response time of the client which is located in the area without overload (a client connected to TS3). In this case we see that our approach and the PS replication approach exhibit similar performance because the requests are serviced at PS.

Now we increase L_L at TS11 ans TS13 to 100 in the initial state in Figure 3 and experiment our algorithm with a more complex scenario as follows.

- L_L on TS12, TS14, TS15 increases to 200, 100, 300, respectively. PS becomes overloaded with L(PS) = 900 and installs RS on TS12 with L(TS12) = 600.
- L_L on TS9, TS10, TS11, TS13 increases to 200, 100, 200, 200, respectively. PS gets overloaded again with L(PS) = 800 and installs RS on TS9 with L(TS9) = 700.
- L_L on TS14, TS15 increases to 300, 300, respectively. RS at TS12 gets overloaded with L(TS12) = 800 and installs RS on TS14 with L(TS14) = 600.
- L_L on TS14 and TS15 decreases and RS on TS14 is uninstalled.
- L_L on TS12 decreases and RS on TS12 is uninstalled.

Figure 5 (a) shows the response time of requests from clients on TSs on which RSs are installed and Figure 5 (b) shows the queue length at PS and TSs on which RSs are installed. We can see as client requests increase, overloaded servers install RSs on proper TSs at 52sec, 82sec, and 118sec and, therefore, response times and queue lengths are prevented from increasing too high. When RS is installed on TS14 at 118sec, the total load of RS on TS12 goes below L_{UNDER} and TS12 tries to uninstall its RS. But if RS on TS12 is uninstalled, the total load of RS on TS9 becomes 720 ($>L_{ALERT}$) and the uninstallation attempt is



(a) Response time (b) Queue length

Fig. 5. Simulation results for a more complex scenario

rejected. In the meantime L_L on TS14 and TS15 decreases and the total load on TS14 goes below L_{UNDER}. Then with the consent of RS on TS12 that is lightly loaded, TS14 uninstalls its RS and as a result TS12's load increases. As time passes, L_L on TS12 and TS14 decreases and at 180sec TS12 uninstalls its RS with the consent of RS on TS12. Now the remaining servers are PS and RS on TS9. We can see that the proposed algorithm installs and uninstalls replicated servers properly and helps maintain the response time and queue length at a proper level.

5 Conclusion

Traditional approaches such as web caching and static replication of web servers are not suitable solutions to the hot spot web sites that appear and disappear rapidly. In this paper we described our approach to automatically replicate web servers for hot spot web sites using rent-a-servers. We explained algorithms to install replicated web servers on proper rent-a-servers when web servers get overloaded and uninstall them when they become unnecessary. The proposed method not only distributes web requests among replicated web servers but also significantly reduces traffic on the Internet. We showed the effectiveness of the proposed method through simulation.

References

1. Rabinovich,M., Spatscheck,O.: Web Caching and Replication. Addison Wesley. (2002)
2. Vahdat,A., et al.: WebOS: Operating System Services for Wide Area Applications. UC Berkeley Technical Report UCB CSD-97-938. (1997)
3. Chankhunthod,A., et al.: A Hierarchical Internet Object Cache. USENIX. (1996)
4. Heddaya,A., Mirdad,S.: WebWave: Globally Load Balanced Fully Distributed Caching of Hot Published Documents. IEEE ICDCS. (1997)
5. Cardellini,V. Colajanni,M., Yu,P.S.: Dynamic Load Balancing on Web-Server Systems. IEEE Internet Computing. (1999)
6. Anderson,E., Patterson,D., Brewer,E.: The Magicrouter: An Application of Fast Packet Interposing. UC Berkeley Technical Report. (1996)
7. Cardellini,V., Colajanni,M., Yu,P.S.: Redirection Algorithms for Load-Sharing in Distributed Web-Server Systems. IEEE ICDCS. (1999)
8. Rabinovich,M., Aggarwal,A.: RaDaR: A Scalable Architecture for A Global Web Hosting Service. 8th WWW Conf. (1999)
9. Tennenhouse,D.L., Wetherall,D.J.: Towards an Active Network Architecture. DARPA Active Networks Conf. and Exposition. (2002)
10. The Nerwork Simulator: www.isi.edu/nsnam/ns.

Survey of Parallel and Distributed Volume Rendering: Revisited

Jiawan Zhang[1], Jizhou Sun[1], Zhou Jin[1], Yi Zhang [1], and Qi Zhai[2]

[1] IBM Computer Technology Center, Department of Computer Science, Tianjin University, Tianjin 300072, P.R.China
jwzhang@tju.edu.cn
http://ibm.tju.edu.cn
[2] School of Material Science and Engineering, Tianjin University, Tianjin 300072, P.R.China

Abstract. This paper covers a number of considerations about the subject of parallel volume rendering according to the rendering pipeline, including the choice of parallel architectures, the parallel volume rendering algorithms, the strategies for data distribution, and sorting and composition methods to achieve load balancing. Through the survey of recent parallel implementations, the general concerns and current research trend on the design of a parallel volume rendering system are discussed.

Keywords: Survey, Parallel Volume Rendering, Parallel Rendering System.

1 Introduction

Direct volume rendering (DVR) has gained great popularity in the process of mapping three-dimensional datasets to meaningful and visual information in a two-dimensional image plane without generating geometry models. It has been widely used in not only medical imaging, but also scientific trials such as geographical information analysis and unstructured aerodynamics calculations. The advantage of using direct volume rendering techniques is that both opaque and translucent structures can be visualized at the same time. Unfortunately, DVR is a memory and computationally intensive task.

Visualizing large-scale datasets remains one of the most challenging research problems. In order to achieve receivable performance, parallel volume rendering become a natural solution by increasing the number of the processing units to speed up the rendering task. The challenges of designing an efficient parallel visualization system consist of optimizing every stage in the overall rendering pipeline and balancing competing resources and numerous tradeoffs. Classifying the various schemes is important to characterize the behavior of each. And many researchers have given contributions to the development of parallel systems including Ma[MA99][MA95], Coelho and Nascimento [AC05], Mavriplis and Pirzadeh [MP99], Djurcilov [DJ98], Gonzato and Saec [GS00], Prakash and Kaufman [PK97], and Watson [WA02].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 435–444, 2005.

This paper discusses each stage in an overall parallel visualization system in detail. First, general architectures of the parallel system are reviewed. The parallel volume rendering algorithms including ray-casting, splatting and shear-warp are summarized. After that, the classification of the data distribution is expounded following with sorting strategy and composition methods in order to achieve the load balance problems. A number of parallel volume rendering systems are given as examples and conclusions of the current development and research trend in the field are discussed at last.

2 Parallel Volume Rendering Architectures

Parallel volume rendering application may run on parallel hardware or in a distributed computing environment. The architecture of the target system, including the memory organization and programming paradigm, has a major impact on the design of volume rendering system. Crockett [CT95]devised a more detailed survey.

The taxonomy of parallel computer architectures could be classified based on the number of instruction and data streams in the dependent system, such as Single Instruction Multiple Data and Multiple Instruction Multiple Data. Numerous volume renderers have been implemented on such platforms including Gribble [CG04], Leo[LS02], Schmid [SB00], and Meissner [MM01]. Another choice is shared memory architectures which could provide relatively efficient access to a global address space. Compared with shared memory architectures, distributed-memory systems offer improved architectural scalability, but often with higher costs for remote memory references. For this class of machines, managing communication is a primary consideration.

Currently, because of share memory system's isolation and lacking of expansibility, more and more researchers prefer to distributed memory systems, especially commercial PC-clusters. Many implements have been developed on this common platform like Garcia [AG02], Muraki [MS03], and Bajaj [BC02][BC00]. However, the interconnection network for communications between processors is observably slower than internal CPU and memory compared with supercomputers. So it is paramount to reduce the amount of bandwidth needed by global operations such as compositing and final assembly of the image in parallel volume rendering systems.

3 Parallel Volume Rendering Algorithms

Although new parallel algorithms could be developed, it is more convenient to design effective parallel rendering algorithms from existing sequential ones by introducing special methods or algorithms to deal with communication among processors, increased storage requirements, and other parallel issues. Many classic volume rendering algorithms such as ray casting, splatting, shear-warp, cell projection and so on, are chosen to paralled. Wittenbrink[WC98] proposed a detailed algorithm classification and description for parallel volume rendering.

Ray casting is probably the most popular approach for direct volume rendering that could produce the highest quality of rendered images. It is one of the most memory-intensive tasks for volume visualization. Ray casting has been parallelized on many parallel architectures using either image space partition or object space partition methods by Law and Yagel[LA96], Ma [MA94][MA93], Bajaj[BC00], Palmer and Totty[PM98] , and Ray[RH99].

Under the context of Splatting originated by Westover [LW89] [LW90] which is a widely used object-order volume rendering algorithm, there are much work has been done to realize parallelism. Elvins[TE92] was the first one to use splatting on the NCube machine which is a general purpose multiprocessor system. Johnson and Genetti [GJ95] implemented a similar splatting renderer on the Cray T3D. And Li[LP97] developed the parallel volume rendering system, ParVox.

Lacroute [LP94] proposed an accelerated shearing algorithm, called shear-warp factorization. A parallel implementation of the shear-warp algorithm is demonstrated on the SGI which challenged the capable of real-time frame rate[LP96].

The associated meshes of three-dimensional unstructured data from scientific trials such as aerodynamics calculations and Adaptive Mesh Refinement (AMR) are typically large in size and irregular in both shape and resolution. Ma and Crockett [MA97] proposed a parallel cell-projection method to solve this problem on IBM SP2 system. Weber [**?**] also presents a software algorithm for parallel volume rendering of AMR data using a cell-projection technique on several different parallel platforms.

As datasets continue to increase in size, new tools and techniques are developed to visualize such dataset in interactive ways. Methods utilizing hardware features to accelerate and optimize the slow software volume rendering algorithms are proposed by the research society. Cullip's method [CT94]defines the volume data as a 3D texture and utilizes the parallel texturing hardware in the Silicon Graphics RealityEngine workstations to perform reconstruction and resampling on polygons embedded in the texture. Cabral [CB94] proposed a method to use texture mapping hardware to accelerate volume rendering and tomographic reconstruction operation. Several of parallel texture based method for volume rendering have been presented including Magallon[MH01], Gribble [CG04] and Kniss[KJ01].

4 Task Subdivisions

A large amount of data should be distributed to the processors in the parallel environment, which would take a tremendous amount of time. Parallel renderers have to effectively manipulate and deliver those data to obtain the needed computational power for a variety of challenging applications. There are different approaches we can take to alleviate the data distribution problem by task subdivision. It can be generally divided into two categories based on how the workload is distributed among the processors.

One is the image-space partitioning scheme which subdivides the screen into small tiles and assigns each processor one or multiple tiles. Hsu [HW93]presented a segmented ray casting for data parallel volume rendering. In his method, the pixel values in the image plane are computed by casting rays through the volume data.

The other is the object-space partitioning scheme which subdivides the volume into small subvolumes and each processor is responsible for rendering one or multiple subvolumes. Ma [MA94]presented a divide-and-conquer algorithm for parallel volume rendering. His method is a pure object-space partitioning scheme, where each processor receives a partition of the volume.

Some researchers also use hybrid method to combines the benefits of both image-space and object-space partition schemes. Antonio Garcia[AG02] presented a parallel volume rendering algorithm based on the idea of pixel interleaving which can efficiently reduces communication overhead and volume interleaving which provides storage scalability.

5 Sorting Strategy

Sorting is a process to determine the visible-surface order to assure correct composition result in the rendering pipeline. It may happen during geometry preprocessing, between geometry preprocessing and rasterization, or during rasterization. Molnar [MS94]presented a parallel rendering classification method based on where sorting operation takes place. The three strategies are sort-first, sort-middle, and sort-last.

Sort-first approach divides the 2D screen into disjoint regions, and assigns each region to a different processor, which is responsible for all of the rendering in its region. This strategy has been used by Humphreys [HG01], Mueller [MC95], Samanta [SR00][SR99], Correa [WT02], and Bethel [BE03]. Sort-first could take advantage of frame-to-frame coherence well, since few primitives tend to move between tiles from one frame to the next. But it could cause load balancing problems between renderers because primitives may cluster into regions.

Poulton[FH89] and Montrym [MJ97] used sort-middle approaches to assign an arbitrary subset of primitives to each geometry processor, and a portion of the screen to each rasterizer. The geometry processor transforms and lights its primitives, and then sends them to the appropriate rasterizers.

The algorithms developed by Heirich and Moll[HA99], Molnar [MS92], Wei[WB98], Kirihata [YK04][YA04], and Moreland[MK01]adopted sort-last strategies to distribute pixels during rasterization. In sort-last method, the primitives with an arbitrary subset are assigned to each renderer which computes pixel values for its subset, no matter where they fall in the screen, and then transfer these pixels to compositing processors to obtain the final rendering results.

6 Composition Methods

Composition is the essential step of volume rendering which fuses the resamples generated in the local rendering step to obtain the final pixels. Parallel image compositing can be performed in many different ways.

Beside the simplest direct send approach used in Hsu [HW93] and Neumann [NU93], which has each processor send ray segments to the one responsible for compositing, Ma [MA94] introduced several other composition methods. Binary tree is a simple method which suites to hardware implementation. Muraki[MS01] designed and built special image compositing hardware to support the depth order composition for the translucent volume rendering, which implements the binary-tree subimage composition process. A straightforward binary compositing called binary swap is developed by Ma [MA94], which could keep all processors busy during the whole course of the compositing. Many sort-last methods including Sano [SK97], Wylie[WB01], and Yang[YD99] use this method. The projection method composites the final image by propagating the sub-images through processors in a front-to-back order. Silva [SC96] provided the implementation of this method. Ino [IF03]proposed another method called Divided-Screenwise Hierarchical method for sort-last parallel volume rendering on distributed memory multiprocessors. The DSH method divides the screen to tile-shaped regions and composites them in concurrent according to a dynamically determined hierarchical order.

7 Load Balance

Load balance is one of standards to evaluate parallel applications performance. Static load balancing is generally used by many researches in the field of parallel volume rendering to solve load imbalance in a distributed memory parallel computer because of the high overhead required to redistribute the data at runtime. A preprocessing step is often needed to partition the spatial domain based on not only the visualization parameters but also the characteristics of the grids and data for the unstructured dataset [MA95]. Ma and Crockett [MA97] use a round-robin cell distribution method to average out the differences in loads for different regions of the spatial domain to avoid preprocessing.

Dynamic load balancing techniques are essential solutions to load imbalance. Lacroute [LP95] presents a parallel volume rendering algorithm with dynamic load balancing strategy and finds that data redistribution and communication costs do not dominate rendering time on shared-memory architectures. At the same time, Amin [AM95]'s parallel algorithm based on the shear-warp algorithm and its implementation on a 128 processor TMC CM-5 distributed-memory parallel computer could renders a large medical dataset at an appropriate frames rate with the aid of dynamic load balancing.

8 Examples of Parallel Volume Rendering Systems

Many parallel volume rendering systems have been developed to obtain the needed computational power for a variety of challenging applications in medical imaging, engineering and scientific visualization. Li[LP97]'s ParVox is a parallel volume renderer using image-order as well as object-order splatting algorithm on Cray 3D. Recently, Gribble [CG03],and Kniss [PS04]developed their parallel

volume rendering applications based on Simian, which is originally proposed by Kniss [KJ04]. It provides a set of functions including transfer function configuration, clipping, shading and classification controls. Systems developed in special hardware, such as VIZARD II by Meißner [MM02] and Sepia by Moll[ML99],have also been developed.

9 Conclusions

Designing an efficient parallel visualization system remains one of the most challenging research problems in the field of scientific visualization. PC-Cluster, hardware-assisted rendering, real-time and large scale volume rendering is becoming the main trend of not only the parallel volume rendering but also the whole visualization society. More and more applications make use of data encoding techniques to decreasing the communication amounts. Load balance is an attention-getting problem in the parallel research at all times. Many challenges remain, and the discipline of parallel rendering is likely to be an active one for years to come.

Acknowledgements

This work is supported by Natural Science Foundation of China under grant No. 60373061.

References

[AC05] Alexandre Coelho and Marcio Nascimento: Cristiana Bentes. Parallel Volume Rendering for Ocean Visualization in a Cluster of PCs. BDBComp 2005

[AG02] Antonio Garcia, Han-Wei Shen: An Interleaved Parallel Volume Renderer With PC-clusters. Proceedings of the Fourth Eurographics Workshop on Parallel Graphics and Visualization. (2002)51–59

[AM95] Amin M., Grama A., and Singh V.: Fast volume rendering using an efficient, scalable parallel formulation of the shear-warp algorithm. Proceedings of the 1998 Parallel Rendering Symposium, (1995)7–14

[BC00] Bajaj C., Insung Ihm, and Sanghun Park: Compression-Based Ray Casting of Very Large Volume Data in Distributed Environments. The Proceedings of Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region. **2**(2000) 720–725

[BC02] Bajaj C., Park S., and Thane A. G.: Parallel Multi-PC Volume Rendering System. CS & ICES Technical Report. University of Texas at Austin. (2002)

[BE03] Bethel E.W., Humphreys G., Paul B., and Brederson J.D.: Sort-First, Distributed Memory Parallel Visualization and Rendering. IEEE Symposium on Parallel and Large-Data Visualization and Graphics. (2003)41–50

[CB94] Cabral B., Cam N., Foran J.: Accelerated volume rendering and tomographic reconstruction using texture mapping hardware. Proceedings of the 1994 symposium on Volume visualization.(1994) 91–98

[CG03] Christiaan Gribble, Xavier Cavin, Mark Hartner, and Charles Hansen: Cluster-Based Interactive Volume Rendering with Simian. Technical Report: UUCS-03-017.(2003)

[CG04] Christiaan Gribble, Steven Parker, and Charles Hansen: Interactive Volume Rendering of Large Datasets using the Silicon Graphics Onyx4 Visualization System. (2004)

[CT94] Cullip T. J., Neumann U. A.: Accelerating volume reconstruction with 3D texture mapping hardware. Technical Report: TR93-027. (1994)

[CT95] Crockett T. W.: Parallel Rendering. Technical Report: TR-95-31. Institute for Computer Applications in Science and Engineering.(1995)

[DJ98] Djurcilov S., Kim K., Lermusiaux P. F. J., and Pang A.: Visualizing scalar volumetric data with uncertainty. Computers and Graphics, **2**(26)(1998)239–248.

[FH89] Fuchs H., Poulton J., and Eyles J.: Pixel-planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-enhanced Memories. Computer Graphics. **23**(3)1989 79–88

[GJ95] Greg Johnson and Jon Genetti. Medical Diagnosis using the Cray T3D. Proceedings of 1995 Parallel Rendering Symposium.(1995) 70–77, 1995

[GS00] Gonzato J. C. and Saec B. L.: On modeling and rendering ocean scenes. Journal of Visualisation and Computer Simulation **11**(1) (2000) 27–37

[HA99] Heirich A. and Moll L.: Scalable distributed visualization using off-the-shelf components. Symposium on Parallel Visualization and Graphics. (1999) 55–59

[HG01] Humphreys G., Eldridge M., and Buck I.: WireGL: A scalable graphics system for clusters. Proceedings of SIGGRAPH 2001. (2001) 129–140

[HW93] Hsu W. M.: Segmented Ray Casting for Data Parallel Volume Rendering. Proceedings of 1993 Parallel Rendering Symposium. (1993)7–14

[IF03] Ino F., Sasaki T., Takeuchi A., and Hagihara K.: A Divided-Screenwise Hierarchical Compositing for Sort-Last Parallel Volume Rendering. Proceedings of 2003 Parallel and Distributed Processing Symposium. (2003)

[KJ01] Kniss J., McCormick P., McPherson A.: Interactive Texture-Based Volume Rendering for Large Data Sets. IEEE Computer Graphics and Applications.**21**(4) (2001)52–61

[KJ04] Kniss J., Jürgen P. Schulze, and Uwe Wössner: Medical Applications of Multi-field Volume Rendering and VR Techniques. Proceedings of the Joint Eurographics-IEEE TCVG Symposium on Visualization. (2004) 249–254

[LA96] Law A., and Yagel R.: Multi-frame thrashless ray casting with advancing ray-front. Proceedings of Graphics Interface '96. (1996) 70–77

[LP94] Lacroute P. and Levoy M.: Fast Volume Rendering Using a Shear-Warp Factorization of the Viewing Transformation. Proceedings of SIGGRAPH '94.

[LP95] Lacroute P.: Real-time volume rendering on shared memory multiprocessors using the shear warp factorization. Proceedings of the 1998 Parallel Rendering Symposium. (1995)15–22

[LP96] Lacroute P.: Analysis of a Parallel Volume Rendering System Based on the Shear-Warp Factorization. IEEE Transactions on Visualization and Computer Graphics.**2**(3)(1996) 218–231

[LP97] Li P. P., Whitman S., and Mendoza R.: ParVox-A Parallel Splatting Volume Rendering System for Distributed Visualization. Proceedings of IEEE Symposium on Parallel Rendering. (1997). 7– 14

[LS02] Leo C. S., and Schroder H.: Fast Processing Of Medical Images Using A New Parallel Architecture.Proceedings of Fifth IEEE Southwest Symposium on The Hybrid System Image Analysis and Interpretation(2002)148–152

[LW89] Lee Westover: Interactive Volume Rendering. Proceedings of Chapel Hill Workshop on Volume Visualization Workshop.(1989)9–16

[LW90] Lee Westover: Footprint Evaluation for Volume Rendering. Computer Graphics. **24**(4) (1990) 367–376

[MA93] Ma K. L., Painter J. S., and Hansen, C. D.: A Data Distributed, Parallel Algorithm for Ray-Traced Volume Rendering. Parallel Rendering Symposium. 1993 15–22

[MA94] Ma K. L., Painter J., Hansen C., and Krogh M.: Parallel volume rendering using binary-swap compositing. IEEE Computer Graphics and Applications. **14**(4)(1994) 59–68

[MA94] Ma K. L., Painter J. S., Hansen C. D., and Krogh M. F.: Parallel Volume Rendering Using Binary-Swap Compositing. IEEE Computer Graphics and Applications.**14**(4)(1994) 59–68

[MA95] Ma K. L.: Parallel volume ray-casting for unstructured-grid data on distributed-memory architectures. ACM SIGGRAPH Proceedings of the 1995 Parallel Rendering Symposium. 1995 23–30

[MA97] Ma K. L. and Crockett T. W.: A scalable parallel cell-projection volume rendering algorithm for three-dimensional unstructured data. ACM SIGGRAPH Proceedings of the 1997 Symposium on Parallel Rendering. 1997 95–104

[MA99] Ma K. L. and Crockett T. W.: Parallel Visualization of Large-scale Aerodynamics Calculations: A Case Study on the Cray T3E. Proceedings of IEEE Parallel Visualization and Graphics Symposium. 1999 15–20

[MC95] Mueller C.: The sort-first rendering architecture for high-performance graphics. 1995 Symposium on Interactive 3D Graphics. (1995) 75–84

[MH01] Magallon M., Hopf M. and Ertl T. : Parallel volume rendering using PC graphics hardware. Proceedings of Ninth Pacific Conference on Computer Graphics and Applications. (2001)384–389

[MJ97] Montrym J. S., Baum D. R., Dignam D. L. and Migdal C. J. : InfiniteReality: A Real-time Graphics System. Proceedings of SIGGRAPH 97. (1997) 293–302

[MK01] Moreland K., Wylie B., Pavlakos C.: Sort-Last Parallel Rendering for Viewing Extremely Large Data Sets on Tile Displays. Proceedings of IEEE 2001 Symposium on Parallel and Large-Data Visualization and Graphics. (2001) 85–154

[MP99] Mavriplis D. J. and Pirzadeh S.: Large-scale parallel unstructured mesh computations for 3D high-lift analysis. AIAA Journal ofilircraji. 1999

[ML99] Moll L., Heirich A., Shand M. Sepia: Scalable 3D Compositing Using PCI Pamette. Proceedings of Seventh Annual IEEE Symposium on Field-Programmable Custom Computing Machines.(1999)146–155

[MM02] Meissner M., Kanus U., Wetekam G. , Hirche J., Ehlert A. , Strasser W., Doggett M., Forthmann P., and Proksa R.: VIZARD II: A Reconfigurable Interactive Volume Rendering System. Eurographics Workshop on Graphics Hardware (2002)137–146.

[MM01] Meissner M., Grimm S., Strasser W.: Parallel volume rendering on a single-chip SIMD architecture. Proceedings of IEEE Symposium on Parallel and Large-Data Visualization and Graphics, (2001)107–157

[MS92] Molnar S., Eyles J. and Poulton J.: Pixelflow: Highspeed rendering using image composition. Computer Graphics. **26** (2) (1992) 231–240

[MS94] Molnar S., Cox M., and Ellsworth D.: A sorting classification of parallel rendering. IEEE Computer Graphics and Applications, **14**(4)(1994) 23–32

[MS01] Muraki S., Ogata M., Ma K. L., Koshizuka K.: Next-Generation Visual Supercomputing using PC Clusters with Volume Graphics Hardware Devices. Proceedings of High Performance Networking and Computing Conference. (2001)

[MS03] Muraki S., Lum E.B., and Ma K. L. A PC Cluster System for Simultaneous Interactive Volumetric Modeling and Visualization. IEEE Symposium on Parallel and Large-Data Visualization and Graphics. (2003)95–102

[NU93] Neumann U.: Parallel Volume-Rendering Algorithm Performance on Mesh-Connected Multicomputers. Proceeding of 1993 Parallel Rendering Symp. (1993) 97-104

[PK97] Prakash C. E. and Kaufman A. E. : Volume terrain modeling. Technical report, SUNYSB Technical Report, 1997.

[PM98] Palmer M. E., Totty B., and Taylor S.: Ray casting on shared-memory architectures: memory-hierarchy considerations in volume rendering. Concurrency, IEEE, **6**(1)(1998)20–35

[PS04] Patrick S. McCormick, Jeff Inman, and Charles Hansen: Scout: A Hardware-Accelerated System for Quantitatively Driven Visualization and Analysis. Proceedings IEEE Visualization 2004, (2004) 171–178

[RH99] Ray H., Pfister H., and Silver D.: Ray Casting Architectures for Volume Visualization. IEEE Transactions on Visualization and Computer Graphics. **5**(3)(1999) 210–223

[SB00] Schmidt B.: Design of a Parallel Accelerator for Volume Rendering. Euro-Par'00, Munich, Germany, Springer LNCS **1900**(2000) 1095–1104

[SC96] Silva C. T., Kaufman A. E., and Pavlakos C.: PVR: High- Performance Volume Rendering. 1996 IEEE Computational Science and Engineering Winter.**3**(4) (1996) 18–28

[SK97] Sano K., Kitajima H., Kobayashi H., and Nakamura T.: Parallel Processing of the Shear-Warp Factorization with the Binary-Swap Method on a Distributed-Memory Multiprocessor System. Proceedings of 1997 Parallel Rendering Symposium. (1997) 87–94

[SR99] Samanta R., Zheng J., Funkhouser T., Li K. and Singh J. P.: Load balancing for multi-projector rendering systems. 1999 SIGGRAPH/Eurographics Workshop on Graphics Hardware. (1999) 107–116

[SR00] Samanta R., Funkhouser T., and LiK.: Sort- first parallel rendering with a cluster of PCs. ACN SIGGRAPH 2000.(2000)

[TE92] Todd Elvins: Volume Rendering on a Distributed Memory Parallel Computer. Proceedings of Visualization '92.(1992)

[WA02] Watson, Andrew I., Lerico T. P., Fournier, J. D., and Szoke, E. J.: The use of d3d when examining tropical cyclones. Interactive Symposium on AWIPS.(2002) 131–135

[WB98] Wei B., Clark D. W., Felten E. W., Li K., and Stoll G.: Performance issues of a distributed frame buffer on a multicomputer. 1998 SIGGRAPH / Eurographics Workshop on Graphics Hardware. (1998) 87–96

[WB01] Wylie B., Pavlakos C., Lewis V., and Moreland K.: Scalable Rendering on PC Clusters. IEEE Computer Graphics and Applications, **21**(4)(2001) 62–70

[WC98] Wittenbrink, and Craig M.: Survey of Parallel Volume Rendering Algorithms. HP Labs Technical Reports. HPL-98-49R1. (1998)

[WG03] Weber G. H., Ohler M., and Kreylos O.: Parallel Cell Projection Rendering of Adaptive Mesh Refinement Data. Proceedings of IEEE Symposium on Parallel and Large-Data Visualization and Graphics 2003.(2003) 51–60

[WT02] Wagner T. C., James T. K., and Claudio T. S.: Out-Of-Core Sort-First Parallel Rendering for Cluster-Based Tiled Displays. 2002 Fourth Eurographics Workshop on Parallel Graphics and Visualization. (2002)

[YA04] Yasuhiro Kirihata, Jason Leigh, Chaoyue Xiong, and Tadao Murata. A Sort-Last Rendering System over an Optical Backplane. International Conference on Cybernetics and Information Technologies, Systems and Applications: CITSA 2004

[YK04] Yasuhiro Kirihata: An Implementation of Sort-Last Volume Rendering over an Optical Network Backplane. Electronic Visualization Laboratory Technical Report. MS Project. University of Illinois.

[YD99] Yang D. L., Yu J. C., Chung Y. C.: Efficient Compositing Methods for the Sort-Last-Sparse Parallel Volume Rendering System on Distributed Memory Multicomputers. Proceedings of 1999 International Conference on Parallel Processing.(1999) 200–207

Scheduling Pipelined Multiprocessor Tasks: An Experimental Study with Vision Architecture

M. Fikret Ercan

School of Electrical and Electronic Engineering,
Singapore Polytechnic, 500 Dover Rd., S139651, Singapore
mfercan@sp.edu.sg

Abstract. This paper presents the application scheduling algorithms on a class of multiprocessor architectures that exploit temporal and spatial parallelism simultaneously. The hardware platform is a multi-level or partitionable architecture. Spatial parallelism is exploited with MIMD type processor clusters (or layers) and temporal parallelism is exploited by pipelining operations on those independent clusters. In order to fully exploit system's capacity, multi processor tasks (MPTs) that are executed on such system should be scheduled appropriately. In our earlier study, we have proposed scheduling algorithms based on well known local search heuristic algorithms such as simulated annealing, tabu search and genetic algorithm and their performances were tested computationally by using a set of randomly generated test data. In this paper, we present application of these scheduling algorithms on a multilayer architecture which is designed as a visual perception unit of an autonomous robot and evaluate performance improvement achieved.

1 Introduction

In many engineering applications dedicated parallel architectures are employed to overcome heavy computational demand. A particular class of these architectures exploits parallelism in space (data or control) and time (temporal) simultaneously and they are mainly utilized in applications such as machine vision, robotics, and power system simulation [6, 12]. There are many architectures of this type reported in literature either by employing a pool of processors that can be partitioned into independent clusters or by organizing processors in multiple hierarchically organized layers (see for instance [4,14]). Processor clusters typically communicate via a shared bus or via direct (or configurable) processor to processor communication links. A typical example to this type of computing structure is real-time computer vision, where overall structure is made of a stream of related tasks. Operations performed on each image frame can be categorized as low, intermediate and high level. The result of an algorithm at the low level initiates another algorithm at the intermediate level and so on. By exploiting available spatial parallelism, algorithms at each level can be split into smaller grains to reduce their computation time. In addition, computing performance can be improved further by exploiting temporal parallelism when the continuous image frames are being processed by the system. That is, algorithms at each level can be mapped into a processing cluster and executed simultaneously to create a pipelining effect.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 445 – 452, 2005.

An important issue that affects the performance of parallel computing is proper scheduling of tasks to system processors. A significant amount of study deals with partitioning and mapping of tasks considering network topology, processor, link, memory parameters and processor load balance whereas another group of study deals with sequencing and scheduling multiprocessor tasks (MPT) so that the completion time of all the tasks can be minimized [1, 2, 3, 5]. Task partitioning problem is well studied in literature whereas MPT scheduling studies are generally targeted for architectures made of a single layer. However, a special scheduling problem arises for the class of architecture considered in this paper. Here, we have to consider scheduling of MPTs not only for one layer but for all the layers of the systems. Computations performed on these systems are made of multiple interdependent MPTs and they will be defined as jobs. Hence, the job-scheduling problem is basically finding a sequence of jobs that can be processed on the system in the shortest possible time.

In our earlier studies, we tackled with the above defined problem and developed several scheduling algorithms based on well known local search algorithms, such as simulated annealing, tabu search, and genetic algorithms [7, 8, 9]. In this paper, we present application of these scheduling techniques on the actual system. The application platform is a multiprogrammable multiprocessor system designed as a visual perception unit of an autonomous robot. In the following sections, we will present the system architecture and the structure of the computations performed on it. Then, we will briefly discuss scheduling algorithms employed and present the result of our empirical studies.

2 Multiprocessor System and the Structure of the Operations

A multilayer architecture is basically made of m number of processing layers, each of which holds k_i ($i = 1,2...,m$) number of processors. The number of layers, processor type and the parallelism model employed are all depend on the application.

2.1 Multiprocessor System Used for Experiments

We employed a parallel architecture, developed for machine vision and autonomous navigation purpose, as our application base. The system was built with latest parallel DSPs employing a combination of distributed and shared memory MIMD parallelism. In order to achieve multiprogramming, processor arrays are organized in a hierarchical multi-layered manner. Image is acquired by a frame grabber and fed into a pair of DSPs which then scatter image frame to processors simultaneously. This way four processors, performing low-level image processing algorithms, receive image data in very short time. Two inter-processor connection schemes are employed for the communication. A reconfigurable network provides one-to-one connection of remote processors. Meanwhile, communication links of DSPs are used to communicate with neighboring processors directly. Furthermore, a shared memory block is used to exchange data between layers. The basic computing element used in this system is the high performance parallel DSPs. These processors are capable of performing various forms of computations

such as number crunching to symbolic type of operations efficiently. Analogue Devices' SHARC DSPs (TM-001 at 250MHz) were used for the system [13]. Each processor holds 6 Mbytes of local memory and 128Mbytes memory is shared between the four DSP. Figure 1 illustrates the vision architecture used as test bed.

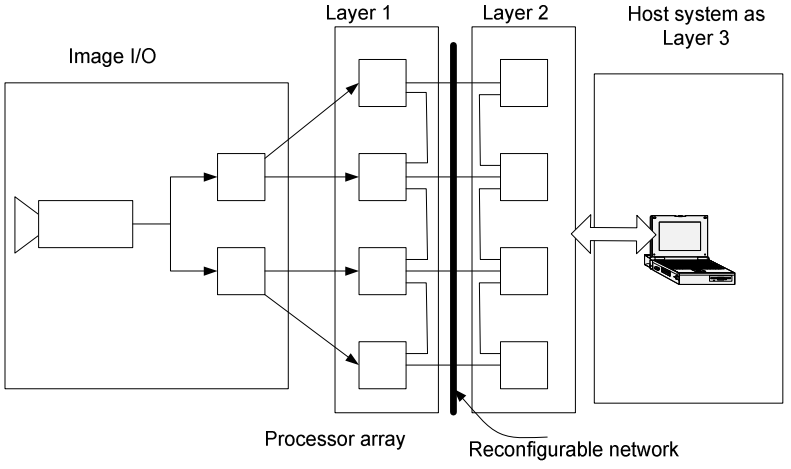


Fig. 1. A block diagram of the system

Data transfer cost between the layers is negligible compared to the processing times of the tasks. When a task at one layer is completed, its successor at the subsequent layer will be initiated without delay, provided that there is sufficient amount of vacant processors. Otherwise, tasks has to be placed in a queue (or stored in a buffer) until some processors become available. For example, consider the case where two tasks are available to be executed at one layer. If each task requires large number of processors, apparently both tasks can not be executed on that layer simultaneously. One task will be executed while the other one is waiting for its turn. An important criterion is therefore to decide on task sequence so that processor idle time can be minimized.

2.2 Modeling of the Operations Performed on the System

In this system, a knowledge based image interpretation algorithm is performed and continuous image frames are processed by a series of multiprocessor tasks. The number of hypothesized object models determines the number of jobs to be executed. It can easily be seen from the Figure 1 that the first and second layers of the system are made of multiprocessors and the third layer has single processor. In Figure 2 tasks performed on the system is expressed graphically. Here, due to distributed nature of the image interpretation algorithm developed, there is no precedence relationship among the jobs. Data dependency occurs for the final task at the third layer. Hence,

the problem can be defined as scheduling pipelined multiprocessor tasks of two layers. Apparently, the efficient scheduling of layer 1 and layer 2 tasks will consequently benefit layer 3 tasks. That is, by minimizing completion times of tasks at the first and the second layers, third layer tasks can start earlier and have a larger time slot to perform model matching algorithms. The number of processors required by each task depends on the available parallelism that algorithms can exploit. Using scheduling terminology each pipelined MPT will be named as job. The problem is formally defined as follows: There is a set J of n independent and simultaneously available jobs to be processed in a computing platform with two multiprocessor layers, where layer j has m_j identical parallel processors, $j = 1,2$. The level of pipeline in each job is the same and compatible with the number of processing layers available in the computing platform. Each job $J_i \in J$ has two multiprocessor tasks (T_{ij}), namely T_{i1} and T_{i2}. T_{ij} should be processed on P_{ij} number of processors simultaneously in layer for a period of t_{ij} without interruption $(i = 1,2,...,n$ and $j = 1,2)$. Hence, each T_{ij} is characterized by its processing time, t_{ij}, and its processor requirement, P_{ij} $(i = 1,2,...,n$ and $j = 1,2)$. All the processors are continuously available from time 0 onwards, and each processor can handle no more than one MPT at a time. Multiprocessor tasks are performed without pre-emption. The objective is to find an optimal schedule for the jobs so as to minimize the maximum completion time among all jobs.

3 Scheduling Algorithms

In order to solve the scheduling problem presented above, we have developed algorithms based on three well known local search heuristics. In our earlier study, tabu search and simulated annealing approaches were considered and their performance was evaluated by means of intensive computational experiments using randomly generated problem instances [7, 8]. For each combination of processing time ratio and processor configuration of the architecture, 25 problems were generated. In another study, we have developed a genetic algorithm for this problem [9]. Genetic algorithms have been shown to be effective solving scheduling problems in literature. We have evaluated the performance of algorithm with the same data set and a comparison of performances is presented at the following Table 1 and Table 2.

All the algorithms implemented using C++ and run on a PC with a 350 MHz Pentium II processor. Results are presented in terms of Average Percentage Deviation (APD) of solution from the lower bound. This deviation is defined as $((C_{\max}(HE) - LB)/LB) \times 100$, where $C_{\max}(HE)$ denotes schedule length obtained by heuristic algorithms and LB denotes the lower bound for the schedule. The details of the lower bounds developed for this problem are presented in [11].

From the Tables 1 and 2, it can be observed that our genetic algorithm outperformed our simulated annealing and tabu search based algorithms.

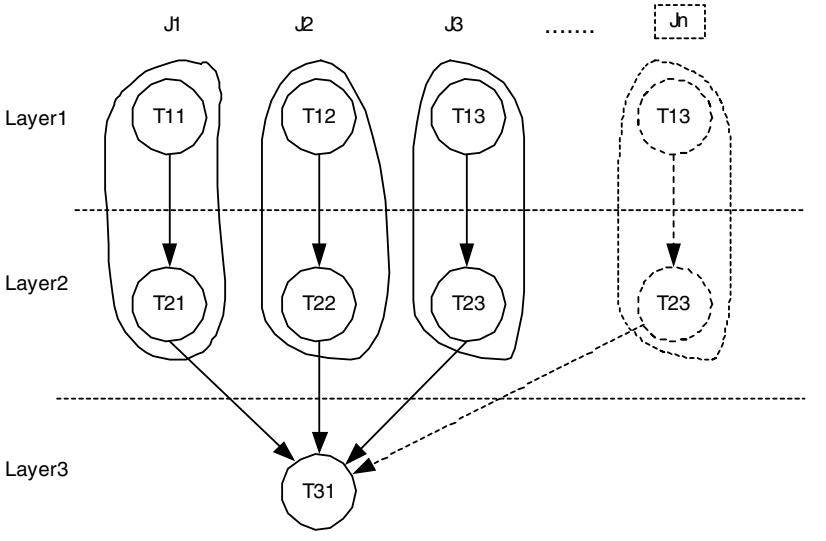


Fig. 2. Data dependency graph between tasks in the perception unit

Table 1. A comparison of TS, SA and GA algorithms for machine configuration 2:1

jobs	# of Processors: Layer1=2, Layer2=1			# of Processors: Layer1=4, Layer2=2			# of Processors: Layer1=8, Layer2=4		
	SA	TS	GA	SA	TS	GA	SA	TS	GA
10	0.94	1.14	0.91	4.21	5.06	5.31	10.15	10.72	10.34
30	0.52	0.47	0.1	2.64	3.41	0.247	4.65	5.82	1.44
50	0.42	0.44	0	1.71	2.3	0.222	4.53	5.59	1.78

Table 2. A comparison of TS, SA and GA algorithms for machine configuration 1:1

jobs	# of Processors Layer1=2, Layer2=2			# of Processors Layer1=4, Layer2=4			# of Processors Layer1=8, Layer2=8		
	SA	TS	GA	SA	TS	GA	SA	TS	GA
10	2.62	3.28	3.05	7.07	8.49	7.6	8.88	9.21	11.51
30	2.49	3.25	0.288	3.5	4.69	0.88	7.81	9.32	3.37
50	2.69	3.06	0.19	4.2	4.79	1.17	7	7.96	2.93

4 Experimental Study

As mentioned, our earlier study was based on computational experiments. In order to make a reasonable comparison of the search methods described above, other than the randomly generated problem sets, we considered a set of test instances obtained

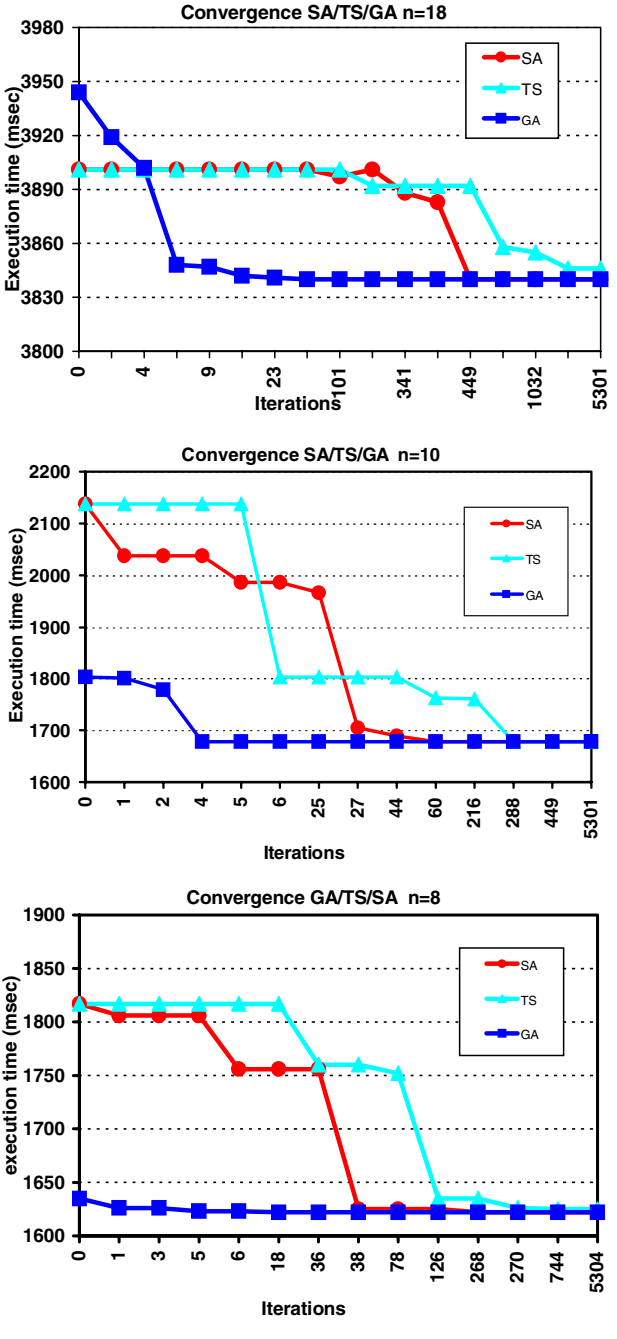


Fig. 3. Convergence of algorithms for four test cases

from the machine vision system described above. The analysis of the set has practical relevance and complements the results of the randomly generated problem sets. The particular configuration of the machine vision system has four processors at each stage. We obtained the processing times and the processor requirements for jobs from the machine vision system where each job comprises two tasks. The task at the first level is an object detection algorithm that particularly searches evidence in the image for a predefined object. The task at the second layer elaborates on the results and groups object-related data and creates symbolic data about the object. The processing technique described here employs top-down control mechanism instead of traditional bottom up processing methods. Therefore, the number of jobs to be implemented in this system is related to the variety of objects to be searched in a given image. The processor requirements for all the MPTs were deterministic, though some of the task execution times were content dependent; however they did not vary greatly. Hence, for those tasks we used their average execution times in the scheduling heuristic. The data for each test instance is provided in [10].

The following Figure 3 depicts the maximum completion times obtained in each search iteration (or generation in GA) for three search heuristics. Among the four test cases, the best performance improvement was achieved for 10 jobs. The execution time was shortened by around 22% and the number of frames processed within a 10-second period by the system reached to 11.5 frames, which were around 9 frames. However, for 18 jobs, improvement was no better than 1.5%. An important conclusion that can be drawn from the practical experiments is that our GA converges very quickly compared to our TS and SA algorithms.

5 Summary

In this paper, the job-scheduling problem on a multi-tasking multiprocessor environment is considered. A job is made up of interrelated multiprocessor tasks where a multiprocessor task is modeled with its processing requirement and processing time. Three well-known local search heuristics, Simulated Annealing, Tabu Search and Genetic Algorithms have been applied for the solution of this problem. In this paper, we evaluate the performance of the algorithms mainly based on their capacity to minimize maximum completion time of all the jobs. We applied these scheduling algorithms to multilayer system and observed that performance of the system significantly improved. However, due to their large computation times, SA, TS, or GA can be used in deterministic cases where scheduling is done before executing the programs on the system.

References

1. Błażewicz J. and Drozdowski M.: Scheduling Divisible Jobs on Hypercubes. Parallel Computing, Vol. 21. (1995) 1945-1956
2. Błażewicz J., Ecker K. H., Pesch E., Schmidt G. and Węglarz J.: Scheduling Computer and Manufacturing Processes. Springer-Verlag, Berlin (1996)
3. Bokhari S. H.: Assignment Problems in Parallel and Distributed Computing. Kluwer Academic, Boston (1987)

4. Choudhary A. N., Patel J. H and Ahuja N.: NETRA: A Hierarchical and Partitionable Architecture for Computer Vision Systems. IEEE Transactions on Parallel and Distributed Systems, Vol. 4. (1993) 1092-1104
5. El-Revini H.: Partitioning and Scheduling. In: Zomaya A. D. (ed): Parallel and Distributed Computing Handbook. McGraw-Hill, New York (1996) 239-273
6. Ercan M. F. and. Fung Y. F: Real-time Image Interpretation on a Multi-layer Architecture. IEEE TENCON'99, Vol. 2. (1999) 303-1306
7. Ercan M.F., Oğuz C., Fung Y.F.: Performance Evaluation of Heuristics for Scheduling Pipelined Multiprocessor Tasks. Lecture Notes in Computer Science. Vol. 2073. (2001) 61-70
8. Ercan M. F. and Fung Y. F.: Tabu Search and Simulated Annealing on the Scheduling of Pipelined Multiprocessor Tasks. Proceedings of ICSEC'02 (2002) 785-788
9. Oğuz C., Fung Y.F., Ercan M. F. and Qi X.T.: Parallel Genetic Algorithm for a Flow-Shop Problem with Multiprocessor Tasks. Lecture Notes in Computer Science, Vol. 2667. (2003) 987-997
10. Oğuz C.: http://www.acad.polyu.edu.hk/~msceyda/Research/data_set_EJOR.htm, 2002.
11. Oğuz C., Ercan M.F., Cheng T.C.E. and Fung Y.F.: Multiprocessor Task Scheduling in Multi Layer Computer Systems. European Journal of Operations Research, Vol. 149. (2003) 390-403
12. Scala M. L., Bose A., Tylavsky J. and Chai J. S.: A Highly Parallel Method for Transient Stability Analysis. IEEE Transactions on Power Systems, Vol. 5. (1990)1439-1446
13. TigerSharc DSP Hardware Specifications, Analog Devices (2002)
14. Weems C. C., Riseman E. M. and Hanson A. R.: Image Understanding Architecture: Exploiting Potential Parallelism in Machine Vision. IEEE Computer, Vol. 25. (1992) 65-68

Universal Properties Verification of Parameterized Parallel Systems

Cecilia E. Nugraheni

Dept. of Computer Science, Fac. of Mathematics and Natural Sciences,
Parahyangan Catholic University, Bandung, Indonesia
cheni@home.unpar.ac.id

Abstract. This paper presents a method for verifying universal properties of parameterized parallel systems using Parameterized Predicate Diagrams [10]. Parameterized Predicate Diagrams are diagrams which are used to represent the abstractions of such systems described by specifications written in temporal logic. This method presented here integrates deductive verification and algorithmic techniques. Non-temporal proof obligations establish the correspondence between the original specification and the diagram, whereas model checking can be used to verify properties over finite-state abstractions.

Keywords: parameterized parallel system, verification, formal method, diagram, temporal logics.

1 Introduction

Parameterized parallel systems, or parameterized systems for short, have become a very important subject of research in the area of computer-aided verification. A parameterized system consists of several similar processes whose number is determined by an input parameter. Many interesting systems are of this form, for example, mutual exclusion algorithms for an arbitrary number of processes wanting to use a common resource. Many distributed systems, in particular those that control communication and synchronization of networks, also have as their body a parallel composition of many identical processes. A challenging problem is to provide methods for the uniform verification of such systems, i.e. prove by a *single* proof that a system is correct for any value of the parameter. The key to such a uniform treatment is parameterization, i.e. presenting a single syntactic object that actually represents a family of objects.

There are basically two approaches to formal verification, which are the deductive approach and the algorithmic approach. The deductive approach is based on *verification rules*, which reduce the system validity of a temporal property to the general validity of a set of first-order *verification conditions*. The most popular algorithmic verification method is *model checking*. Although this method is fully automatic for finite-state systems, it suffers from the so-called *state-explosion* problem. The size of the state space is typically exponential in the

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 453–462, 2005.

number of components, and therefore the class of systems that can be handled by this method is limited.

The ability to conduct a uniform verification of a parameterized system is one of the striking advantages of the deductive method for temporal verification over algorithmic techniques such as model-checking techniques. Let n denote an arbitrary but finite number of identical processes. Usually, the model checker's memory capacity is exceeded for values of n smaller than 100 [8]. Furthermore, in the general case, nothing can be concluded about the property holding for any value of n from the fact that it holds for some finite set of values. In comparison, the deductive method establishes in one fell swoop the validity of the property for any value of n [8, 5].

The need for a more intuitive approach to verification leads to the use of *diagram-based formalisms*. Usually, these diagrams are graphs whose vertices are labelled with first-order formulas, representing sets of system states, and whose edges represent possible system transitions. This approach combines some of the advantages of deductive and algorithmic verification: the process is goal-directed, incremental and can handle infinite-state systems.

In the context of parameterized systems, there are two classes of properties may be considered, namely the properties related to the whole processes and the ones related to a single process in the systems. The latter class is sometimes called the *universal* property. For example, given a parameterized system which consists of n processes and some property P, the universal properties are expressed as formulas of the form $\forall\, k \in 1..n \,:\, P(k)$.

This paper proposes a verification technique for parameterized systems using a class of diagrams called Parameterized Predicate Diagrams (PPD) [10]. PPDs is an extension of (the original) predicate diagrams from Cansell et.al. [4]. Predicate diagrams can be used to prove the properties related to the whole processes. Unfortunately, this approach suffers from the limitation that it cannot be used for the verification of universal properties. We use TLA* [9] to formalize our approach and use TLA$^+$ style [7] for writing specifications.

This paper is structured as follows. We begin with the specification of parameterized system in TLA*. Section 3 describes the definition and the use of PPDs in the verification of parameterized systems. As illustration we take the Ticket protocols [2] as case study. This is explained in Section 5. Conclusion and future work will be given in Section 6.

2 Specification of Parameterized Systems

In this work, we restrict on the parameterized systems which are *interleaving* and consist of finitely, but arbitrarily, *discrete* components.

Let M denotes a finite and non-empty set of processes running in the system being considered. A parameterized system can be describe as a formula of the form:

$$parSpec \equiv Init \wedge \Box[\exists k \in M : Next(k)]_v \wedge \forall k \in M : L(k) \qquad (1)$$

where

- *Init* is a state predicate that describes the global initial condition,
- $Next(k)$ is an action that characterizes the next-state relation of a process k,
- v is a state function representing the variables of the system and
- $L(k)$ is a formula stating the liveness conditions expected from the process k.

Formulas such as $Next(k)$ and $L(k)$ are called parameterized actions.

3 Parameterized Predicate Diagrams

In proving universal properties we have to find a way that enables us to keep track the behaviors of some particular process. The idea is to view the systems as collections of two components, which are a particular process and the collection of the rest of the processes.

Given a specification of parameterized system, *parSpec*, and a property, P, our goal is to prove the validity of $parSpec \rightarrow \forall\, k \in M \,:\, P(k)$. Let $i \in M$ be some process. We reduce the proof to the proof of $parSpec \wedge i \in M \rightarrow P(i)$[1]. If the proof succeeds then, since we apply the standard quantifier introduction rule of first-order logic, we can conclude that the property holds over each process in the system, i.e. $\forall\, k \in M \,:\, P(k)$ is valid.

Now we present a class of diagrams that can be used for the verification of parameterized systems. The underlying assertion language, by assumption, contains a finite set $\mathcal{O}$ of binary relation symbols $\prec$ that are interpreted by well-founded orderings. For $\prec \in \mathcal{O}$, its reflexive closure is denoted by $\preceq$. We write $\mathcal{O}^{=}$ to denote the set of relation symbols $\prec$ and $\preceq$ for $\prec$ in $\mathcal{O}$.

For the sake of the presentation, in the following we will denote by $A(i)$ for some action of process i and $A(\mathbf{k})$ for formula $\exists\, k \in M \setminus \{i\} \,:\, A(k)$ for some action of any process other than i.

Definition 1. *(quantified-actions) For a set of parameterized actions $\mathcal{A}$, we denote by $\Phi(\mathcal{A})$, the set of quantified-actions which are formulas of the form $A(i)$ or $A(\mathbf{k})$ for $A(k)$ some parameterized action in $\mathcal{A}$.*

Formally, the definition of parameterized predicate diagram is relative to finite sets $\mathcal{P}$, $\mathcal{A}$ and $\Phi(\mathcal{A})$ that contain the state predicates, the (names of) parameterized actions of interest and quantified-actions over parameterized actions in $\mathcal{A}$, respectively. We will later use $\tau \notin \mathcal{A}$ to denote a special stuttering action. We write $\overline{\mathcal{P}}$ to denote the set of literals formed by the predicates in $\mathcal{P}$, that is, the union of $\mathcal{P}$ and the negations of the predicates in $\mathcal{P}$.

Definition 2. *(PPD) Given a set of state predicates $\mathcal{P}$, a set of parameterized actions $\mathcal{A}$ and the set of quantified-actions over parameterized actions in $\mathcal{A}$, $\Phi(\mathcal{A})$,* PPD *over $\mathcal{P}, \mathcal{A}$, and $\Phi(\mathcal{A})$, $\mathcal{G}$, consists of*

[1] We call such method *Skolemization*.

- *a finite set* $N \subseteq 2^{\overline{\mathcal{P}}}$ *of nodes,*
- *a finite set* $I \subseteq N$ *of initial nodes,*
- *a family* $\delta = (\delta_B)_{B \in \Phi(\mathcal{A})}$ *of relations* $\delta_B \subseteq N \times N$*; we also denote by* δ *the union of the relations* δ_B *for* $B \in \Phi(\mathcal{A})$*,*
- *an edge labeling* o *that associates a finite set* $\{(t_1, \prec_1), \ldots, (t_k, \prec_k)\}$*, of terms* t_i *paired with a relation* $\prec_i \in \mathcal{O}^=$ *with every edge* $(n, m) \in \delta$*, and*
- *a mapping* $\zeta : \mathcal{A} \to \{\mathrm{NF}, \mathrm{WF}, \mathrm{SF}\}$ *that associates a fairness condition with every parameterized action in* $\mathcal{A}$*; the possible values represent no fairness, weak fairness, and strong fairness.*

We say that the quantified-action $B \in \Phi(\mathcal{A})$ can be taken at node $n \in N$ *iff* $(n, m) \in \delta_B$ *holds for some* $m \in N$*, and denote by* $En(B) \subseteq N$ *the set of nodes where* B *can be taken.*

A PPD is a finite graph whose nodes are labelled with sets of (possibly negated) predicates, and whose edges are labelled with (the names of) quantified actions as well as optional annotations that assert certain expressions to decrease with respect to an ordering in $\mathcal{O}^=$. Intuitively, a node of a PPD represents the set of system states that satisfy the formulas contained in the node. (We indifferently write n for the set and the conjunction of its elements.) An edge (n, m) is labelled with a quantified action A if A can cause a transition from a state represented by n to a state represented by m. A parameterized action may have an associated fairness condition.

We now define *runs* and *traces* through a PPD as the set of those behaviors that correspond to fair runs satisfying the node and edge labels. To evaluate the fairness conditions we identify the enabling condition of an action $A \in \mathcal{A}$ with the existence of A-labelled edges at a given node. For a term x and two states s and t, if x holds at the state s or the pair of states s and t, then we write $s[\![x]\!]$ or $s[\![x]\!]t$, respectively.

Definition 3. *Let* $\mathcal{G} = (N, I, \delta, o, \zeta)$ *be a* PPD *over* $\mathcal{P}, \mathcal{A}$ *and* $\Phi(\mathcal{A})$*. A run of* $\mathcal{G}$ *is an* ω*-sequence* $\rho = (s_0, n_0, A_0)\ (s_1, n_1, A_1) \ldots$ *of triples where* s_i *is a state,* $n_i \in N$ *is a node and* $A_i \in \Phi(\mathcal{A}) \cup \{\tau\}$ *is an action such that all of the following conditions hold:*

1. $n_0 \in I$ *is an initial node.*
2. $s_i[\![n_i]\!]$ *holds for all* $i \in \mathbb{N}$*.*
3. *For all* $i \in \mathbb{N}$ *either* $A_i = \tau$ *and* $n_i = n_{i+1}$ *or* $A_i \in \Phi(\mathcal{A})$ *and* $(n_i, n_{i+1}) \in \delta_{A_i}$*.*
4. *If* $A_i \in \Phi(\mathcal{A})$ *and* $(t, \prec) \in o(n_i, n_{i+1})$*, then* $s_{i+1}[\![t]\!] \prec s_i[\![t]\!]$*.*
5. *If* $A_i = \tau$ *then* $s_{i+1}[\![t]\!] \preceq s_i[\![t]\!]$ *holds whenever* $(t, \prec) \in o(n_i, m)$ *for some* $m \in N$*.*
6. *For every quantified-action* $B \in \Phi(\mathcal{A})$ *of parameterized action* $A(k) \in \mathcal{A}$ *such that* $\zeta(A(k)) = \mathrm{WF}$ *there are infinitely many* $i \in \mathbb{N}$ *such that either* $A_i = B$ *or* $n_i \notin En(B)$*.*
7. *For every action* $B \in \Phi(\mathcal{A})$ *of parameterized action* $A(k) \in \mathcal{A}$ *such that* $\zeta(A(k)) = \mathrm{SF}$*, either* $A_i = B$ *holds for infinitely many* $i \in \mathbb{N}$ *or* $n_i \in En(B)$ *holds for only finitely many* $i \in \mathbb{N}$*.*

We write runs$(\mathcal{G})$ *to denote the set of runs of* $\mathcal{G}$*.*

The set $tr(\mathcal{G})$ traces through $\mathcal{G}$ consists of all behaviors $\sigma = s_0 s_1 \ldots$ such that there exists a run $\rho = (s_0, n_0, A_0)(s_1, n_1, A_1)\ldots$ of $\mathcal{G}$ based on the states in σ.

4 Verification Using PPDs

In linear-time formalisms such as TLA*, trace inclusion is the appropriate implementation relation. Thus, a specification *parSpec* implements a property or high level specification F if and only if the implication $parSpec \rightarrow F$ is valid [7]. This implication can be refine into two conditions by using PPDs: first, all behaviors allowed by *parSpec* must also be traces through the diagram and second, every trace through the diagram must satisfy F. Although both conditions are stated in terms of trace inclusion, following CANSELL ET.AL. [4], two different techniques are used here. To show that a PPD is a correct representation of a specification, we consider the node and edge labels of the diagram as predicates on the concrete state space of *parSpec*, and reduce trace inclusion to a set of proof obligations that concern individual states and transitions. On the other hand, to show that the diagram implies the high level property, we regard all labels as Boolean variables. The parameterized predicate diagram can therefore be encoded as a finite labelled transition system, whose temporal properties are established by model checking.

4.1 Relating Diagrams with Specifications

A PPD $\mathcal{G}$ is said to conform to a specification *parSpec* if every behavior that satisfies *parSpec* is a trace through $\mathcal{G}$. In general, reasoning about entire behaviors are required for proving this conformance. The following theorem essentially introduces a set of first-order ("local") verification conditions that are sufficient to establish conformance of a PPD to a parameterized system specification in standard form (Formula 1). The proof of Theorem 1 is given in [10].

Theorem 1. *Let $\mathcal{G} = (N, I, \delta, o, \zeta)$ be a PPD over $\mathcal{P}, \mathcal{A}$ and $\Phi(\mathcal{A})$ as defined and let $parSpec \equiv Init \wedge \Box[\exists\, k \in M : Next(k)]_v \wedge \forall\, k \in M : L(k)$ be a parameterized system. If the following conditions hold then $\mathcal{G}$ conforms to parSpec:*

1. $\models Init \rightarrow \bigvee_{n \in I} n.$
2. $\approx n \wedge [\exists\, k \in M : Next(k)]_v \rightarrow n' \vee \bigvee_{(m,B):(n,m)\in\delta_B} \langle B \rangle_v \wedge m'$
3. *For all $n, m \in N$ and all $(t, \prec) \in o(n, m)$*
   (a) $\approx n \wedge m' \wedge \bigvee_{B:(n,m)\in\delta_B} \langle B \rangle_v \rightarrow t' \prec t$
   (b) $\approx n \wedge [\exists k \in M : Next(k)]_v \wedge n' \rightarrow t' \preceq t.$
4. *For every parameterized action $A(k) \in \mathcal{A}$ such that $\zeta(A(k)) \neq$ NF*
   (a) *If $\zeta(A(k)) =$ WF then for every quantified action B of $A(k)$, $\models parSpec \rightarrow \mathrm{WF}_v(B)$.*
   (b) *If $\zeta(A(k)) =$ SF then for every quantified action B of $A(k)$, $\models parSpec \rightarrow \mathrm{SF}_v(B)$.*

(c) $|\approx n \rightarrow \langle \text{ENABLED } B \rangle_v$ *holds for every quantified action of* $A(k)$, B, *whenever* $n \in En(B)$.
(d) $|\approx n \wedge \langle B \rangle_v \rightarrow \neg m'$ *holds for all* $n, m \in N$ *and for every quantified action of* $A(k)$, B, *such that* $(n, m) \notin \delta_B$.

Condition 1 asserts that every initial state of the system must be covered by some initial node. This ensures that every run of the system can start at some initial node of the diagram. Condition 2 asserts that from every node, every transition, if it is enabled then it must have a place to go, i.e., there is a successor node which represents the successor state of the transition. It proves that every run of the system can stay in the diagram. Condition 3 is related to the ordering annotations and Condition 4 is related to the fairness conditions.

4.2 Model Checking PPDs

For the proof that all traces through a PPD satisfy some property F we view the diagram as a finite transition system that is amenable to model checking. All predicates and actions that appear as labels of nodes or edges are then viewed as atomic propositions.

Regarding PPDs as finite labelled transition systems, their runs can be encoded in the input language of standard model checkers such as SPIN [6]. Two variables indicate the current node and the last action taken. The predicates in $\mathcal{P}$ are represented by boolean variables, which are updated according to the label of the current node, nondeterministically, if that label contains neither P nor $\neg P$. We also add variables $b_{(t,\prec)}$, for every term t and relation $\prec \in \mathcal{O}$ such that $(t, \prec)$ appears in some ordering annotation $o(n, m)$. These variables are set to 2 if the last transition taken is labelled by $(t, \prec)$, to 1 if it is labelled by $(t, \preceq)$ or is stuttering transition and to 0 otherwise. Whereas the fairness conditions associated with the actions of a diagram are easily expressed as LTL (Linear Temporal Logic) assumptions for SPIN.

5 Tickets Protocol: A Case Study

The Tickets protocol is a mutual exclusion protocol designed for multi-client systems operating on a shared memory. In order to access the critical section, every client executes the protocol based on the first-in first-served access policy. The protocol works as follows. Initially, t and s store the same initial value. When requesting the access to the critical section, a client stores the value of the current ticket t in its local variable a. A new ticket is then emitted by incrementing t. Clients wait for their turn until the value of their local variable a is equal to the value of s. After the elaboration inside the critical section, a process releases it and the current turn is updated by incrementing s. During the execution the global state of the protocol consists of the internal state (current value of the local variable) of each process together with the current value of s and t.

module TICKETS

Specification

$$
\begin{aligned}
\mathit{Init} &\equiv \forall\, k \in M \,:\, pc[k] = 1 \wedge a[k] = 0 \wedge s = 1 \wedge t = 1 \\
\mathit{Request}(k) &\equiv \wedge\ pc[k] = 1 \wedge pc' = [pc \text{ EXCEPT } !k = 2] \\
&\quad \wedge\ a' = [a \text{ EXCEPT } !k = t] \wedge t' = t + 1 \wedge s' = s \\
\mathit{Grant}(k) &\equiv \wedge\ pc[k] = 2 \wedge a[k] = s \wedge pc' = [pc \text{ EXCEPT } !k = 3] \\
&\quad \wedge\ s' = s \wedge t' = t \wedge a' = a \\
\mathit{Release}(k) &\equiv \wedge\ pc[k] = 3 \wedge pc' = [pc \text{ EXCEPT } !k = 1] \wedge s' = s + 1 \\
&\quad \wedge\ a' = a \wedge t' = t \\
v &\equiv \langle pc, a, s, t \rangle \\
\mathit{Next}(k) &\equiv \mathit{Request}(k) \vee \mathit{Grant}(k) \vee \mathit{Release}(k) \\
L(k) &\equiv \wedge\ \mathrm{WF}_v(\mathit{Grant}(k)) \\
&\quad \wedge\ \mathrm{WF}_v(\mathit{Release}(k)) \\
\mathit{Tickets} &\equiv \mathit{Init} \wedge \Box[\exists\, k \in M \,:\, \mathit{Next}(k)]_v \wedge \forall\, k \in M \,:\, L(k)
\end{aligned}
$$

Theorem

$\mathit{Tickets} \rightarrow \forall\, k \in M \,:\, \Box(pc[k] = 2 \rightarrow \Diamond pc[k] = 3)$

Fig. 1. Tickets protocol for $n \geq 1$ processes

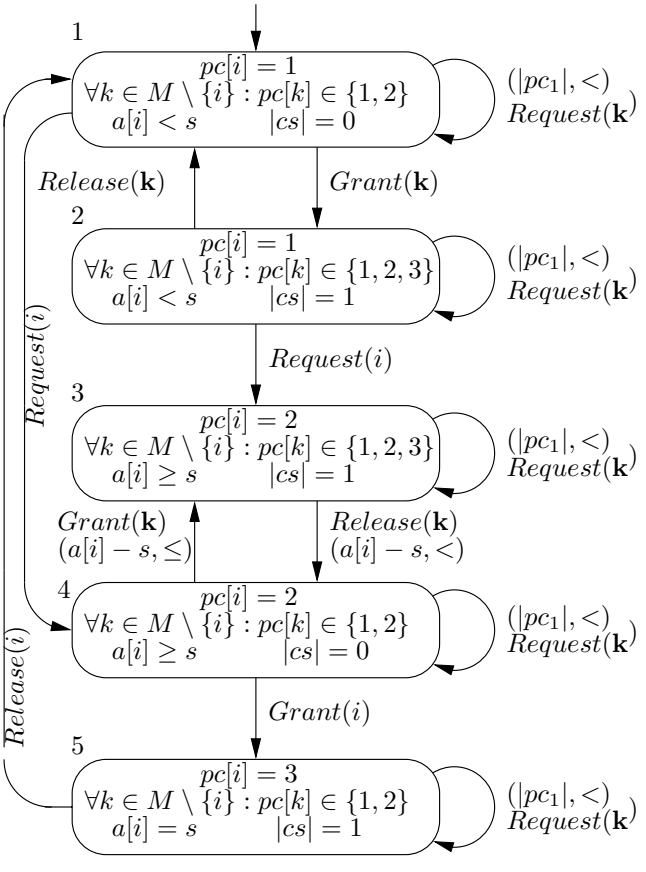


Fig. 2. PPD for Tickets protocol for $n \geq 1$ processes

Our objective is to prove that the protocol satisfies the so-called *individual accessibility*, that says that every time a process requests to enter its critical section, it will be eventually allowed to enter its critical section. This property is a universal property.

The specification of the Tickets protocol as well as the property to be proved in TLA$^+$ [7] style are shown in Figure 1.

Let pc_1 and cs be the set of processes whose pc is equal to 1 and 3, respectively, the diagram in Figure 2 is a suitable PPD for Tickets protocol for $n \geq 1$ processes. We associate an ordering annotation $(|pc_1|, <)$ with the loop of every node to ensure that eventually the system will leave the loops due to the finiteness of M. We also associate the ordering annotation $(a[i] - s, <)$ with the edge from node 3 to node 4 and associate the ordering annotation $(a[i] - s, \leq)$ with the edge from node 4 to node 3 for avoiding loops that may happen between the pair of nodes. Thus, we can ensure that eventually process i is allowed to enter its critical section. The choice of the orderings is based on the fact that whenever $pc[i] = 2$ then the difference between $a[i]$ and s decreases whenever some process other than i leaves its critical section and increments the value of s by 1.

The PPD in Figure 2 conforms to the *Tickets* specification in Figure 1. We can use Theorem 1 for proving this conformance. For example we have:

$$
\begin{aligned}
&- \ \mathit{Init} \rightarrow \begin{pmatrix} pc[i] = 1 \\ \land\ \forall k \in M \setminus \{i\} \ :\ pc[k] \in \{1,2\} \\ \land\ a[i] < s \\ \land\ |cs| = 0 \end{pmatrix}. \\
&- \ \begin{pmatrix} pc[i] = 1 \\ \land\ \forall k \in M \setminus \{i\} \ :\ pc[k] \in \{1,2\} \\ \land\ a[i] < s \\ \land\ |cs| = 0 \end{pmatrix} \land [\exists\, k \in M \ :\ \mathit{Next}(k)]_v \rightarrow \\
&\quad \lor \begin{pmatrix} pc[i]' = 1 \\ \land\ \forall k \in M \setminus \{i\} \ :\ pc[k]' \in \{1,2\} \\ \land\ a[i]' < s \\ \land\ |cs'| = 0 \end{pmatrix} \\
&\quad \lor \langle \mathit{Request}(i) \rangle_v \land \begin{pmatrix} pc[i]' = 2 \\ \land\ \forall k \in M \setminus \{i\} \ :\ pc[k]' \in \{1,2\} \\ \land\ a[i]' \geq s \\ \land\ |cs'| = 0 \end{pmatrix} \\
&\quad \lor \langle \exists\, k \in M \setminus \{i\} : \mathit{Request}(k) \rangle_v \land \begin{pmatrix} pc[i]' = 1 \\ \land\ \forall k \in M \setminus \{i\} : pc[k]' \in \{1,2\} \\ \land\ a[i]' < s \\ \land\ |cs'| = 0 \end{pmatrix} \\
&\quad \lor \langle \exists\, k \in M \setminus \{i\} : \mathit{Grant}(k) \rangle_v \land \begin{pmatrix} pc[i]' = 1 \\ \land\ \forall k \in M \setminus \{i\} : pc[k]' \in \{1,2,3\} \\ \land\ a[i]' < s \\ \land |cs'| = 1 \end{pmatrix}.
\end{aligned}
$$

$$- \begin{pmatrix} \wedge \wedge\ pc[i] = 2 \\ \quad \wedge\ \forall k \in M \setminus \{i\} : pc[k] \in \{1,2,3\} \\ \quad \wedge\ a[i] \geq s \\ \quad \wedge\ |cs| = 1 \\ \wedge \wedge\ pc[i]' = 2 \\ \quad \wedge\ \forall k \in M \setminus \{i\} : pc[k]' \in \{1,2\} \\ \quad \wedge\ a[i]' \geq s \\ \quad \wedge\ |cs'| = 0 \\ \wedge\ \langle \exists\, k \in M \setminus \{i\} : Release(k) \rangle_v \end{pmatrix} \longrightarrow (a[i]' - s') < (a[i] - s).$$

Using the diagram in Figure 2 we can prove that it is always the case that whenever process i request to enter its critical section, it will eventually enters its critical section, i.e. we can prove the validity of formula $Tickets \rightarrow \Box(pc[i] = 2 \rightarrow \Diamond pc[i] = 3)$. Moreover, since we have just applied the standard quantifier introduction rule of first-order logic, this implies the validity of formula $\forall\, k \in M : Tickets \rightarrow \Box(pc_k = 2 \rightarrow \Diamond pc_k = 3)$ as required.

Encoding the PPD in Promela, the input language of SPIN, as described in Section 4.2, and then model-checking the resulted transition system using SPIN, we can verify the individual accessibility of the Tickets protocol.

6 Conclusion and Future Work

We have proposed a method for verifying universal properties of parameterized parallel systems using Parameterized Predicate Diagrams. In this work we have restricted to a class of parameterized systems that are interleaving and consist of a finitely, but arbitrarily, discrete components. The parameterized systems are represented as parameterized TLA* specifications. The verification is done deductively and algorithmically by means of diagrams. Our diagrams can be viewed as the abstract representation of parameterized systems, i.e. we represent a family of processes in a single diagram. The same spirit but using difference formalism is the work from BAUKUS ET.AL.[1]. They propose a method for the verification of universal properties of parameterized networks based on the transformation of an infinite family of systems into a single WS1S [3] transition system and applying abstraction techniques on this system.

Using the Tickets protocol as case study, we have shown that PPDs can be used to prove the universal properties. For handling the universal properties we distinguish some single arbitrary process from the rest of processes. This can be extended for proving the properties that related to some set of particular processes. The idea is to consider those processes separately from the rest of the processes. In this case, some more complex reasoning might be necessary to do, such as induction on the number of processes, depending on the protocol at hand.

For the practical application of our method, tool support is essential. We have implemented a prototype tool that can be used to generate PPDs [10]. This tool still needs to improvement, in particular in the aspect of graphical user interface.

References

1. Kai Baukus, Saddek Bensalem, Yassine Lakhnech and Karsten Stahl. Abstracting WS1S Systems to Verify Parameterized Networks. In *Proceeding of the 6th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2000)*, Volume 1785 of *Lecture Notes in Computer Science*, pages 188-203. Springer, 2000.
2. M. Bozzano and G. Delzanno. Beyond Parameterized Verification. In *Proceedings of International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2002*. Volume 2280 of *Lecture Notes in Computer Science*, pages 221-235. Springer, 2002.
3. J.R. Büchi. Weak second-order arithmatic and finite automata. Z. Math. Logik Grundl. Math., 6:66-92, 1960.
4. Dominique Cansell, Dominique Méry and Stephan Merz. Predicate diagrams for the verification of reactive systems. In 2^{nd} *Intl. Conf. on Integrated Formal Methods (IFM 2000)*, vol. 1945 of *Lectures Notes in Computer Science*, Dagstuhl, Germany, November 2000. Springer-Verlag.
5. E.A. Emerson and K.S. Namjoshi. Verification of a parameterized bus arbitration protocol. Volume 1427 of *Lecture Notes in Computer Science*, pp. 452–463. Springer,1998.
6. G. Holzmann. The SPIN model checker. *IEEE Trans. on software engineering*, 16(5):1512-1542. May 1997.
7. Leslie Lamport. The Temporal Logic of Actions. *ACM Transactions on Programming Languages and Systems*, 16(3) : 872-923, May 1994.
8. Zohar Manna and Amir Pnueli. Verification of parameterized programs. In *Specification and Validation Methods (E. Borger, ed.)*, Oxford University Press, pp. 167-230, 1994.
9. Stephan Merz. Logic-based analysis of reactive systems: hiding, composition and abstraction. Habilitationsschrift. Institut für Informatik. Ludwig-Maximillians-Universität, Munich Germany. December 2001.
10. Cecilia E. Nugraheni. Predicate diagrams as basis for the verification of reactive systems. PhD Thesis. Institut für Informatik. Ludwig-Maximillians-Universität, Munich Germany. February 2004.

2d Polynomial Interpolation: A Symbolic Approach with Mathematica

Ali Yazici[1], Irfan Altas[2], and Tanil Ergenc[3]

[1] Computer Engineering Department, TOBB University of Economics & Technology, Ankara - Turkey
aliyazici@etu.edu.tr
[2] Charles Sturt University, Wagga Wagga - Australia
ialtas@csu.edu.au
[3] Mathematics Department, Middle East Technical University, Ankara - Turkey
tanil@metu.edu.tr

Abstract. This paper extends a previous work done by the same authors on teaching 1d polynomial interpolation using Mathematica [1] to higher dimensions. In this work, it is intended to simplify the the theoretical discussions in presenting multidimensional interpolation in a classroom environment by employing Mathematica's symbolic properties. In addition to symbolic derivations, some numerical tests are provided to show the interesting properties of the higher dimensional interpolation problem. Runge's phenomenon was displayed for 2d polynomial interpolation.

1 Introduction

The combination of the power of symbolic computations and the reliability of numerical methods together forms a solid base in the teaching of mathematical methods to students. Teaching using symbolic computations can create opportunity for students to deal with real life problems without worrying about underlaying cumbersome manipulations of algebraic operations [1, 2]. Students use software tools such as Mathematica or Maple to test mathematical ideas and access the Internet to develop an awareness of the wider learning environment. Kaput [3] has suggested that the mathematical thinking ability to recognize translation from one representation of a function to another can be assisted by the use of computers. Our experience in implementing Mathematica in our teaching certainly indicates similar findings. In this work, we will illustrate using Mathematica in teaching of interpolation methods, specifically multidimensional interpolation.

Polynomials are widely used as interpolating functions since it is easy to work with polynomials such as calculating derivatives. There are a number of methods to calculate the interpolating functions such as Lagrange, divided difference and cubic splines[4, 5]. Majority of these algorithms is appropriate for symbolic computations. In [1] we demonstrated teaching these one dimensional interpolation methods with the help of symbolic computation package, Mathematica.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 463–471, 2005.

The generalization of one dimensional polynomial interpolation to the multidimensional one is not straightforward. We can not expect to interpolate uniquely from an arbitrary set of data points. We illustrate this point in our teaching with the help of the symbolic computation as it is demonstrated in the following sections. It is also worth to note that some nice properties used for efficient calculations of one dimensional interpolating functions may not be carried over to multidimensional case. We discuss multidimensional interpolation problem in Section 2. In Section 3, we utilize a symbolic computation package, Mathematica, to illustrate 2d interpolation in our teaching. Conclusions are presented in Section 4.

2 Multidimensional Interpolation

A polynomial interpolates a function if the values of the polynomial and some of its derivatives match the corresponding values and derivatives of the function at some given interpolation points. The theory and application of one dimensional interpolation are well developed and can be found in most books on numerical analysis and approximation theory [6].

Unlike some other numerical methods the generalization of one dimensional interpolation to multidimensional case is not straight forward even though multidimensional interpolation has been widely implemented in many areas such as image rendering and numerical solution of partial differential equations.

For 1d interpolation problem, Lagrange interpolation generates a unique polynomial. However, the problem of multidimensional interpolation is a difficult one since it might not exist for some arbitrarily selected (or given) interpolation points. In our teaching, we emphasize the existence issue of multidimensional interpolation by letting students to perform a number of laboratory exercises in Mathematica.

2.1 Generalized Lagrange Interpolation

Lagrange interpolating polynomials can be used to interpolate 2d data or approximate a given function f(x,y). Let $p(x_i, y_i) = f(x_i, y_i)$, for $i = 0, 1, \ldots, n$, and $j = 0, 1, \ldots, m$. Then, one defines the generalized Lagrange interpolating polynomial $p(x, y)$ for $(n + 1)(m + 1)$ points by

$$p(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} f(x_i, y_j) L_{ij}(x,y) \tag{1}$$

where

$$L_{ij}(x,y) = L_i(x) \overline{L}_j(y) \tag{2}$$

and $L_i(x)$ and$\overline{L}_j(y)$ are the usual Lagrange interpolating polynomials in x and y respectively. Now, the polynomial is written as follows:

$$p(x,y) = \sum_{i=0}^{n} L_i(x) p_i(y) \tag{3}$$

and

$$p_i(y) = \sum_{j=0}^{m} f(x_i, y_j)\overline{L}_j(y) \tag{4}$$

In case of uniformly distributed points in each direction, order of the error term can be depicted by the following theorem [7]:

Theorem 1. *Let f(x,y) be in C^{n+1} with respect to x and in C^{m+1} with respect to y. Also, take uniformly spaced points in each direction, that is, $x_{i+1} - x_i = h_x$ and $y_{i+1} - y_i = h_y$. Then*

$$|f(x,y) - p(x,y)| = O(h_x^{n+1}) + O(h_y^{m+1}) \tag{5}$$

We will demonstrate the symbolic derivation of 2d Lagrange interpolating polynomials and Runge's phenomenon using Mathematica in Sections 3.1 and 3.2 respectively.

2.2 Taylor Series Approach

We provide students with various distributions of interpolation points where some set of interpolation points leads non-existence of multidimensional interpolation. Students employ Mathematica to prove that the interpolation polynomial does not exist. We restrict our discussions to two dimensional interpolation and choose as the interpolating function the Taylor polynomials. This way students will start exploring the existence of interpolation with a very familiar setting, namely Taylor polynomials. Meanwhile, it would be easy for them to see distribution of points. Once students have grasped concepts/issues of multidimensional interpolation, settings can be changed for 3 or more dimensions as well as for other types of constructing of interpolation polynomials.

Given a function $u = u(x,y) \in C^{n+1}(\Omega)$ where $\Omega \subset R^2$. We wish to interpolate u using its values on a set of distinct interpolation points $Z = \{(x_k, y_k) \in \Omega, k = 0, \ldots, m = \frac{n^2+3n}{2}\}$ by a polynomial of degree n. The value of u at (x_k, y_k) is denoted by $u_k = u(x_k, y_k)$. For brevity, we use the notation $p_k = (x_k, y_k)$. From the Taylor series expansion of $u(x,y)$ about the point p_0 (where p_0 can be considered as the origin without loss of generality) we have

$$u(x_k, y_k) \approx u_0 + \sum_{i=1}^{n} \frac{1}{i!}[x_k \frac{\partial}{\partial x} + y_k \frac{\partial}{\partial y}]^i u(x,y)|_{p_0}; k = 1, \ldots, m = \frac{n^2+3n}{2} \tag{6}$$

where the remainder term is neglected.

The n^{th} degree interpolation polynomial, $Q_n(x,y)$, which interpolates the function $u(x,y)$, can be written down explicitly as

$$Q_n(x,y) = u_0 + \frac{1}{1!}(\frac{\partial u}{\partial x})_{(0,0)}x + \frac{1}{1!}(\frac{\partial u}{\partial y})_{(0,0)}y + \cdots + \frac{1}{n!}(\frac{\partial^n u}{\partial x^n})_{(0,0)}y^n \tag{7}$$

or

$$Q_n(x,y) = c_0 + c_1 x + c_2 y + \cdots + c_{\frac{n^2+3n}{2}} y^n \tag{8}$$

Once the unknowns c_0 , c_1, and so on are determined (hence, the derivatives of u at $(0,0)$) the interpolation polynomial can be written down in terms of u_0, u_1,...,u_n. One of the main reasons to define interpolation polynomial in the context of Taylor polynomial is to prepare students to derive finite difference formulas for the solution of differential equations later in the course.

For a given set of interpolation points

$$Q_n(x_k, y_k) = u(x_k, y_k), \;\; k = 0, 1, \ldots, \frac{n^2 + 3n}{2} \tag{9}$$

we can rewrite Eq. 8 in matrix form as

$$Aw = b \tag{10}$$

where b is the left hand side of Eq. 9, w is the vector of the unknown coefficients and A is the coefficient matrix of the system obtained from the same equation. We denote the determinant of A by $det(A)$. For the existence of a unique interpolation polynomial it is necessary that the determinant $det(A) \neq 0$.

A numerical example to this approach will be presented in Section 3.3.

3 Mathematica Experiments

3.1 Deriving Generalized Lagrange Polynomials

In this section, we propose a set of Mathematica commands to derive the generalized product form of the 1d Lagrange interpolating polynomials to interpolate a function F(x,y) of two variables x and y. For the purpose of this study, we do not use the build in interpolation functions provided by Mathematica. With the same purpose in mind, the Mathematica codes are kept as simple as possible avoiding optimization and complex coding.

1. First, define a set of abscissas and ordinates and a general function F.
   $In[1] := Set\,[X, \{x0, x1\}];\; Set\,[Y, \{y0, y1\}];\; F[x\_, y\_] := f@\{x, y\}$

2. Define formally the Lagrange polynomials in y direction
   $In[2] := Ly[j\_, y\_] := Product[If[Equal\,[i, j], 1,\; Divide\,[y - Y[[i]], Y[[j]] - Y[[i]]]],$
   $\{i, 1, Length[Y]\}]$

3. In a similar manner, define Lagrange polynomials in x direction
   $In[3] := Lx[i\_, x\_] := Product[If[Equal\,[i, j], 1,\; Divide\,[x - X[[i]], X[[j]] - X[[i]]]],$
   $\{i, 1, Length[X]\}]$

4. Form the product form of the Lagrange polynomials
   $In[4] := L[i\_, j\_, x\_, y\_] := Times\;[Lx[i, x], Ly[j, y]]$

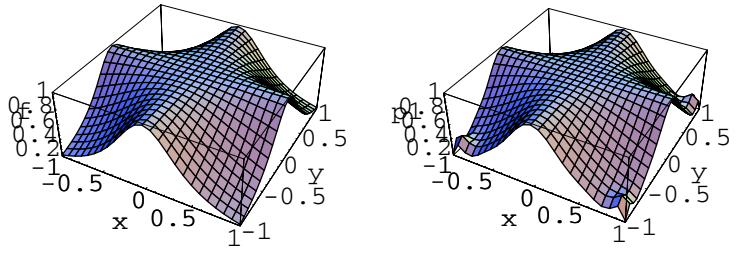


Fig. 1. Plot of f (left) and p1(x,y) (right) for 121 interpolation points

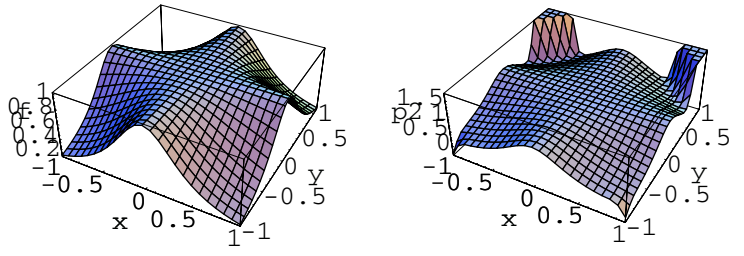


Fig. 2. Plot of f (left) and p2(x,y) (right) for 441 interpolation points

5. Finally, form the interpolating polynomial p(x,y) as depicted by Eq.6 above
$In[5] := Set\,[p[x\_, y\_], Sum\,[\,Sum\,[\,Times\ [F[X[[i]], Y[[j]]], L[i, j, x, y]],$
$\{j, 1, Length[Y]\}], \{i, 1, Length[X]\}]]$

$Out[5] = \frac{(x-x1)(y-y1)f[x0,y0]}{(x0-x1)(y0-y1)} + \frac{(x-x1)(y-y0)f[x0,y1]}{(x0-x1)(-y0+y1)} +$
$\frac{(x-x0)(y-y1)f[x1,y0]}{(-x0+x1)(y0-y1)} + \frac{(x-x0)(y-y0)f[x1,y1]}{(-x0+x1)(-y0+y1)}$

3.2 Exploring Runge's Phenomenon

Runge's phenomenon is a well-known problem in 1d polynomial interpolation which displays divergence at the end points for a specific function [1]. In an attempt to display the same, we have chosen $f(x, y) = \frac{1}{1+5x^2y^2}$, over $[-1, 1]$x$[-1, 1]$ and determined the corresponding 2d Lagrange interpolating polynomials p1 and p2 by employing 121 and 441 interpolation points respectively. The results are illustrated in Fig.1 and Fig2. below. We observed that especially near the corners, the error is considerably high displaying an expected behavior as in the 1d Runge function. For example, the error at $[0.9, 0.9]$ is reported to be 0.143539 and 23.1168 for p1 and p2 respectively.

3.3 Numerical Example for the Taylor's Approach

In this section, a numerical example is provided to show some of the properties of the 2d interpolation problem using Taylor expansions.

Example: A bush walker stands at the point p_0 aiming to reach the point E (see Fig.3). Before he proceeds his journey he wishes to estimate the elevation of the point, E. From Digital Elevation Modelling (DEM) map in his PDA computer he figures out the elevation of 6 surrounding points in Fig.3 as 700m at p_0, 740m at p_1, 860m at p_2, 810m at p_3, 780m at p_4, 910m at p_5. He tries to fit a second degree interpolation polynomial for the data with h=500m and r=800m by employing an interpolation algorithm on his PDA computer. However, he receives an error message indicating that the second degree interpolation polynomial does not exist for the data given.

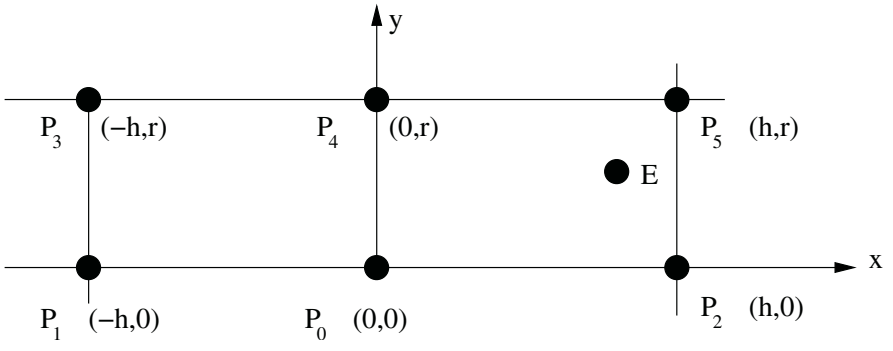


Fig. 3. Estimation of the point of elevation E

1. Can you confirm whether the bush walker obtained the correct result from his PDA?
2. Check whether a linear interpolation can be obtained if the data points are chosen as p0, p2 and p5. If yes, approximate the elevation at the point E=(400,600) from this first degree interpolating polynomial.
3. The bush walker was not sure whether the existence problem was due to his choice of h=500m and r=800m. Can you prove for the bush walker that the existence of the second degree interpolation does not depend on the choice of the values of h and r?
4. Prove the following claim for the interpolation polynomials with degrees $n = 2$ and $n = 3$ via Mathematica: If there are two lines both parallel to the same coordinate axis and with n+1 points on each one, then the determinant det(A) of the system determined by Eq.10 is zero[8]. (The case n=2 is a special case where the conditions only occur when the axis is one of the lines).

Solution: Mathematica instructions for the solution of the problem is outlined below:

1. Define the polynomial Q, and form the coefficient matrix A. Compute the determinant(A)of Eq.10.

In[6]:= $Set[Q, c_0 + c_1x + c_2y + c_3x^2 + c_4xy + c_5y^2]$
Out[6]= $c_0 + c_1x + c_3x^2 + c_2y + c_4xy + c_5y^2$

In[7]:= $Set[A, \{\{1,0,0,0,0,0\}, \{1,-h,0,h^2,0,0\}, \{1,h,0,h^2,0,0\},$
$\{1,-h,r,h^2,-hr,r^2\}, \{1,0,r,0,0,r^2\}, \{1,h,r,h^2,hr,r^2\}\}]$
Out[7]= $\{\{1,0,0,0,0,0\}, \{1,-h,0,h^2,0,0\}, \{1,h,0,h^2,0,0\},$
$\{1,-h,r,h^2,-hr,r^2\}, \{1,0,r,0,0,r^2\}, \{1,h,r,h^2,hr,r^2\}\}$

In[8]:= $MatrixForm[A]$

Out[8]=
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -h & 0 & h^2 & 0 & 0 \\ 1 & h & 0 & h^2 & 0 & 0 \\ 1 & -h & r & h^2 & -hr & r^2 \\ 1 & 0 & r & 0 & 0 & r^2 \\ 1 & h & r & h^2 & hr & r^2 \end{bmatrix}$$

In[9]:= $MatrixForm[A]$ `/. {h->500,r->800}`

Out[9]=
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -500 & 0 & 250000 & 0 & 0 \\ 1 & 500 & 0 & 250000 & 0 & 0 \\ 1 & -500 & 800 & 250000 & -400000 & 640000 \\ 1 & 0 & 800 & 0 & 0 & 640000 \\ 1 & 500 & 800 & 250000 & 400000 & 640000 \end{bmatrix}$$

In[10]:= $Det[A]$
Out[10]= 0

Determinant of the coefficient matrix A is zero independent of h and r. This confirms that the bush walker obtained the correct result from his PDA.

2. Form the 3x3 matrix B for the second part of the question (matrix form is not shown)
In[11]:= $Set[B, \{\{1,0,0\}, \{1,h,0\}, \{1,h,r\}\}]$ `/.{h->500,r->800}`
In[12]:= $Set[detB, Det[B]]$
Out[12]= 400000

In[13]:= $Set[c0m, \{\{700,0,0\}, \{860,h,r\}, \{910,h,r\}\}]$ `/. {h->500,r->800}`
In[14]:= $Set[detc0m, Det[c0m]]$
Out[14]= 280000000

In[15]:= $Set[c0, Divide[detc0m, detB]]$
Out[15]= 700

In[16]:= $Set\,[c1m, \{\{1, 700, 0\}, \{1, 860, 0\}, \{1, 910, r\}\}]$ `/. r->800`
In[17]:= $Set\,[detc1m, Det[c1m]]$
Out[17]= 128000

In[18]:= $Set\,[c1, Divide[detc1m, detB]]$
Out[18]= $\frac{8}{25}$

In[19]:= $Set\,[c2m, \{\{1, 0, 700\}, \{1, h, 860\}, \{1, h, 910\}\}]$ `/. h->500`
In[20]:= $Set\,[detc2m, Det[c2m]]$
In[21]:= $Set\,[c2, Divide[det2cm, detB]]$
Out[21]= $\frac{1}{16}$

In[22]:= $Set\,[Q1, c0 + c1x + c2y]$
Out[22]= $700 + \frac{8x}{25} + \frac{y}{16}$

In[23]:= $Set\,[QatE, c0 + c1x + c2y]$ `/. {x->400,y->600}`
Out[23]= $\frac{1731}{2}$

Thus, the interpolated elevation value at point E is $\frac{1731}{2}$.

3. Determinant of the original matrix (Out[8] above) is 0 proving that, the existence of the second degree interpolation does not depend on the choice of the values of h and r.
4. Without loss of generality one can assume that the lines are parallel to the x-axis and contain the origin. The case for n=2 is similar to part 3 above and hence omitted. For n=3, 10 interpolation points are required: 4 points on the x-axis, 4 on a line parallel to x-axis, and 2 are chosen randomly. For example, for the points $(0,0)$, $(t,0)$, $(k2,t)$, $(k22,t)$, $(k3t,0)$, $(k4t,s)$, $(k5t,s)$, $(k6t,s)$, $(u1,z1)$, $(u2,z2)$, the coefficient matrix is formed as follows:

In[26]:= `part4n3 :=` $\{\{1,0,0,0,0,0,0,0,0,0\}, \{1,t,0,t^2,0,0,t^3,0,0,0\},$
$\{1,k2t,0,k2^2t^2,0,0,k2^3t^3,0,0,0\}, \{1,k22t,0,k22^2t^2,0,0,k22^3t^3,0,0,0\},$
$\{1,k3t,s,k3^2t^2,k3ts,s^2,k3^3t^3,sk3^2t^2,k3ts^2,s^3\},$
$\{1,k4t,s,k4^2t^2,k4ts,s^2,k4^3t^3,sk4^2t^2,k4ts^2,s^3\},$
$\{1,k5t,s,k5^2t^2,k5ts,s^2,k5^3t^3,sk5^2t^2,k5ts^2,s^3\},$
$\{1,k6t,s,k6^2t^2,k6st,s^2,k6^3t^3,sk6^2t^2,k6ts^2,s^3\},$
$\{1,u1,z1,u1^2,u1z1,z1^2,u1^3,z1u1^2,u1z1^2,z1^3\},$
$\{1,u2,z2,u2^2,u2z2,z2^2,u2^3,z2u2^2,u2z2^2,z2^3\}\}$

In[27]:= $Set\,[detPart4, Det[part4n3]]$
Out[27]= 0
This proves the claim for n=3.

4 Conclusions

In this paper, a use of symbolic algebra software is demonstrated for teaching 2d polynomial interpolation problem in an educational setting. It has been demonstrated that, symbolic packages are quite effective in deriving the required

formula, and demonstrating existence issue of multidimensional interpolation. Our experiences with this approach show that students can grasp important and difficult concepts easily in a laboratory environment. The symbolic power provided by Mathematica, has provided a platform to discuss the fundamental and difficult issues related to multidimensional interpolation problem.

References

1. Yazici, A.,Altas, I. and Ergenc, T.: Symbolic Interpolation using Mathematica, Lecture Notes in Computer Science (LNCS), Eds. Sloot, P.M.A. et al., Springer-Verlag, Vol. 3039, Part IV, (2004) 365-370.
2. Reiter, C.A.: Exploring Hermite Interpolation with Mathematica, Primus, 2, 2(1992) 173-182.
3. Kaput, J.: Technology and Mathematics Education, in Handbooks of Research on Mathematics Teaching and Learning (Ed. Grouws, D.A.), MacMillan, New York (1992) 515-556.
4. De Boor, C.: A Practical Guide to Splines, Springer Verlag, (1978).
5. Mathews, J.H.: Numerical Methods For Computer Science, and Mathematics, Prentice-Hall International (1987).
6. Heath, M.T.: Scientific Computing: An Introductory Survey, McGraw-Hill International Editions (1997).
7. Linz, P.: Theoretical Numerical Analysis: An Introduction to Advanced Techniques, John-Wiley & Sons, Ltd. (1979).
8. Altas, I., Stephenson J.W.:Existence of Second Order Discretizations on Irregular Mesh, Appl. Math Lett. Vol. 2. No. 4, (1989) 315-318.

Analyzing the Synchronization of Chaotic Dynamical Systems with Mathematica: Part I

A. Iglesias* and A. Gálvez

Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n, E-39005, Santander, Spain
iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. One of the most interesting and striking issues in dynamical systems is the possibility to synchronize the behavior of several (either identical or different) chaotic systems. This is the first of a series of two papers (both included in this volume) describing a new Mathematica package developed by the authors, `ChaosSynchronization`, for the analysis of chaotic synchronization. In particular, this first paper is devoted to the analysis of the Pecora-Carroll scheme (the first chaotic synchronization scheme reported in the literature) as well as a recent modification based on partial connections. The performance of the package is discussed by means of several illustrative and interesting examples. In our opinion, this package provides the users with an excellent, user-friendly computer tool for learning and analyzing this exciting topic within a unified (symbolic, numerical and graphical) framework.

1 Introduction

The notion of *synchronization* is well known from the viewpoint of the classical mechanics since early 16th century. Since then, many other examples have been reported in the literature (see, for instance, [9] for electrical and mechanical systems). However, the possibility of synchronizing chaotic systems is not so intuitive, since these systems are very sensitive to small perturbations on the initial conditions and, therefore, close orbits of the system quickly become uncorrelated. Surpringly, in 1990 it was shown that certain subsystems of chaotic systems can be synchronized by linking them with common signals [7]. In particular, the authors reported the synchronization of two *identical* (i.e., two copies of the same system with the same parameter values) chaotic systems. They also show that, as the differences between those system parameters increase, synchronization is lost. Subsequent works showed that synchronization of non-identical chaotic systems is also possible.

This paper is the first of a series of two papers (both included in this volume) describing a new Mathematica package, `ChaosSynchronization`, for the analysis

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 472–481, 2005.

of chaotic synchronization. The package can be successfully applied to study the different scenarios (monotonic, oscillating, marginal, sized) appearing in the chaotic synchronization phenomena. The performance of the package is discussed by means of several illustrative and interesting examples. In our opinion, this package provides the users with an excellent, user-friendly computer tool for learning and analyzing this exciting topic within a unified (symbolic, numerical and graphical) framework [10].

This first paper analyzes the Pecora-Carroll (PC) scheme (the first chaotic synchronization scheme reported in the literature) and a recent modification based on partial connections. Such analysis is performed by using a new Mathematica package, `ChaosSynchronization`, developed by the authors. The structure of this paper is as follows: the Pecora-Carroll scheme is analyzed in Section 2. Then, Section 3 discusses a recent generalization based on partial connections. We start our discussion by loading the package [6]:

```
In[1]:=<<DynamicalSystems‘ChaosSynchronization‘
```

2 Pecora-Carroll Scheme for Chaotic Synchronization

In 1990 Pecora and Carroll [7] showed that when a state variable from a chaotic system is input into a replica subsystem of the original one, both systems can be synchronized identically. In mathematical terms, given a couple of autonomous n-dimensional identical chaotic systems $\dot{x_1} = f(x_1)$ and $\dot{x_2} = f(x_2)$ as a drive and response systems respectively, the basic idea of the Pecora-Carroll (PC) scheme is decomposing the drive system into two subsystems, $\dot{x_1} = (\dot{u_1}, \dot{v_1})$, with $x_1 \in \mathbb{R}^n$, $u_1 \in \mathbb{R}^p$ and $v_1 \in \mathbb{R}^q$ (where $n = p + q$) as:

$$\left.\begin{aligned} \dot{u_1} &= g(u_1, v_1) \\ \dot{v_1} &= h(u_1, v_1) \end{aligned}\right\} \text{ drive,} \tag{1}$$

and considering one of the decomposed subsystems as driving signal, say u_1, to be injected into the response system. This reduces the dimensionality of the response becoming

$$\left.\begin{aligned} \dot{v_2} &= h(u_1, v_2) \end{aligned}\right\} \text{ response,} \tag{2}$$

where u_1 is the set of connecting variables. Note that the dimensions of the drive and response systems are n and q respectively, with $n > q$. Note also that the system (1) is independent on the response system, whereas (2) is driven by u_1 (*unidirectional coupling*). The previous scheme given by eqs. (1)-(2) can be generalized by considering a response system given by:

$$\left.\begin{aligned} \dot{v_2} &= k(u_1, v_2) \end{aligned}\right\} \text{ response,} \tag{3}$$

that is, by assuming that the functions h and k describing the dynamics of $\dot{v_1}$ and $\dot{v_2}$ respectively must not be the same. This situation is usually referred to as *heterogeneous* (or *inhomogeneous*) *driving*. A simple example of this heterogeneous driving is a linear oscillator system driving a nonlinear pendulum.

As a first synchronization example, we consider the Lorenz system given by:
`In[2]:=Lorenz[x_,y_,z_]:= {σ(y-x),(r-z)x-y,x y-b z};`

It is useful to write this system as a linear system by taking the first-order approximation, $\dot{X} = J(X).X$ where the square matrix $J(X)$ is called the *Jacobian matrix* of the system. The `JacobianMatrix` command calculates the Jacobian matrix of a system with respect to its list of variables. For example:

`In[3]:=JacobianMatrix[Lorenz[x,y,z],{x,y,z}]]`

$$Out[3] := \begin{pmatrix} -\sigma & \sigma & 0 \\ r-z & -1 & -x \\ y & x & -b \end{pmatrix}$$

This system is known to exhibit chaotic behavior for certain values of the parameters. Given a dynamical system and a list of its variables the `Parameters` command returns all the parameters of such a system:

`In[4]:=Parameters[Lorenz[x,y,z],{x,y,z}]`

$Out[4] := \{b, r, \sigma\}$

With this system viewed as the transmitter or master system, we introduce the drive signal `y` which can be used at the receiver or slave system, to achieve asymptotic synchronization. Now, we consider the Jacobian matrix of the error dynamics between the slave and the master systems, the so-called *Jacobian Conditional Matrix (JCM)*. The `JCMatrix` command calculates such a matrix given: a dynamical system, its list of variables and a specific connection given by the driving variable. For instance:

`In[5]:=JCMatrix[Lorenz[x,y,z],{x,y,z},y]]`

$$Out[5] := \begin{pmatrix} -\sigma & 0 \\ y & -b \end{pmatrix}$$

whose eigenvalues are:

`In[6]:=Eigenvalues[%]`

$Out[6] := \{-\sigma, -b\}$

It can be proved that the master and slave systems will synchronize if those eigenvalues are both negative [8]. From this, it becomes clear that the synchronization depends not only on the given dynamical system and the injected variable, but also on the parameter values of the system. In their original paper, Pecora and Carroll took the values:

`In[7]:= param={b->8/3,r->60,σ->10};`

Note that, from *Out[6]* and `In[7]`, for this particular choice of the system parameters, the connection in variable `y` is synchronizing. On the contrary, injecting variable `x` leads to the Jacobian conditional matrix:

`In[8]:=JCMatrix[Lorenz[x,y,z],{x,y,z},x]]`

$$Out[8] := \begin{pmatrix} -1 & -x \\ x & -b \end{pmatrix}$$

Because the eigenvalues do depend on the system variables, synchronization cannot be so easily determined. Actually, the solution to this question is given by the Lyapunov exponents of the difference system, since they indicate if small

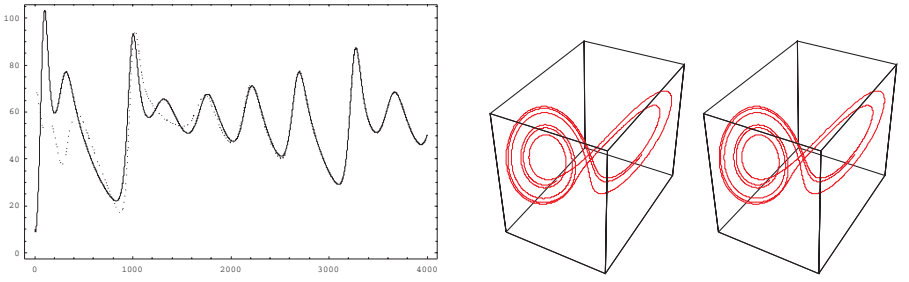


Fig. 1. Synchronization of two Lorenz systems with the Pecora-Carroll scheme: (left) time series for the variable z when injecting the variable x; (right) chaotic attractors of the drive and the receiver systems when injecting the variable y

displacements of trajectories are along stable or unstable directions. In fact, the Lyapunov exponents are the generalization of the Jacobian matrix for stability purposes. Lyapunov exponents of the v_2-subsystem for a particular drive trajectory are called *conditional Lyapunov exponents*. If we are looking for a stable subsystem, then all the exponents must be negative so that all the small perturbations will exponentially decay to zero. Therefore, synchronization occurs only if the conditional Lyapunov exponents are all negative.

Through this new procedure we can determine how many stable connections can we get for the Lorenz system. For our previous choice of the parameter values the corresponding eigenvalues for the x, y and z connections are given by:

```
In[9]:= Eigenvalues[
            JCMatrix[Lorenz[x,y,z] /. param,{x,y,z},#]]& /@ {x,y,z}
```

$$Out[9] := \begin{pmatrix} \frac{1}{6}\left(-\sqrt{25-36x^2}-11\right) & \frac{1}{6}\left(\sqrt{25-36x^2}-11\right) \\ -10 & -\frac{8}{3} \\ \frac{1}{2}\left(-\sqrt{2481-40z}-11\right) & \frac{1}{2}\left(\sqrt{2481-40z}-11\right) \end{pmatrix}$$

meaning that the y-connection is self-synchronizing, while for the x and z connections we need to calculate the conditional Lyapunov exponents. The next input computes numerically such exponents for the Lorenz system:

```
In[10]:= NCLyapunovExponents[
             JCMatrix[Lorenz[x,y,z] /. param,{x,y,z},#]]& /@ {x,y,z}
```

$$Out[10] := \begin{pmatrix} -1.81 & -1.86 \\ -10.0 & -2.66667 \\ 0.0108 & -11.01 \end{pmatrix}$$

From this output we can conclude that synchronization will occur for either x or y driving signal. Note also the agreement between the second rows of outputs

9 and 10, as expected from our previous discussion. Our package also includes a command, `Synchronize`, that admits two equal or different dynamical systems, the list of their variables, two initial conditions (one for each system), the time and the integration step, and the variable to be connected and returns the temporal series for the drive and response systems when the Pecora-Carroll (PC) scheme is applied. For example, we already know that (for our choice of parameter values) the connection through variable `x` is self-synchronizing for the Lorenz systems, even if they have different initial conditions. Each system is described by the system equations, the list of its variables and one initial condition:

```
In[11]:={sys1,sys2}={Lorenz[x,y,z] /. param,{x,y,z},#}& /@
                        {{35,-10,10},{12,7,70}};
```

The variables `drivx` and `respx` store the temporal series of the drive and response systems, respectively:

```
In[12]:={drivx,respx}=Synchronize[sys1,sys2,{4,0.001},x];
```

Now, one of the variables of both systems, let us say z, is used to check the synchronization:

```
In[13]:={zdrivx, zrespx}=Map[Last,{drivx,respx},{2}];
```

Then, we load the package:

```
In[14]:=<< Graphics`MultipleListPlot`
```

which provides a convenient way to plot several lists on the same axes. For example, Figure 1(left) shows the temporal series of the z variable for both the drive and response systems, showing that they become eventually synchronized.

```
In[15]:=MultipleListPlot[zdrivx,zrespx,SymbolShape->{None,None},
                         PlotJoined->True,Frame->True,Axes->False]
```

$Out[15] := See\ Figure\ 1(left)$

Another synchronizing connection is given by the variable y, leading to the (x, z) Lorenz subsystem. We can also display the trajectories of the y-driven Lorenz systems from the drive and the response: firstly, we remove a transient of the first 2000 iterates. Then, we join all the trajectory points by lines and finally display the trajectories of the drive and response Lorenz systems for the y-connection (see Figure 1(right)). Note that they are identical, meaning that the synchronization has been attained after that transient of 2000 iterates.

```
In[16]:= {drivy,respy}=Synchronize[sys1,sys2,{6,0.001},y];
         Drop[#,2000]& /@ %;
         Graphics3D[{RGBColor[1,0,0],Line[#]}]& /@ %;
         Show[GraphicsArray[%]]
```

$Out[16] := See\ Figure\ 1(right)$

The same behavior can be obtained for other chaotic systems. For example, in [5] application of both control and synchronization to small arrays of chaotic electronic circuits (namely, Chua's circuit) is described. See also [1, 2, 4].

3 Partial Connections Scheme

The Pecora-Carroll method is specially suited for the case of analog circuits, as one-way connections can be easily implemented through operational amplifiers. However, a problem is that the drive and the response must share a part in common, meaning that the response has, in practice, a reduced dimensionality compared to the drive.

In 1995 an extension of the PC scheme, based on the idea of introducing the driving signal at a given place of the dynamical equations of the response, was proposed [3]. The main difference with the PC method is that in the PC scheme the driving signal is introduced in all possible places, suppressing completely the dynamical evolution of the variable that is homologous to the driving signal in the response dynamical system. In order to understand better the basic idea of the method, it will be illustrated in the case that several Lorenz systems are connected through the variable y. By looking at their equations it is possible that y enters generically at four different places in the evolution equations. One of these possible connections (in bold) is:

$$\left.\begin{array}{l}\dot{x}_1 = \sigma(y_1 - x_1)\\ \dot{y}_1 = x_1(r - z_1) - y_1\\ \dot{z}_1 = -bz_1 + x_1 y_1\end{array}\right\} drive \qquad \left.\begin{array}{l}\dot{x}_2 = \sigma(\mathbf{y_1} - x_2)\\ \dot{y}_2 = x_2(r - z_2) - y_2\\ \dot{z}_2 = -bz_2 + x_2 y_2\end{array}\right\} response \tag{4}$$

In Mathematica, expression (4) can be written as:

```
In[17]:= sys1={x1'[t]==10 (y1[t]-x1[t]),
               y1'[t]==(60-z1[t]) x1[t]-y1[t],
               z1'[t]==x1[t] y1[t]-8/3 z1[t]};
In[18]:= sys2={x2'[t]==10 (y1[t]-x2[t]),
               y2'[t]==(60-z2[t]) x2[t]-y2[t],
               z2'[t]==x2[t] y2[t]-8/3 z2[t]};
```

In this case the corresponding linearized equation for the time evolution of the relative errors between the drive and the response systems is obtained as:

$$\begin{pmatrix}\dot{e}_1\\ \dot{e}_2\\ \dot{e}_3\end{pmatrix} = \begin{pmatrix}-\sigma & \mathbf{0} & 0\\ r - z & -1 & -x\\ y & x & -b\end{pmatrix}\begin{pmatrix}e_1\\ e_2\\ e_3\end{pmatrix} \tag{5}$$

where $(e_1, e_2, e_3) = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ (i.e. the difference between the response and the drive systems in (4)). Note that the square matrix in (5) is analogous to that in *Out[3]* except for the fact that a 0 entry (in bold in (5)) appears at the same place at which the driving signal enters. In other words, if we input the second variable of the drive system into the first equation of the response, one 0 appears at the position (1,2) of the Jacobian matrix.

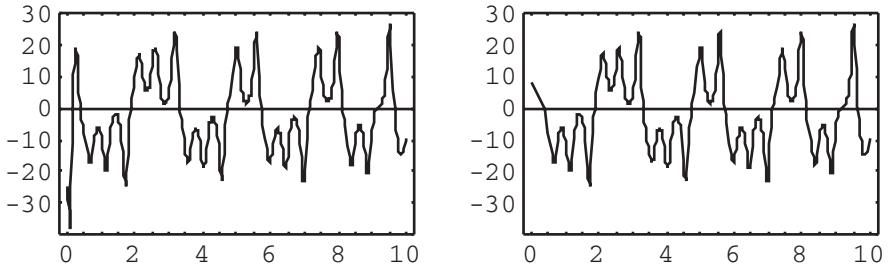


Fig. 2. Synchronization of two Lorenz systems in the variable x with the partial connection given by (4): (left) drive system; (right) response system

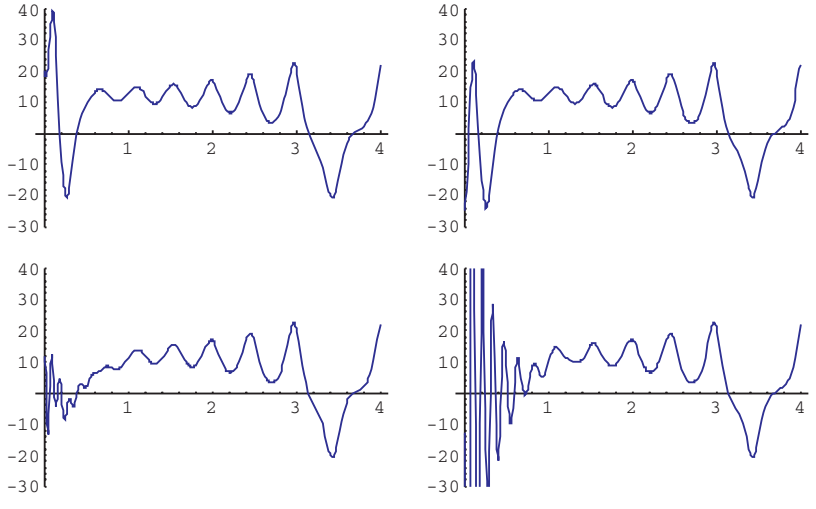


Fig. 3. (left-right,top-bottom) Synchronization of four Lorenz systems with the partial connection scheme: temporal series of variable x

Now, the six-dimensional drive-response couple can be integrated numerically by applying the `NDSolve` command:

```
In[19]:=rl=NDSolve[Union[sys1,sys2,x1[0]==20,y1[0]==-3,z1[0]==-12,
      x2[0]==35,y2[0]==-10,z2[0]==-7],{t,0,25},MaxSteps->10000];
In[20]:=Plot[# /. rl,{t,0,10},Frame->True,PlotRange->{-40,30}]&
                  /@ {x1[t],x2[t]}
```

Out[20] := *See Figure* 2

One of the most interesting potential applications of this scheme is to arrays of chaotic systems, obtaining a highly complex network in which several connections among the different units coexist. For example, we consider two more Lorenz systems connected with the previous ones as:

```
In[21]:= sys3={x3'[t]==10 (y3[t]-x3[t]),
              y3'[t]==60 x2[t]-x3[t] z3[t]-y3[t],
```

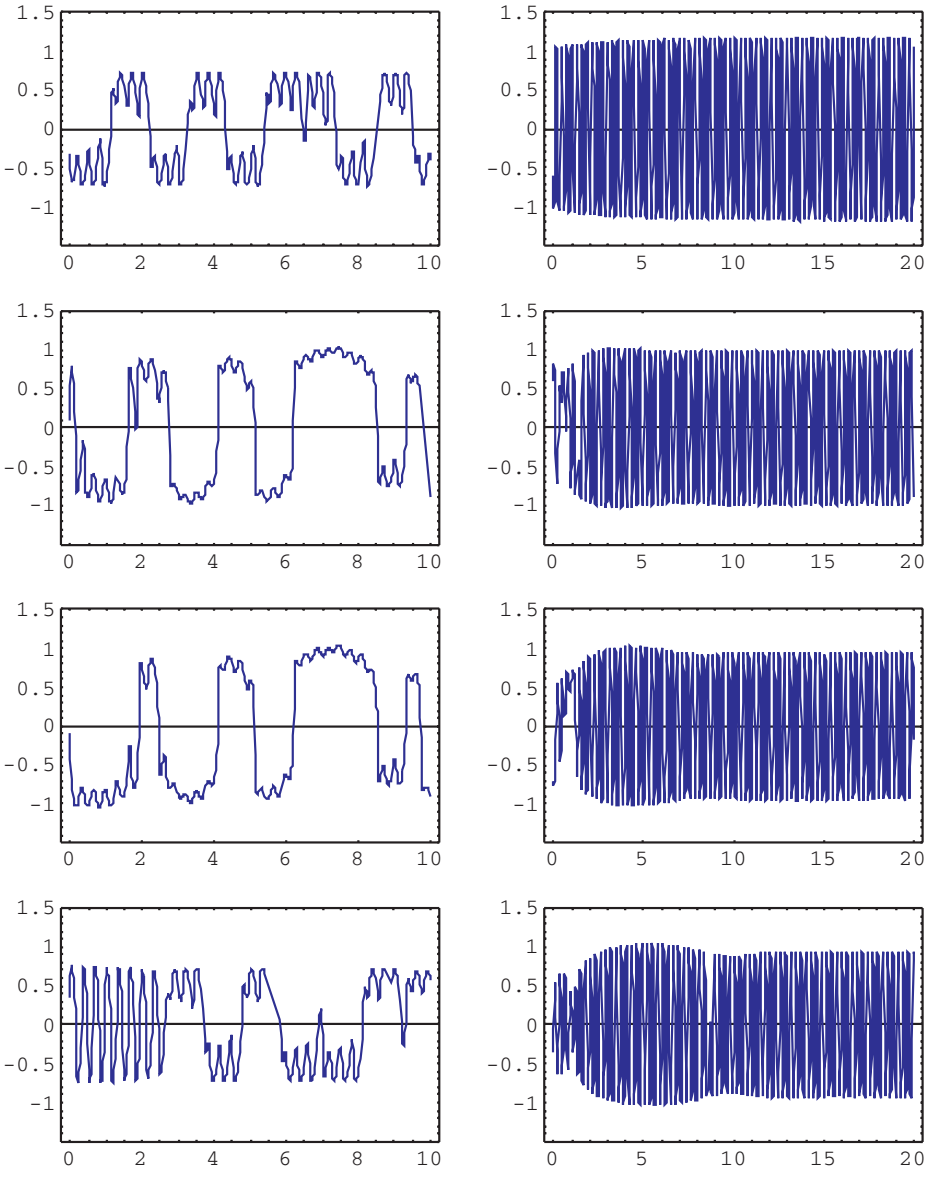


Fig. 4. Two examples of synchronization of four Van der Pol systems with the partial connection scheme: (left) Signals from the first and fourth oscillators are injected into the second and the third systems; (right) The periodic behavior of the first system induces, by inhomogeneous cascading, a periodic behavior in the other three chaotic systems once synchronization is achieved

```
                    z3'[t]==x3[t] y3[t]-8/3 z3[t]};
In[22]:= sys4={x4'[t]==10 (y4[t]-x4[t]),
          y4'[t]==60 x3[t]-x4[t] z4[t]-y4[t],
          z4'[t]==x4[t] y4[t]-8/3 z4[t]};
```

As in the previous example, the synchronization can be analyzed by solving the system of twelve ODEs with the corresponding initial conditions:

```
In[23]:=NDSolve[Union[sys1,sys2,sys3,sys4,x1[0]==20,y1[0]==-3,
          z1[0]==-12,x2[0]==35,y2[0]==-10,z2[0]==-7,x3[0]==12,
          y3[0]==30,z3[0]==600,x4[0]==-13,y4[0]==-210,
          z4[0]==-400],{t,0,4},MaxSteps->10000];
In[24]:=Plot[# /. %,{t,0,4},PlotRange->{-30,40}]&
            /@ {x1[t],x2[t],x3[t],x4[t]};
In[25]:=Show[GraphicsArray[Partition[%, 2, 2]]]
```

$Out[25] :=$ *See Figure* 3

The last two examples of synchronization of arrays of systems are displayed in Figure 4. The corresponding Mathematica codes are very similar to those of the previous examples and hence, are not included here because of limitations of space. The example on the left shows four Van der Pol-Duffing (VPD) oscillators connected in such a way that the first and the fourth systems evolves independently, whereas the second and the third oscillators are driven by the other ones. In this kind of competitive connection in which one oscillator is simultaneously driven by several different signals the observed behavior is a sort of compromise between the inputs that cannot be predicted a priori, but it appears that each connection has a particular strength and, hence, some of them may have the ability to dominate the others. Note also that by using the original PC method it would not be possible to establish this kind of connection for a three-dimensional circuit, because by injecting two signals one would have a one-dimensional, and hence trivial, circuit.

Another remarkable feature of synchronization is given by networks in which one of the oscillators has different values of parameters (inhomogeneous driving). In particular, we can consider the situation in which one of the oscillators is in a periodic window or on the road to chaos. For example, we consider the case of a cascade of four VPD elements with undirectional coupling from the first oscillator, which has been chosen such that it is in a periodic orbit, while the rest of the systems of the network are in a chaotic regime. As shown in Figure 4(right), this new driving scheme induces a transition in the last three chaotic oscillators from the strange attractor to the coexisting stable cycle limit. This transition comes from the driving signal of the first system, that is in a periodic regime itself.

This analysis of the behavior of the network in situations of inhomogeneous driving is useful to determine the response of the network in the case where some external input acts on some oscillator changing its state. For example, a behavior similar to that of Figure 4(right) can be simply obtained by adding some external noise specially chosen to induce the transition from the strange attractor to the cycle limit. This indicates that changes in the parameter values might produce a signal that is equivalent to a noisy signal.

References

1. Chua, L.O., Kocarev, L., Eckert, K., Itoh, M.: Experimental chaos synchronization in Chua's circuit, Int. J. of Bifurcation and Chaos **2** (1992) 705-708
2. Cuomo, K.M., Oppenheim, A.V.: Circuit implementation of synchronized chaos with applications to communications, Phys. Rev. Lett. **71** (1993) 65-68
3. J. Güémez, J., Matías, M.A.: Modified method for synchronizing and cascading chaotic systems, Phys. Rev. E **52** (1995) 2145-2148
4. Kocarev, L.J., Halle, K.S., Eckert, K., Chua, L.O., Parlitz, U.: Experimental demonstration of secure communications via chaotic synchronization. Int. J. of Bifurcation and Chaos **2** (1992) 709-713
5. Iglesias, A.: Synchronization and control in small networks of chaotic electronic circuits. In: Villacampa, Y., Carlomagno, G.M., Brebbia, C.A. (Eds.): Computational Methods and Experimental Measurements, X. WIT Press/Computational Mechanics Publications, Southampton-Boston (Series: Computational Engineering) Vol. 3 (2001) 799-808
6. Maeder, R.: Programming in Mathematica, Second Edition, Addison-Wesley, Redwood City, CA (1991)
7. Pecora, L.M., Carroll, T.L.: Synchronization in chaotic systems, Phys. Rev. Lett. **64** (1990) 821-823
8. Pecora, L.M., Carroll, T.L.: Driving systems with chaotic signals, Phys. Rev. A, **44** (1991) 2374-2383
9. Van der Pol, B.: Theory of the amplitude of free and forced triode vibration. Radio Rev. **1** (1920) 701-710
10. Wolfram, S.: The Mathematica Book, Fourth Edition, Wolfram Media, Champaign, IL & Cambridge University Press, Cambridge (1999)

Analyzing the Synchronization of Chaotic Dynamical Systems with Mathematica: Part II

A. Iglesias* and A. Gálvez

Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n, E-39005, Santander, Spain
iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. This work concerns the synchronization of chaotic dynamical systems by using the program Mathematica and follows up a previous one (also included in this volume) devoted to the same topic [2]. In particular, this second paper classifies and illustrates the wide range of chaotic synchronization phenomena. Some examples of each kind are briefly analyzed in this paper with the help of a Mathematica package already introduced in the previous paper.

1 Introduction

This paper is the second one of a series of two papers devoted to analyze the synchronization of chaotic dynamical systems by using the program Mathematica. In the first paper, the main concepts of chaotic synchronization as well as two popular approaches to this problem (Pecora-Carroll (PC) and partial connections) have been analyzed. This second paper is devoted to classify the wide range of chaotic synchronization phenomena described in the literature. Some examples of each kind of phenomena are briefly analyzed in this paper with the help of the Mathematica package already introduced in the previous paper.

The structure of the paper is as follows: in Section 2 we introduce some mathematical background. Then, Section 3 analyzes the cases of monotonic, oscillating and marginal synchronization, which are briefly described with the help of some illustrative examples carried out in Mathematica.

2 Mathematical Background

During the last few years, a number of attempts to classify the vaste range of synchronization phenomena have been published. Among them, those based on the analysis of the conditional Lyapunov exponents (see the previous paper for a definition) are gaining more and more popularity. For instance, the authors in [1] analyzed the case of a couple master-slave of identical chaotic systems.

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 482–491, 2005.

In particular, they considered only those systems connected by either of the two chaotic synchronization schemes (namely, the Pecora-Carroll and the partial connections) described in our previous paper. This makes this classification especially suitable for our purposes. Due to this reason, this is the classification considered in this paper. We also remark that this discussion is restricted to three-dimensional dissipative dynamical systems driven by a one-dimensional signal. This implies that the dimension of the synchronization manifold will be 2 in the case of the PC method and 3 in the case of the partial connection method.

This classification is based on the linear stability analysis of the synchronization error system $\delta\dot{\mathbf{x}} = \dot{\mathbf{x}}_{drive} - \dot{\mathbf{x}}_{response}$, which can be written as:

$$\delta\dot{\mathbf{x}} = \mathbf{Z}.\delta\mathbf{x} + O((\delta\mathbf{x})^2) \tag{1}$$

where the solution of this expression is given by:

$$\delta\mathbf{x}(t) = \delta\mathbf{x}(0)exp(t\,\mathbf{Z}) \tag{2}$$

where $\mathbf{Z}$ is the evolution matrix. Note that in the case that the two systems are not connected, this matrix $\mathbf{Z}$ has some time-dependent coefficient having a very complicated form (as it represents the chaotic signal) and, hence, the asymptotic behavior of the system is given by the Lyapunov spectra. In addition,there is a (approximate) procedure of getting more information about the system: it consists of replacing the instantaneous values of the corresponding chaotic signals by their corresponding averaged values. In the case of the PC method, $\mathbf{Z}$ is a two-dimensional matrix:

$$\mathbf{Z} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3}$$

with two eigenvalues

$$\lambda_{1,2} = \frac{1}{2}\left[(a+d) \pm \sqrt{(a-d)^2 + 4bc}\right] \tag{4}$$

The analysis of this expression indicates that one can safely assure that the two eigenvalues are real only if $(a-d) > 4bc$ as the expression will be linear in $(a-d)$. In other cases, the replacement by the mean values will probably work well if the time-dependent coefficient appears in a or d. The analysis of the eigenvalues is more complicated in the case of the modified method of the partial connections and the interested reader is referred to [1] for a more detailed discussion.

In the following section we will describe many of the phenomena associated with the synchronization. Roughly speaking, the imaginary part of the eigenvalues informs about the approximation to the synchronized state (steady or oscillatory) whereas the sign of the real part informs about the following behaviors: synchronization (negative), marginal synchronization (zero) and nonsynchronization. This classification will be used in the next paragraphs.

3 Synchronization Cases

The synchronization state is characterized by the fact that the real part of the eigenvalues of **Z** is negative. In addition, several subcases can be defined depending on the imaginary part of these eigenvalues: if all the eigenvalues are real, the approach to the synchronized state will be monotonic, while it will be oscillatory if there are two complex-conjugate eigenvalues.

3.1 Monotonic Synchronization

This case can be illustrated by considering two Lorenz systems with different initial conditions both leading to chaotic behavior:

```
In[1]:=lor1={x1'[t]==10 (y1[t]-x1[t]),y1'[t]==28 x1[t]-z1[t] x1[t]-
             y1[t],z1'[t]==x1[t] y1[t]-8/3 z1[t],x1[0]==20,
             y1[0]==-3,z1[0]==-12};
In[2]:=lor2={x2'[t]==10 (y2[t]-x2[t]),y2'[t]==28 x2[t]-z2[t] x2[t]-
             y2[t],z2'[t]==x2[t] y2[t]-8/3 z2[t],x2[0]==35,
             y2[0]==-10,z2[0]==-7};
```

Now, we will consider the connection through the variable y by employing the PC method. In this case, the system has two real eigenvalues $\lambda_1 = -\frac{8}{3}, \lambda_2 = -10$ (see our previous paper [2] for details) implying synchronization. Moreover, the difference between the variables decreases as $\delta\mathbf{x}(t) = \delta\mathbf{x}(0)exp(-|\lambda|t)$ as shown in Figure 1. In order to analyze the synchronization phenomena, we have defined the `PlotSynchronizationFlow` command. Its syntax is:

`PlotSynchronizationFlow[`$\{sys_1, sys_2\}$`,`var_c`,`var_d`,`$\{t, t_{ini}, t_{end}, t_{app}\}$`,`$opts$`]`

where $\{sys_1, sys_2\}$ are the master and slave systems respectively, var_c is the variable to be connected between them, var_d is the variable to be displayed, t means the time, which is assumed to evolve between t_{ini} and t_{end}. The additional temporal value, t_{app}, represents an intermediate time at which synchronization algorithm is applied. It is intended to show the comparison between the independent and the synchronized states of both the master and the slave systems. This time will be indicated by the dashed line in the figures throughout this paper. Finally, *opts* allows the users to input some graphical options. For example:

```
In[3]:= PlotSynchronizationFlow[{lor1,lor2},y,x,{t,0,20,5},
          PlotRange->{-20,20}]
```

$Out[3] :=$ *See Figure* 1

Figure 1 displays the temporal series of the x variable for the master, the slave and the difference between them, respectively. Note the exponential decay of this difference to zero once the synchronization method is applied. Similar behavior is obtained for the variable z.

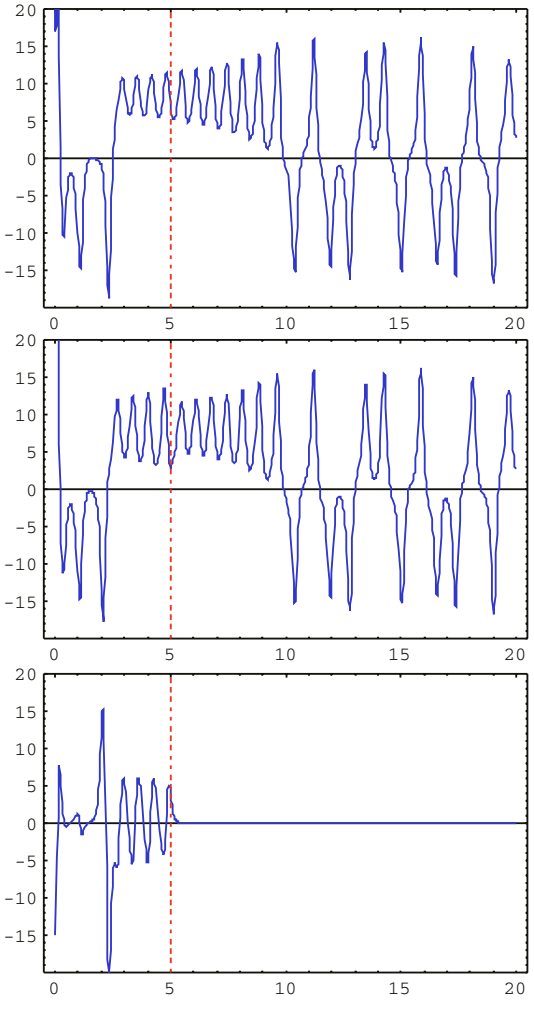


Fig. 1. Monotonic synchronization: two Lorenz systems are connected through the y variable. The figure shows the temporal series of the x variable for the drive and response systems and the difference between them

3.2 Oscillating Synchronization

An example of oscillating behavior is given by two Chua's circuits with parameters:

```
In[4]:=f[x_]:=-0.68 x-0.295(Abs[x+1]-Abs[x-1]);
In[5]:=chua1={x1'[t]==10 (y1[t]-x1[t]-f[x1[t]]),y1'[t]==x1[t]-
              y1[t]+z1[t],z1'[t]==-14.87 y1[t],x1[0]==-0.4,
              y1[0]==0.2,z1[0]==0.1};
```

```
In[6]:=chua2={x2'[t]==10 (y2[t]-x2[t]-f[x2[t]]),y2'[t]==x2[t]-
          y2[t]+z2[t],z2'[t]==-14.87 y2[t],x2[0]==0.3,
          y2[0]==-0.2,z2[0]==-0.7};
```

which, for the PC connection with the x variable, is characterized by a pair ot complex-conjugate eigenvalues:

```
In[7]:= JCMatrix[chua1,{x1[t],y1[t],z1[t]},x1[t]]
```

$$Out[7] := \begin{pmatrix} -1 & 1 \\ -14.87 & 0 \end{pmatrix}$$

```
In[8]:= Eigenvalues[%]
```

$$Out[8] := \{-0.5 + 3.82361\,i, -0.5 - 3.82361\,i\}$$

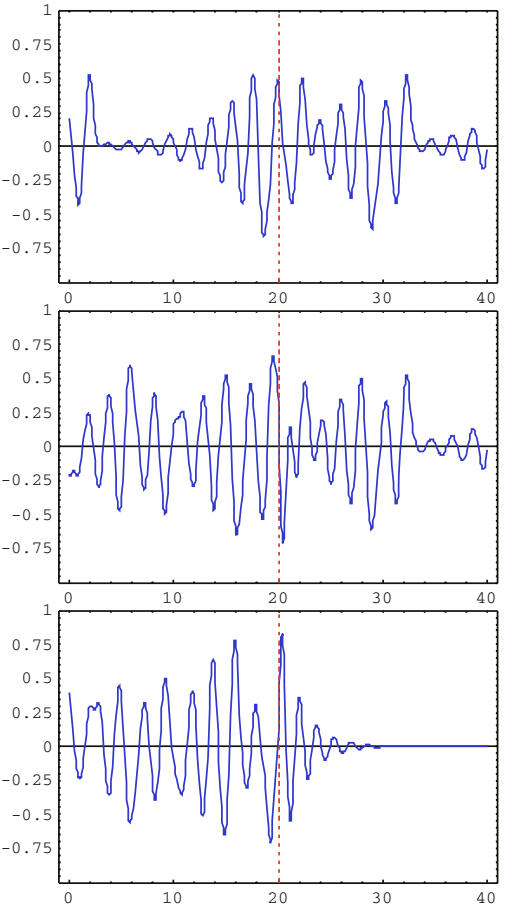


Fig. 2. Oscillating synchronization: two Chua's circuits are connected through the x variable. The figure shows the temporal series of the y variable for the drive and response systems and the difference between them

Therefore, the corresponding conditional Lyapunov exponents for this connection are $(-0.5, -0.5)$ implying synchronization. However, these complex-conjugate eigenvalues induce the oscillatory behavior shown in Figure 2 for the variable y. The same behavior also occurs for the variable z.

```
In[9]:= PlotSynchronizationFlow[{chua1,chua2},x,y,{t,0,40,20},
          PlotRange->{-1,1}]
```

$Out[9] :=$ *See Figure* 2

3.3 Marginal Synchronization

This situation is characterized by the fact that the highest conditional Lyapunov exponent is zero, while the rest of the exponents might be zero or negative.

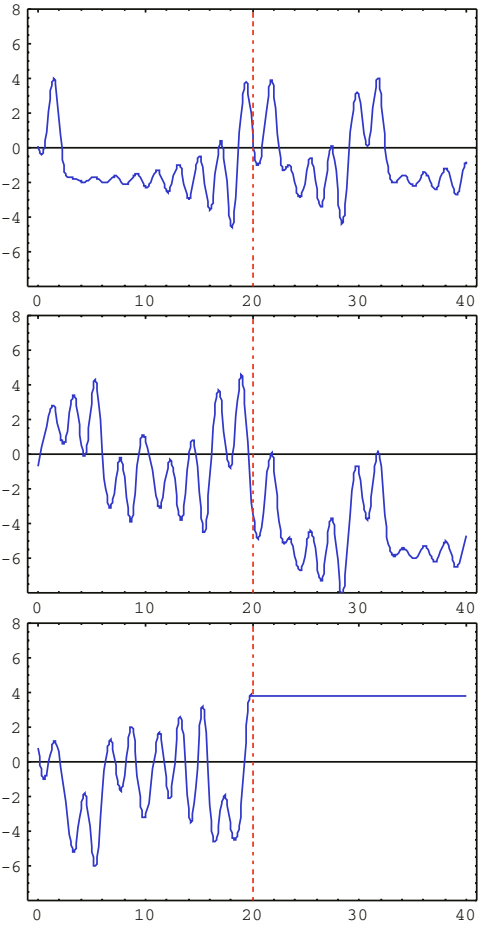


Fig. 3. Marginal constant synchronization: two Chua's circuits are connected through the y variable. The figure shows the temporal series of the z variable for the drive and response systems and the difference between them

Moreover, some qualitatively different behaviors may arise depending on whether this null Lyapunov exponent is zero at a local level (the corresponding **Z** has a zero eigenvalue) or only in average. The behavior can also change if more than one Lyapunov exponents are null.

Marginal Constant Synchronization

In this case, there is a zero Lyapunov exponent while the another one is strictly negative. The observed behavior is that one of the variables in the response system becomes synchronized with the drive, while the other also synchronizes with the drive but with a constant separation. The connection does not have the ability to reduce the initial separation at the time of connection (in other words, $\delta\mathbf{x}(t) = \delta\mathbf{x}(0)$). An example of this behavior is given by Chua's circuit for the y connection. The corresponding eigenvalues of the matrix **Z** are:

```
In[10]:= (m=JCMatrix[chua1,{x1[t],y1[t],z1[t]},y1[t]])//Eigenvalues
```

$Out[10] := \{10[-f'(x_1(t)) - 1], 0\}$

meaning that one eigenvalue depends on the variable x_1. The Lyapunov exponents for this case are:

```
In[11]:= NCLyapunovExponents[m]
```

$Out[11] := \{0, -2.488\}$

In Figure 3 we can see the evolution of the variable z.

Marginal Sized Synchronization

Other possibility when the system has a single zero conditional Lyapunov exponent is that, while the system takes this value in average, at the local level it is alternatively positive and negative. In this case the observed behavior is that the response system exhibits the same qualitative behavior as the drive, but with different size (and sometimes with different symmetry). This different size is related to the differences in the variables referred to the time at which the connection starts. An example of this behavior is given by the Lorenz system (already described in Sec. 3.1) for the z connection, whose eigenvalues are:

```
In[12]:= (n=JCMatrix[lor1,{x1[t],y1[t],z1[t]},z1[t]])//Eigenvalues
```

$$Out[12] := \left\{\frac{1}{2}\left(-\sqrt{1201 - 40\,z_1(t)} - 11\right), \frac{1}{2}\left(\sqrt{1201 - 40\,z_1(t)} - 11\right)\right\}$$

meaning that the eigenvalues depend on the variable z_1. The Lyapunov exponents for this case are:

```
In[13]:= NCLyapunovExponents[%]
```

$Out[13] := \{0, -11\}$

In Figure 4 we can see the evolution of the y variable of the drive, the response and the difference between them. The same behavior can be found for the variable x. The observed behavior is that the signals for both the drive and the response are very similar but symmetric. Therefore, the difference is oscillatory. Note also

that this difference does not oscillate around the origin, but it is slightly shifted from it.

Marginal Oscillatory Synchronization

This last case of marginal synchronization is characterized by the fact that **Z** has a pair of complex-conjugate eigenvalues with zero real part. As a consequence, the difference between the master and the slave changes in an oscillatory way with constant amplitude related to the differences in the variables at the time in which the connection starts. On the contrary, the frequency will depend on the

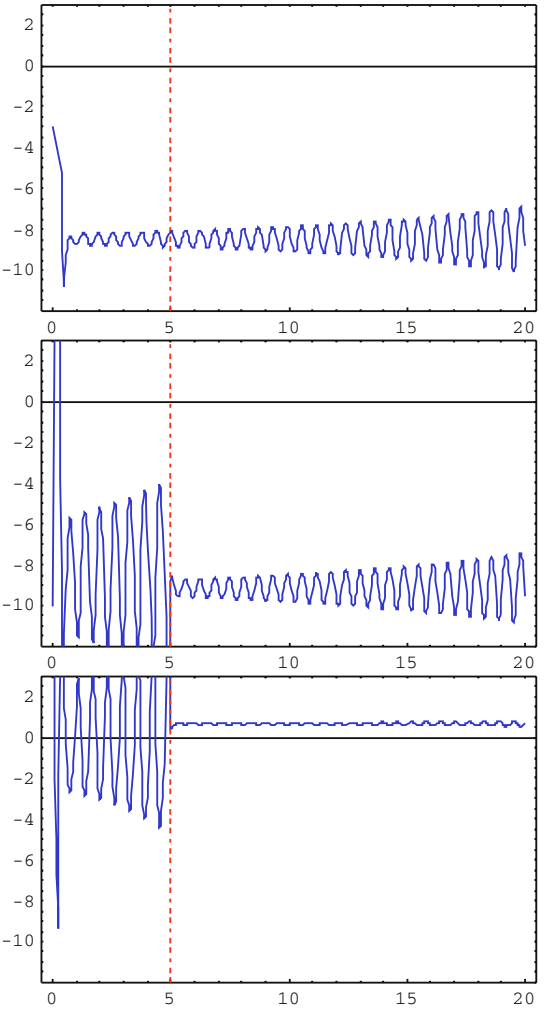


Fig. 4. Marginal sized synchronization: two Lorenz systems are connected through the z variable. The figure shows the temporal series of the x variable for the drive and response systems and the difference between them

imaginary part of the eigenvalues of **Z**. As an example, we consider here the I Sprott system [3] connected by means of the partial connection scheme described in the previous paper. In this case, we insert the connecting variable z_1 into the second equation of the response:

```
In[14]:= spi1={x1'[t]==-0.2 y1[t],y1'[t]==x1[t]+z1[t],z1'[t]==x1[t]-
          z1[t]+y1[t]^2,x1[0]==-0.6,y1[0]==-0.8,z1[0]==0.9};
In[15]:= spi2={x2'[t]==-0.2 y2[t],y2'[t]==x2[t]+z1[t],z2'[t]==x2[t]-
          z2[t]+y2[t]^2,x2[0]==0.6,y2[0]==0.3,z2[0]==-0.2};
```

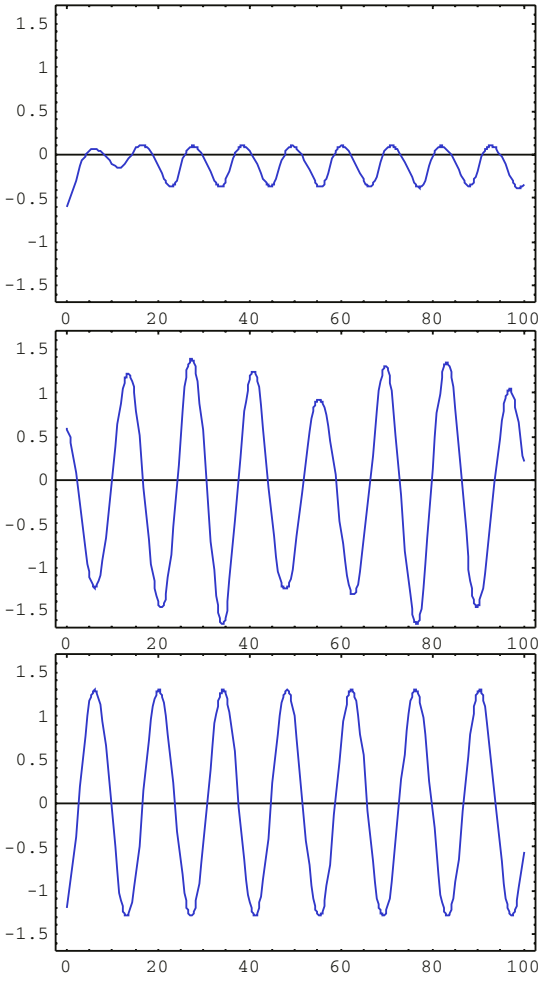


Fig. 5. Marginal oscillatory synchronization of two Sprott systems I coupled by using the modified partial connection method through the z variable into the second equation of the response. Time evolution for x variable of (top-bottom): the drive, the response and the difference between them

The eigenvalues of the resulting Jacobian matrix are:

```
In[16]:= JCMatrix[{spi1,spi2},{x1,y1,z1,x2,y2,z2},t] //Eigenvalues
```

$Out[16] := \{-1, -0.447214\,i, 0.447214\,i\}$

The two last eigenvalues are complex-conjugate with zero real part. This implies that two conditional Lyapunov exponents are zero, both locally and asymptotically. An example of this behavior can be found in Figure 5.

Generalized Marginal Synchronization

The notion of marginal synchronization has been recently extended to a new concept, the *generalized marginal synchronization* (GMS), in the sense that allowing the response system to be different from the drive system [4]. The drive and the response show GMS if there is a function which transforms points of the drive system attractor into points on the response system attractor. Once again, this function is independent on time, but now GMS is known up to a certain constant which depends on the initial conditions of both the drive and the response systems. Finally, GSM can be applied not only to unidirectionally but also to mutually coupled systems [4].

References

1. Guémez, J., Martín, C., Matías, M.A.: Approach to the synchronized state of some driving methods. Physical Review E **55** (1997) 124-134
2. Iglesias, A., Gálvez, A.: Analyzing the Synchronization of Chaotic Dynamical Systems with Mathematica: Part I. ICCSA - Computational Science. Springer-Verlag, Lecture Notes in Computer Science (2005) (this volume)
3. Sprott, J.C.: Some simple chaotic flows. Phys. Rev. E. **50** (1994) 647-650
4. Krawiecki, A., Matyjaskiewicz, S.: Generalizations of the concept of marginal synchronization of chaos. Chaos, Solitons and Fractals **11**(9) (2000) 1445-1458

A Mathematica Package for Computing and Visualizing the Gauss Map of Surfaces

R. Ipanaqué[1] and A. Iglesias[2,3,⋆]

[1] Department of Mathematics, National University of Piura,
Urb. Miraflores s/n, Castilla, Piura, Perú
robertchero@hotmail.com
[2] Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n, E-39005, Santander, Spain
[3] Laboratory of Advanced Research, University of Tsukuba, Tsukuba, Japan
iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. One of the most interesting and striking concepts in Differential Geometry is that of the *Gauss map*. In the case of surfaces, this map projects surface normals to a unit sphere. This strategy is especially useful when analyzing the shape structure of a smooth surface. This paper describes a new Mathematica package, `GaussMap`, for computing and displaying the tangent and normal vector fields and the Gauss map of surfaces described symbolically in either implicit or parametric form. The performance of the package is discussed by means of several illustrative and interesting examples. The package presented here can be applied for visualizing and studying the geometry of a surface under analysis, thus providing the users with an excellent computer tool for teaching and visualization purposes.

1 Introduction

One of the most interesting and striking tools in Differential Geometry is the so-called *Gauss map*. Although this concept can be defined for hypersurfaces of any dimension [5], the most interesting applications appear for the 2-dimensional (curves) and, especially, for the 3-dimensional (surfaces) cases. In particular, in this paper we restrict ourselves to the case of smooth (differentiable) surfaces.

Roughly speaking, the Gauss map projects surface normals to a unit sphere, providing a powerful visualization of the geometry of a graphical object. On the other hand, dynamic visualization of the Gauss map speeds understanding of complex surface properties. Therefore, it can be efficiently used to illuminate the natural structure of surface shape [3]. It can also be applied to predict visual events caused by changes in lighting, shading, and camera control. More details about applications of the Gauss map to computer graphics and visualization can be found in [2, 3, 6] and references therein.

⋆ Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 492–501, 2005.

This paper presents a new Mathematica package, `GaussMap`, for computing and displaying the tangent and normal vector fields and the Gauss map of surfaces described symbolically in either implicit or parametric form. The output obtained is consistent with Mathematica's notation and results. The performance of the package is discussed by means of several illustrative and interesting examples. They show that the package presented here can be applied for visualizing and studying the geometry of a surface under analysis, thus providing the users with an excellent computer tool for teaching and visualization purposes.

The structure of this paper is as follows: Section 2 introduced some mathematical definitions about the normal vector field and the Gauss map of implicit and parametric surfaces to be used throughout the paper. For the sake of illustration, two theoretical examples of the computation of the Gauss map for simple implicit surfaces are also briefly described in this section. Then, Section 3 introduces the new Mathematica package, `GaussMap`, and describes the commands implemented within. The performance of the package is also discussed in this section by using some illustrative examples for implicit and parametric surfaces. Finally, Section 4 closes with the main conclusions of this paper and some further remarks.

2 Mathematical Preliminaries

Let M be an implicit surface in $\mathbb{E}^3$ defined as $\boldsymbol{\phi}(x, y, z) = 0$. A *normal vector field* on M is given by the gradient vector field of $\boldsymbol{\phi}$ as:

$$\boldsymbol{\nabla\phi} = \left(\frac{\partial\boldsymbol{\phi}}{\partial x}, \frac{\partial\boldsymbol{\phi}}{\partial y}, \frac{\partial\boldsymbol{\phi}}{\partial z}\right). \tag{1}$$

The *unit normal vector field*, $\boldsymbol{U}$, on M is simply defined from Eq. (1) as

$$\boldsymbol{U} = \sum_{i=1}^{3} n_i \boldsymbol{U}_i = (n_1, n_2, n_3) = \frac{\boldsymbol{\nabla\phi}}{||\boldsymbol{\nabla\phi}||}. \tag{2}$$

where $\{\boldsymbol{U}_i\}_{i=1,2,3}$ form the natural frame for the tangent space of M.

These definitions can be trivially extended to the parametric case. Let $\boldsymbol{S}$ be a local parameterization of a surface M in $\mathbb{E}^3$ given by

$$\boldsymbol{S}(u, v) = (x(u, v), y(u, v), z(u, v)) \tag{3}$$

The *normal vector field* on M is then given by $\boldsymbol{S}_u(u, v) \times \boldsymbol{S}_v(u, v)$. Similarly, the unit normal vector field, $\boldsymbol{U}$, is described by

$$\boldsymbol{U} = (n_1, n_2, n_3) = \frac{\boldsymbol{S}_u(u, v) \times \boldsymbol{S}_v(u, v)}{||\boldsymbol{S}_u(u, v) \times \boldsymbol{S}_v(u, v)||}. \tag{4}$$

Given an oriented surface M in the Euclidean 3-space with unit normal vector field $\boldsymbol{U}$, its *Gauss map, G,* is a transformation that associates to each point

$\boldsymbol{p} \in M$ its unit normal vector $\boldsymbol{U}(\boldsymbol{p})$, considered as a point on the unit two-sphere $\sum = S^2 \subset \mathbb{R}^3$. In mathematical terms, G is a map $G : M \to \sum$ defined by:

$$G(\boldsymbol{p}) = (n_1(\boldsymbol{p}), n_2(\boldsymbol{p}), n_3(\boldsymbol{p})) \tag{5}$$

where $\boldsymbol{p} \in M$ is a point on the surface M. From a graphical point of view, the Gauss map can be thought of as moving the normal of a point on a surface to a congruent parallel vector at the origin of the unit sphere (see Figure 1).

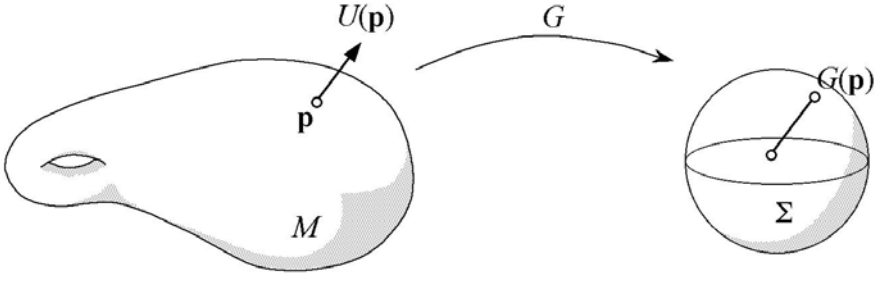


Fig. 1. Graphical interpretation of the Gauss map of a point $\boldsymbol{p}$ on a surface M onto the unit two-sphere $\sum = S^2$

We point out that the Gauss map can be defined (globally) if and only if the surface is orientable, but it always can be defined locally (i.e. on a small piece of the surface). We also remark that the Jacobian of the Gauss map is equal to Gauss curvature, and that the differential of the Gauss map is called shape operator [1].

Example 1: *M is the cylinder of radius r.* Its implicit equation is $x^2 + y^2 = r^2$ or, equivalently, $\boldsymbol{\phi}(x, y, z) = x^2 + y^2 - r^2 = 0$. The gradient function is $\boldsymbol{\nabla\phi} = (2x, 2y, 0)$. Hence, $\boldsymbol{U} = \dfrac{1}{r}(x, y, 0)$ is the unit normal vector field on M. In other words, given a point $\boldsymbol{p} = (p_1, p_2, p_3)$, $G(\boldsymbol{p}) = \dfrac{1}{r}(p_1, p_2, 0)$. On the other hand, the coordinates of this vector hold:

$$\left(\frac{p_1}{r}\right)^2 + \left(\frac{p_2}{r}\right)^2 = \frac{p_1^2 + p_2^2}{r^2} = \frac{r^2}{r^2} = 1$$

This means that the Gauss map image of the cylinder $x^2 + y^2 = r^2$ onto the unit sphere is the unit circle $x^2 + y^2 = 1$, $z = 0$. ∎

Example 2: *M is the paraboloid $z = x^2 + y^2$.* In this case, $\boldsymbol{\phi}(x, y, z) = x^2 + y^2 - z = 0$. The gradient function is $\boldsymbol{\nabla\phi} = (2x, 2y, -1)$. Hence, $\boldsymbol{U} = \dfrac{1}{\sqrt{4x^2 + 4y^2 + 1}}(2x, 2y, -1)$ is the unit normal vector field on M. In other words, given a point $\boldsymbol{p} = (p_1, p_2, p_3)$, $G(\boldsymbol{p}) = \dfrac{1}{\sqrt{4p_1^2 + 4p_2^2 + 1}}(2p_1, 2p_2, -1)$. Because

the third coordinate is negative, the image will lie onto the south hemisphere of $\sum$ only. An additional analysis is required in order to know whether or not such an image actually fills the whole south hemisphere. To this aim, we consider straight lines $p_2 = mp_1$, $m \in \mathbb{R}$ and compute the limit:

$$\lim_{p_1 \to \infty} \frac{1}{\sqrt{4p_1^2 + 4m^2p_1^2 + 1}}(2p_1, 2mp_1, -1) = \frac{1}{\sqrt{4 + 4m^2 + 1}}(2, 2m, 0)$$

This means that for points far away from the origin, their image onto the sphere is given by those points whose third coordinate vanishes and whose two first coordinates satisfy:

$$\left(\frac{2}{\sqrt{4 + 4m^2}}\right)^2 + \left(\frac{2m}{\sqrt{4 + 4m^2}}\right)^2 = \frac{4 + 4m^2}{4 + 4m^2} = 1$$

In other words, the Gauss map image of $z = x^2 + y^2$ onto the unit sphere $\sum$ is the set of points (x, y, z) such that $x^2 + y^2 + z^2 = 1$, with $z < 0$. ∎

Note that similar considerations can be carried out for many other examples, which are not included here because of limitations of space. The interested reader is referred to [5] for a detailed analysis on this interesting question.

3 The Package `GaussMap`: Some Illustrative Examples

This section describes some examples of the application of this package to the cases of implicit and parametric surfaces. Firstly, we load the package:

```
In[1]:=<<DifferentialGeometry‘GaussMap‘
```

3.1 Implicit Surfaces

The first example is given by the paraboloid $z = x^2 + y^2$, already analyzed, from a theoretical point of view, in Example 2 of the previous section. The `ImplicitNormalField` command calculates the normal vector field of any implicit surface in the form $f(x, y, z) = 0$ and returns a graphical output comprised of the surface and such a normal vector field, as displayed in Figure 2:

```
In[2]:=ImplicitNormalField[x^ 2+y^ 2-z,{x,-2,2},{y,-2,2},{z,-1,2},
     Surface->True, PlotPointsSurface->{4,7},VectorHead->Polygon]
```

Out[2] := *See Figure* 2

As shown in the previous section, the Gauss map image of this surface is given by $x^2 + y^2 + z^2 = 1$, $z < 0$. The `ImplicitGaussMap` command of an implicit surface returns the original surface and its Gauss map on the unit sphere $\sum$:

```
In[3]:=ImplicitGaussMap[x^ 2+y^ 2-z,{x,-2,2},{y,-2,2},{z,-1,2},
         PlotPoints->{4,7}]
```

Out[3] := *See Figure* 3

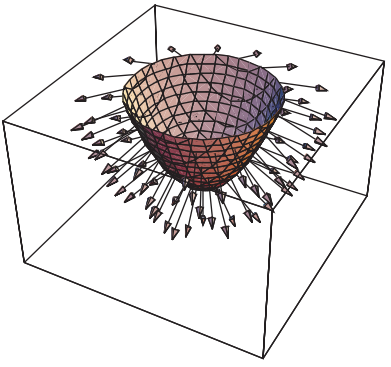


Fig. 2. The paraboloid $z = x^2 + y^2$ and its vector normal field

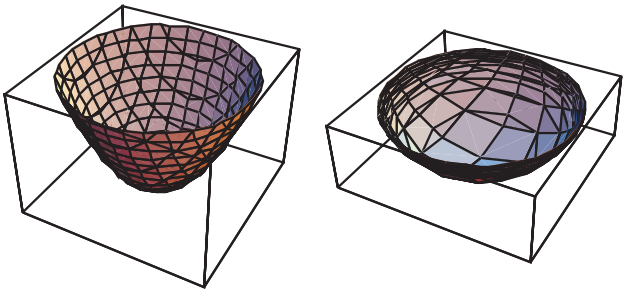


Fig. 3. Applying the Gauss map to the implicit surface $z = x^2 + y^2$: (left) original surface; (right) its Gauss map image onto the unit sphere

The next example calculates the normal vector field of the surface $z = x^2 - y^2$:

```
In[4]:=ImplicitNormalField[z-x^2+y^2,{x,-2,2},{y,-2,2},{z,-2,2},
    Surface->True, PlotPointsSurface->{4,6},VectorHead->Polygon,
    PlotPointsNormalField->{3,4}]
```

Out[4] := *See Figure* 4

Now, the Gauss map image of such a surface is obtained as:

```
In[5]:=ImplicitGaussMap[z-x^2+y^2,{x,-2,2},{y,-2,2},{z,-2,2},
      PlotPoints->{4,6}]
```

Out[5] := *See Figure* 5

A more complicated example is given by the double torus surface, defined implicitly by $\left(\sqrt{\left(\sqrt{x^2+y^2}-2\right)^2+y^2}-2\right)^2+z^2-1=0$ and whose normal vector field is shown in Figure 6:

```
In[6]:=ImplicitNormalField[(Sqrt[(Sqrt[x^2+y^2]-2)^2+y^2]-2)^2+
    z^2-1,{x,-6,6},{y,-4,4},{z,-1,1},Surface->True,
    PlotPointsSurface->{5,5},VectorHead->Polygon]
```

Out[6] := *See Figure* 6

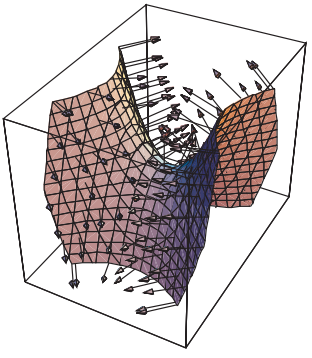


Fig. 4. The surface $z = x^2 - y^2$ and its vector normal field

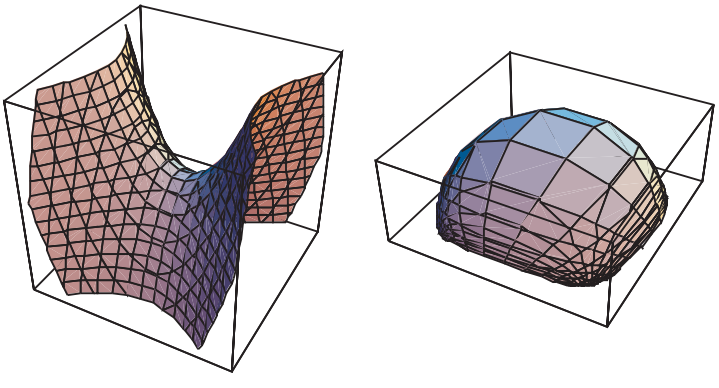


Fig. 5. Applying the Gauss map to the implicit surface $z = x^2 - y^2$: (left) original surface; (right) its Gauss map image onto the unit sphere

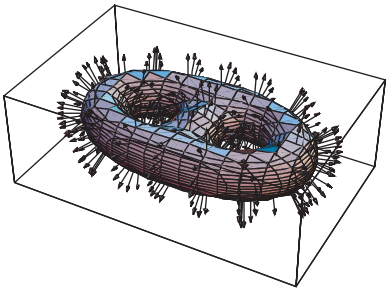


Fig. 6. The double torus surface and its vector normal field

```
In[7]:=ImplicitGaussMap[(Sqrt[(Sqrt[x^2+y^2]-2)^2+y^2]-2)^2+
      z^2-1,{x,-6,6},{y,-4,4},{z,-1,1},PlotPoints->{5,5}]
```

Out[7] := *See Figure* 7

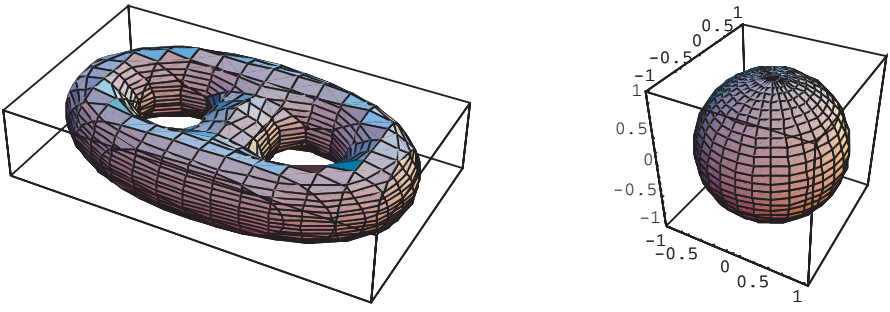


Fig. 7. Applying the Gauss map to the double torus surface: (left) original surface; (right) its Gauss map image onto the unit sphere

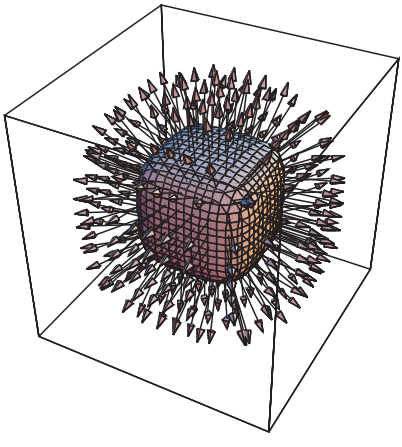


Fig. 8. The surface $x^4 + y^4 + z^4 = 1$ and its vector normal field

It is interesting to remark that, because the Gauss map is a continuous (in fact, differentiable) function, it is closed for compact sets, i.e. it transforms compact sets into compact sets. Since the double torus is compact, its image is actually the unit sphere. Another example of this remarkable property is given by the Lamé surface of fourth degree $x^4 + y^4 + z^4 = 1$:

```
In[8]:=ImplicitNormalField[x^4+y^4+z^4-1,{x,-1,1},{y,-1,1},{z,-1,1},
    Surface->True, PlotPointsSurface->{4,6},VectorHead->Polygon]
```

Out[8] := *See Figure* 8

Its corresponding Gauss map image can displayed as:

```
In[9]:=ImplicitGaussMap[x^4+y^4+z^4-1,{x,-1,1},{y,-1,1},
        {z,-1,1},PlotPoints->{4,6}]
```

Out[9] := *See Figure* 9

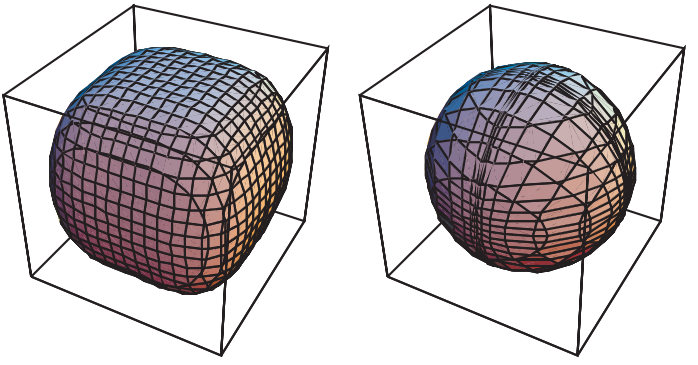


Fig. 9. Applying the Gauss map to the surface $x^4+y^4+z^4=1$: (left) original surface; (right) its Gauss map image onto the unit sphere

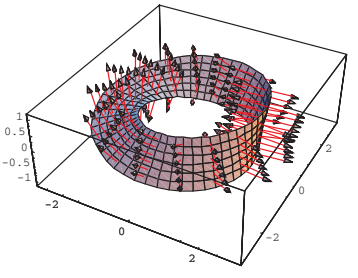


Fig. 10. The Möbius strip and its vector normal field

3.2 Parametric Surfaces

As indicated above, the package also deals with surfaces given in parametric form. In the following example, we consider the Möbius strip, parameterized by

$$\boldsymbol{S}(u,v)=\left(cos(u)+v\,cos(u)sin\left(\frac{u}{2}\right),sin(u)+v\,sin\left(\frac{u}{2}\right)sin(u),vcos\left(\frac{u}{2}\right)\right)$$

Figures 10 and 11 show the normal vector field and the Gauss map image of the Möbius strip, respectively.

```
In[10]:= ParametricNormalField[{Cos[u]+v Cos[u] Sin[u/2],
     Sin[u]+v Sin[u/2] Sin[u],v Cos[u/2]},{u,0,4 Pi,Pi/10},
     {v,-1/2,1/2,0.1},Surface->True,PlotPoints->{50,7},
     SurfaceDomain->{{0,2 Pi},{-1/2,1/2}},VectorHead->Polygon,
     HeadLength->0.25,HeadWidth->0.1,VectorColor->RGBColor[1,0,0]]
```

Out[10] := *See Figure* 10

```
In[11]:=ParametricGaussMap[{Cos[u]+v Cos[u] Sin[u/2],Sin[u]+
     v Sin[u/2] Sin[u],v Cos[u/2]},{u,0,4 Pi},{v,-1/2,1/2},
     SurfaceDomain->{{0,2 Pi},{-1/2,1/2}},PlotPoints->{50,7}]
```

Out[11] := *See Figure* 11

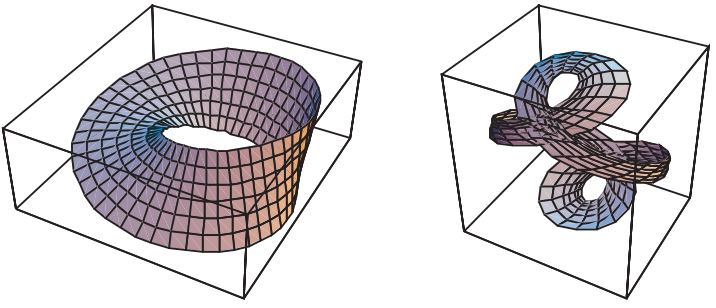


Fig. 11. Applying the Gauss map to the Möbius strip: (left) original surface; (right) its Gauss map image onto the unit sphere

4 Conclusions and Further Remarks

In this paper, a new Mathematica package for computing and displaying the normal vector field and the Gauss map image of implicit and parametric surfaces is introduced. The package is useful to analyze complicated surfaces, especially in those cases in which their Gauss map cannot be easily displayed without appropriate computer tools. The performance of the package is discussed by means of some illustrative and interesting examples.

All the commands have been implemented in Mathematica version 4 [7] and are consistent with Mathematica's notation and results. The powerful Mathematica functional programming [4] features have been extensively used to make the program shorter and more efficient. From our experience, Mathematica provides an excellent framework for this kind of developments. Perhaps the weakest feature of the package is the computation time, which typically ranges from several seconds to a few minutes. Further work will be carried out in order to improve this issue and other features.

Acknowledgements

This paper has been written while the second author was at the Laboratory of Advanced Research of the University of Tsukuba (Japan) for a sabbatical year stay. He would like to thank the laboratory staff, and especially Prof. Tetsuo Ida, for their wonderful hospitality and great collaboration. *Domo arigato gozaimazu!*

References

1. do Carmo, M.P.: Differential Geometry of Curves and Surfaces. Prentice-Hall, NJ (1976)
2. Koenderink, J.J.: Solid Shape. MIT Press, Cambridge, MA (1990)

3. Lowekamp, B., Rheingans, P., Yoo, T.S.: Exploring Surface Characteristics with Interactive Gaussian Images (A Case Study). In: Proceedings of IEEE Visualization'2002 IEEE Computer Society Press, Los Alamitos, CA, (2002) 553-556
4. Maeder, R.: Programming in Mathematica, Second Edition, Addison-Wesley, Redwood City, CA (1991)
5. O'Neill, B.: Elementary Differential Geometry. Second Edition, Academic Press, San Diego, CA (1997)
6. Tanaka, H.T., Ikeda, M., Chiaki, H.: Automatic face and gesture recognition. In: Proceedings of the 15th International Conference on Pattern Recognition'2000. IEEE Computer Society Press, Los Alamitos, CA (1998) 372-377
7. Wolfram, S.: The Mathematica Book, Fourth Edition, Wolfram Media, Champaign, IL & Cambridge University Press, Cambridge (1999)

Numerical-Symbolic *Matlab* Toolbox for Computer Graphics and Differential Geometry

Akemi Gálvez[1,⋆] and Andrés Iglesias[1,2]

[1] Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n
E-39005, Santander, Spain
[2] Department of Computer Science, University of Tsukuba,
Laboratory of Advanced Research, Building B, Room # 1025,
Kaede Dori, 305-8573, Tsukuba, Japan
uc8031@alumnos.unican.es, iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. In the last few years, computer algebra systems (CAS) have become standard and very powerful tools for scientific computing. One of their most remarkable features is their ability to integrate numerical, symbolic and graphical capabilities within a uniform framework. In addition, in most cases, these systems also incorporate a nice user interface making them specially valuable for educational purposes. In this work we introduce a user-friendly *Matlab* toolbox for dealing with many of the most important topics in Computer Graphics and Differential Geometry. The paper describes the main features of this program (such as the toolbox architecture, its simulation flow, some implementation issues and the possibility to generate standalone applications) and how the symbolic, numerical and graphical *Matlab* capabilities have been effectively used in this process.

1 Introduction

Much progress has been made in the world of computer algebra systems (CAS) during the last few years [2, 8, 9, 10, 11]. Nowadays, computer algebra systems (such as *Mathematica*, *Maple*, *Derive*, *Axiom*, *Matlab*, *Scilab*, etc.) have become standard and very powerful tools for scientific computing. As a consequence, they imply strong changes, almost a revolution, in the way of understanding scientific computing, with new mathematical algorithms, advanced programming features, numerical/symbolic integration tools, nice user interfaces, etc.

In this paper we will focus on the system *Matlab* [1], a numerical-oriented package that also includes a number of symbolic computation features. In fact, its most recent versions incorporate a *Maple* kernel totally integrated into the numerical layer. Even although such symbolic capabilities are not as powerful

⋆ Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 502–511, 2005.

as their counterpart in *Maple*, they are much faster and still very suitable for many purposes. In addition, the graphical *Matlab* capabilities (based on Open GL graphical routines and, hence, much better than those from many other CAS) and its excellent graphical user interface improve substantially the power of the system.

All these features have been exploited in this work to implement a unified symbolic/numerical *Matlab* Graphical User Interface (GUI) and toolbox for Computer Graphics and Differential Geometry. In this paper we describe the main features of this program and how the symbolic, numerical and graphical *Matlab* facilities have been applied to tackle this issue.

The structure of this paper is as follows: Section 2 describes the architecture of this symbolic/numerical *Matlab* toolbox, its main components and the simulation flow. Some implementation issues and the possibility of compiling and linking the toolbox files for generating standalone applications are also discussed in this section. Then, Section 3 analyzes the performance of this system by describing some illustrative examples. Finally, Section 4 closes with the main conclusions of this work.

2 The *Matlab* Toolbox Architecture

This section analyzes the architecture of the *Matlab* toolbox introduced in this paper. In addition, some implementation issues as well as the simulation flow are briefly described in this section. Finally, we discuss the possibility of compiling and linking the toolbox files for standalone applications.

2.1 System Architecture

The main system components as well as the simulation flow are depicted in Figure 1. In that figure, a dashed line is used to enclose those modules associated specifically with our *Matlab* toolbox. Basically, it consists of the following elements (see Figure 1):

- a *graphical user interface (GUI)*: *Matlab* provides a set of routines to create GUIs, incorporating menus, push and radio buttons, text boxes and many other interface devices. Most of computer algebra systems do not support GUIs. This fact was specially valuable for us, since our pedagogical orientation requires an interactive and user-friendly program. In addition, the GUI takes advantage of the powerful *Matlab* graphical capabilities. This was one of the most important features we are interested to deal with, and finally, one of the most important reasons to choose *Matlab*. Graphical *Matlab* features allow interactivity in the way of manipulating graphics, support many drawing commands and provide the components necessary to create computer graphics, including shading, texture mapping, hidden line removal, lighting, z-buffering, etc. In fact, we think they improve the usual graphical facilities available in other computer algebra systems. Reference [3] discusses

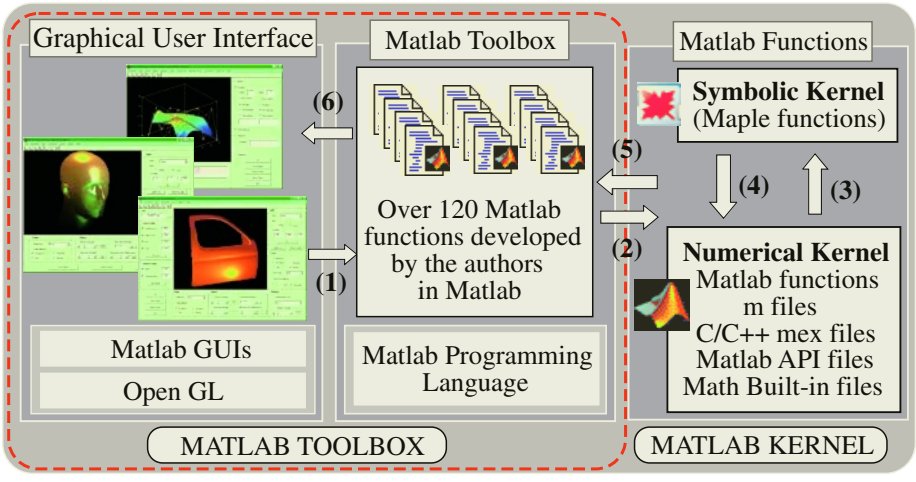


Fig. 1. Architecture of the *Matlab* GUI and toolbox described in this paper

these and other (oriented to computer graphics) interesting graphical *Matlab* features.

In our system, the GUI is comprised of a set of different windows associated with different tasks. Some windows are for differential geometry calculations and include some dialogue boxes in order to allow users to input the symbolic equations of the geometric entities. Other windows are mostly intended for graphical outputs and include a number of different commands for visualization. Those windows will be illustrated in Section 3.

– a *Matlab toolbox*: In our system, we implemented more than 120 functions for both Computer Graphics and Differential Geometry. Those functions call for auxiliar functions defined in *Matlab* kernel and allow us to deal with parametric curves and surfaces and other geometric entities. Of course, the libraries must be continuously updated, so the system must be flexible enough to allow the programmer to improve the algorithms and codes in an efficient, quick and easy way. Interesting enough is the fact that we can deal with those entities in either numerical or symbolic way. In this last case, we invoke the symbolic Math Toolbox, which is based on the kernel of *Maple*, one of the most popular symbolic computation programs.

  Another important factor to choose *Matlab* concerns the efficiency. Since *Matlab* is based on C, it runs faster than other analyzed symbolic and numerical programs. Moreover, its basic element is an array that does not require dimensioning, so it takes less time to be computed. Although the symbolic operations are not suitable for these based-on-array advanced features, numerical and graphical operations (which are the key ingredient of many commands for computer graphics and scientific visualization) have been greatly improved by applying those features on a regular basis.

– *Matlab graphics library*: This module includes the graphical commands needed for our setup. For instance, in the case of PC platforms, it is a set of DLLs

(Dynamic Link Libraries). In general, they are not included in the *Matlab* standard version and should be purchased separately.

2.2 Simulation Flow

Figure 1 also shows the simulation flow of the system: firstly, the user launches *Matlab* and calls for the GUI, that looks like the three windows shown on the left of that figure. Then, the user inputs some values and/or equations in the corresponding forms (dialogue boxes, radio buttons, check buttons and so on) which are sent to the system to be interpreted. User's input typically requires high-level functions developed by the authors in the toolbox described above (step 1 in Figure 1). At their turn, these functions invoke low-level functions from the numerical kernel (2), which are given either as *Matlab* functions, m-files, C/C++ MEX files, API files or built-in files (see [1] for details). If the user wishes to perform calculations in symbolic way, the symbolic kernel is also invoked; that input is subsequently sent to the symbolic kernel (based on *Maple* functions) in order to compute the corresponding output accordingly (3). It is returned to *Matlab* to be combined with numerical routines if, for instance, the user calls for graphical representations in addition to the symbolic output (4). Those results are sent back to the toolbox for subsequent calculations and/or manipulations (5), and eventually returned to the user as the final output (6).

2.3 Implementation Issues

Regarding the implementation, the toolbox has been developed by the authors in *Matlab* v6.0 by using a Pentium III processor at 2.2 GHz. with 256 MB of RAM. However, the program (whose minimum requirements are listed in Table 1) supports many different platforms, such as PCs (with Windows 9x, 2000, NT, Me and XP), Macintosh (Power Mac, 68040 and MacOSX) and UNIX workstations from Sun, Hewlett-Packard, IBM, Silicon Graphics and Digital. All figures in this paper correspond to the PC platform version.

As shown in Figure 1, the graphical tasks are performed by using the *Matlab* GUI for the higher-level functions (windowing, menus, or input) while it uses the graphical library *Open GL* for rendering purposes. The toolbox itself has been implemented in the native *Matlab* programming language, that is based on C for compiled functions.

Table 1. Minimum hardware configuration to install the toolbox

Hardware	**Requirement**
Operating System	Windows (9x, 2000, NT, Me, XP), MacOS, MacOSX, UNIX
RAM	128 MB
Disk storage	25 MB (60 MB recommended)
Monitor	Super VGA monitor
Screen resolution	minimum of 640 $\times$ 480
Matlab versions supported	5.3 - 7.0

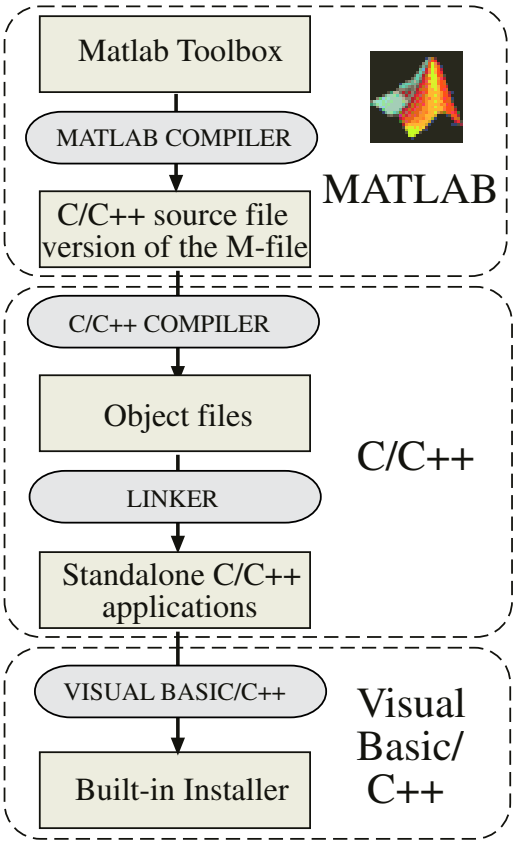


Fig. 2. Process for creating a C/C++ standalone application

2.4 Generation of Standalone Applications

Another remarkable feature of the *Matlab* toolbox is the possibility of compile and link it in order ot generate standalone applications. Figure 2 shows the different steps to be performed in order to get such standalone application. The up-down arrows indicate the compilation flow while dashed rectangles enclose the different programming environments required for each step indicated within. Roughly, this process can be summarized as follows: firstly, the *Matlab* toolbox is compiled by using the *Matlab* compiler. The result is a set of multiple C or C++ source code modules that are actually versions in C/C++ of the initial m-files implemented in *Matlab*. These files are subsequently compiled in a C/C++ environment to generate the object files. Then, those files are linked with the C++ graphics library, M-file library, Built-In library, API library and ANSI C/C++ library files. The final result is an executable program running on our computer.

We should remark here the excellent integration of all these tools to generate optimized code that can easily be invoked from C/C++ via Dynamic Link

Libraries (DLLs), providing both great portability and optimal communication between all modules. Complementary, we can construct an installer for the toolbox by using either Visual C++ or Visual Basic environments. The authors have accomplished all tasks mentioned above. The performance of this system is discussed in the next section.

3 Some Illustrative Examples

In this section the performance of the *Matlab* toolbox is discussed by means of some illustrative examples.

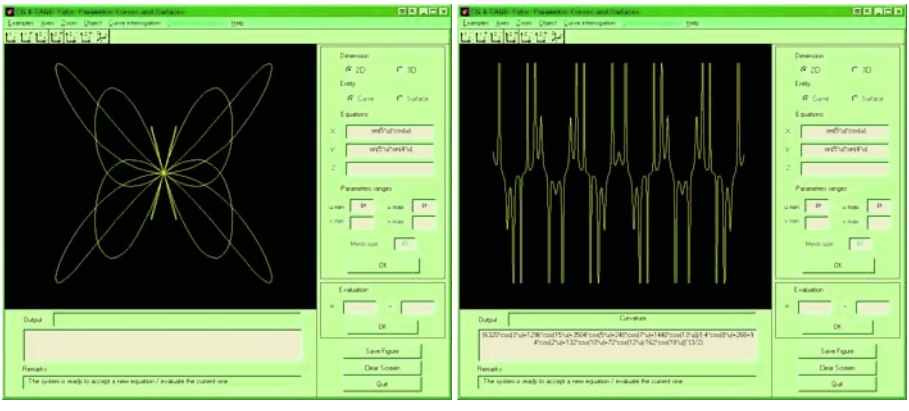


Fig. 3. (left) parametric planar curve and its graphical representation; (right) graphical representation and symbolic expression of its curvature

The first example is displayed in Figure 3: the computation of the curvature of a planar parametric curve. On Fig. 3(left) the parametric equation of the curve, $x = sin(5u)\,cos(u)$, $y = sin(5u)\,cos(4u)$ for $u \in [-\pi, \pi]$ is given and, after pressing the OK button, the graphical representation of the curve is obtained. On the right, the curvature of that curve is displayed and computed in symbolic way. The corrresponding output is shown in the dialogue box below the picture. Additional options allow the user to calculate the symbolic expression of all the curvatures (Gauss, Mean and principals), the first and second fundamental forms (a key topic in Differential Geometry), etc. and their visualization on their prescribed domain. All these topics have a singular importance in research and industry. For instance, we use curvatures for designing smooth surfaces, derivatives for detecting surface irregularities, etc.

Several illumination effects can also be applied to simulate computer images with a high level of quality. In addition to their aesthetic effect, the illumination models are important to detect irregularities on surfaces, for example, by using some illumination lines called *isophotes* [6]. This procedure can simulate the real

situation of looking for the different reflection lines on the surface of a car body. Figure 4 shows the isophotes of the surface $z = sin(u*v)$ with $x, y \in [-6, 6]$. Once again, note the symbolic expression of those isophotes just below the picture in Figure 4.

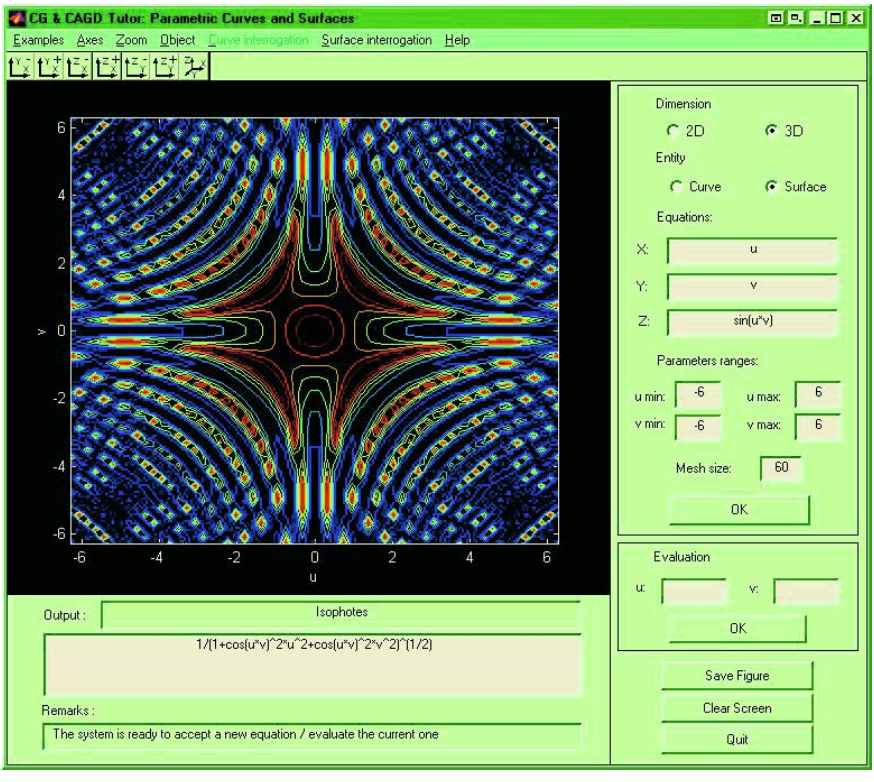


Fig. 4. Isophotes of the surface $z = sin(u * v)$ on the square $[-6, 6] \times [-6, 6]$

A nice feature of our toolbox is that it can be applied not only for educational purposes (which was actually our primary goal for the system) but also for industrial applications. For instance, the system includes a IGES file translator that deals with the most usual geometric entities in industry, including arcs, straight lines, cubic spline curves, B-spline curves, B-splines and NURBS surfaces and trimmed surfaces. Once the file is read, its geometric information is then manipulated by means of the numerical libraries in order to obtain the associated curves and surfaces. The following step is to plot them through the graphical libraries and the graphical *Matlab* options. These options include coloring, zooming, hidden line removal, shading, interactive rotation, etc. In addition, as one or several different geometric entities can be selected at any time, different outputs (wireframe model, surfaces model, trimmed model, etc.) are available from the

system. Figure 5 shows an example of a typical output, in this case, the right front seat door of a car body from a real IGES file. This is a simple example, but even in this case, more than 9500 B-spline curves, 1500 cubic spline curves, 3000 straight lines, almost 3000 arcs, 4182 NURBS surfaces, 628 trimmed surfaces, etc. are read and displayed. Each file might typically include several thousands of curves and surfaces of different types, and require some MB of hard disk to be stored. These data illustrate how complex an IGES file can be.

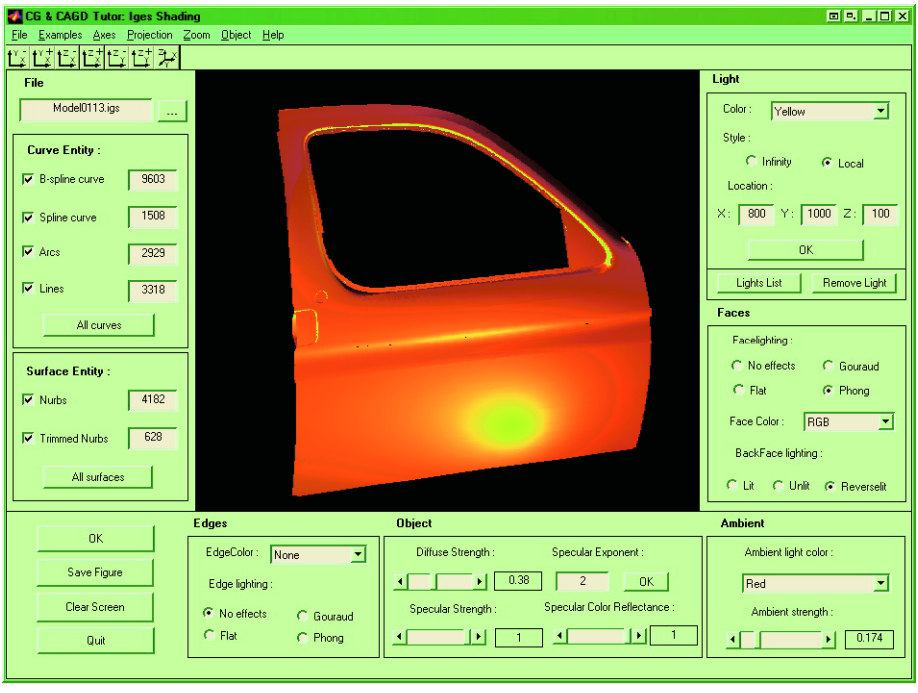


Fig. 5. Visualization of a real piece of a car body from a real IGES file

On the other hand, different effects (as ambient lights, especular and diffusive illumination, several light sources in different positions and with different colors, etc.) contribute to simulate a real image by computer with a high level of quality. Figure 6 shows the window for computer graphics, with a head obtained from a collection of patches joined through the `patch` *Matlab* command. In order to get this final image, several options have been applied: four different light sources (two yellow and one green sources at the infinity and another yellow local source at the point $(3, -4, 0)$) following the `phong` illumination model [5], a `shiny` material and a `copper` colormap, a `white` ambient light, a 80 % of specular light, a 35 % of diffuse light, etc.

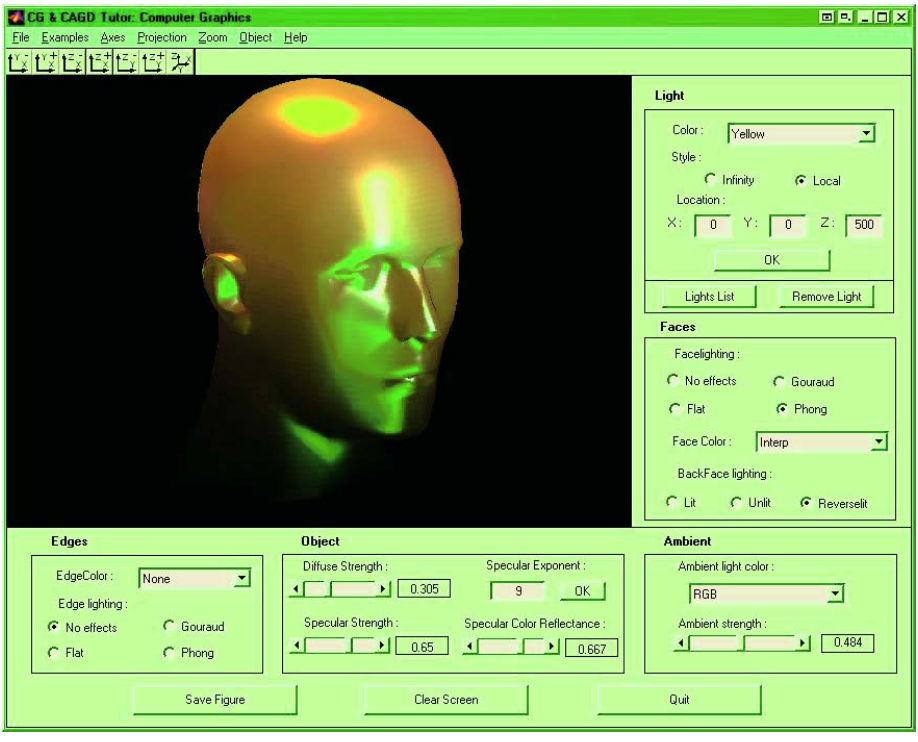


Fig. 6. Example of a computer graphics image generated in our system

4 Conclusions

In this paper we introduce a user-friendly *Matlab* toolbox for dealing with many of the most important topics in Computer Graphics and Differential Geometry. The paper describes the main features of this program (such as the toolbox architecture, its simulation flow, some implementation issues and the possibility to generate standalone applications) and how the symbolic, numerical and graphical *Matlab* capabilities have been effectively used in this process.

The program have been successfully applied for both educational and industrial purposes. In this last case, the initiative is actually part of a project for applying CAS to the automotive industry. This idea has been supported by the recent announcement of the use of *Matlab* by Daimler-Chrysler and Motor Ford Company (see [7] for details). To our knowledge, other similar approaches are expected for the next years. From this point of view, the present work could be helpful to those interested to follow this way.

References

1. The MathWorks Inc: Using Matlab, (1997); see also its Web Page: *http://www.mathworks.com*

2. Dewar, M. C.: Integrating Symbolic and Numeric Computation. In Cohen, A.M. (ed.): Computer Algebra in Industry. John Wiley and Sons, Chichester (1993) 221-232
3. Gálvez, A., Iglesias, A., Gutiérrez, F.: Applying Matlab to Computer Graphics and CAGD. Application to a Visualization Problem in the Automotive Industry. In: IX International Conference on Computer Graphics and Vision, GRAPHICON'99. Moscow (1999) 214-221
4. Gálvez, A., Iglesias, A., Otero, C., Togores, R.: Matlab Toolbox for a first Computer Graphics course for engineers. Lectures Notes in Computer Science **3044** (2004) 641-650
5. Hall, R.: Illumination and Color in Computer Generated Imagery. Springer-Verlag, New York (1989)
6. Poeschl, T.: Detecting surface irregularities using isophotes. Computer Aided Geometric Design **1** (1989) 163-168
7. Web Page: *http://www.mathworks.com/company/pressroom*.
8. Richard, C., Weber, A.: A Symbolic Numeric Environment for Analyzing Measurement Data in Multi-Model Settings. In Ganzha, V.G., Mayr, E.W., Vorozthtsov, E.V. (eds.): Computer Algebra in Scientific Computing, CASC'99. Springer-Verlag, Berlin Heidelberg New York (1999) 343-347
9. Ruhoff, P. T., Proestgaard, E., Perram, J. W.: Symbolic Derivation and Numerical Integration of the Equations of Motion for Constrained Dynamical Systems Using *Mathematica*. In: Keranen, V., Mitic, P. (eds.): Mathematics with Vision. Computational Mechanics Publications, Southampton (1995) 317-324
10. Suhonen, J., Lähde, O.: Combining Symbolic and Numerical Computating in the Multi-Band-Phonon-Projection-Model. In: Keranen, V., Mitic, P. (eds.): Mathematics with Vision. Computational Mechanics Publications, Southampton (1995) 243-250
11. Wang, P.S.: FINGER: A Symbolic System for Automatic Generation of Numerical Programs in Finite Element Analysis. Journal of Symbolic Computation **2** (1986) 305-316

A LiE Subroutine for Computing Prehomogeneous Spaces Associated with Real Nilpotent Orbits

Steven Glenn Jackson and Alfred G. Noël

Department of Mathematics,
University of Massachusetts,
Boston, MA 02125-3393, USA
{jackson, anoel}@math.umb.edu

Abstract. We describe an algorithm for decomposing certain modules attached to real nilpotent orbits into their irreducible components. These modules are prehomogeneous spaces in the sense of Sato and Kimura and arise in the study of nilpotent orbits and the representation theory of Lie groups. The output is a set of LaTeX statements that can be compiled in a LaTeX environment in order to produce tables. Although the algorithm is used to solve the problem in the case of exceptional real reductive Lie groups of inner type it does describe these spaces for the classical cases of inner type also. Complete tables for the exceptional groups can be found at http://www.math.umb.edu/~anoel/publications/tables/.

1 Introduction

In this paper, we continue our program begun in [8], [9] and [10] by describing a fast and stable algorithm for decomposing modules of a Lie subgroup of the Levi factor of Jacobson-Morozov parabolic subgroups defined by real nilpotent orbits. We will solve the problem by working on the other side of the Kostant-Sekiguchi correspondence [15]. In order to continue we need some definitions.

Unless otherwise specified, g will be a real semisimple Lie algebra of inner type with adjoint group G and Cartan decomposition $\mathsf{g} = \mathsf{k} \oplus \mathsf{p}$ relative to a Cartan involution θ. The term inner type means that $\text{rank}(\mathsf{g}) = \text{rank}(\mathsf{k})$. Some authors use the term equal rank to refer to such algebras. We will denote by $\mathsf{g}_{\mathbb{C}}$ the complexification of g. Let σ be the conjugation of $\mathsf{g}_{\mathbb{C}}$ with respect to g. Then $\mathsf{g}_{\mathbb{C}} = \mathsf{k}_{\mathbb{C}} \oplus \mathsf{p}_{\mathbb{C}}$ where $\mathsf{k}_{\mathbb{C}}$ and $\mathsf{p}_{\mathbb{C}}$ are obtained by complexifying k and p respectively. K will be a maximal compact Lie subgroup of G with Lie algebra k and $K_{\mathbb{C}}$ will be the connected subgroup of the adjoint group $G_{\mathbb{C}}$ of $\mathsf{g}_{\mathbb{C}}$, with Lie algebra $\mathsf{k}_{\mathbb{C}}$. It is well known that $K_{\mathbb{C}}$ acts on $\mathsf{p}_{\mathbb{C}}$ and the number of nilpotent orbits of $K_{\mathbb{C}}$ in $\mathsf{p}_{\mathbb{C}}$ is finite. Furthermore, for a nilpotent $e \in \mathsf{p}_{\mathbb{C}}$, $K_{\mathbb{C}} \cdot e$ is a connected component of $G_{\mathbb{C}} \cdot e \cap \mathsf{p}_{\mathbb{C}}$.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 512–521, 2005.

2 The Kostant-Sekiguchi Correspondence

A triple (x, e, f) in $\mathfrak{g}_{\mathbb{C}}$ is called a *standard triple* if $[x, e] = 2e$, $[x, f] = -2f$ and $[e, f] = x$. If $x \in \mathfrak{k}_{\mathbb{C}}$ and e and $f \in \mathfrak{p}_{\mathbb{C}}$ then (x, e, f) is a *normal* triple. It is a result of Kostant and Rallis [7] that any nilpotent e of $\mathfrak{p}_{\mathbb{C}}$ can be embedded in a standard normal triple (x, e, f). Moreover e is $K_{\mathbb{C}}$-conjugate to a nilpotent e' inside of a normal triple (x', e', f') with $\sigma(e') = f'$ [15]. The triple (x', e', f') will be called a $Kostant - Sekiguchi$ or KS-triple .

Every nilpotent E' in $\mathfrak{g}$ is G-conjugate to a nilpotent E embedded in a triple (H, E, F) in $\mathfrak{g}$ with the property that $\theta(H) = -H$ and $\theta(E) = -F$ [15]. Such a triple will be called a KS-triple also.

Define a map c from the set of KS-triples of $\mathfrak{g}$ to the set of normal triples of $\mathfrak{g}_{\mathbb{C}}$ as follows:

$$x = c(H) = i(E - F)$$

$$e = c(E) = \frac{1}{2}(H + i(E + F))$$

$$f = c(F) = \frac{1}{2}(H - i(E + F))$$

The triple (x, e, f) is called the Cayley transform of (H, E, F). It is easy to verify that the triple (x, e, f) is a KS-triple and that $x \in i\mathfrak{k}$. The Kostant-Sekiguchi correspondence [15] gives a one to one map between the set of G-conjugacy classes of nilpotents in $\mathfrak{g}$ and the $K_{\mathbb{C}}$-conjugacy classes of nilpotents in $\mathfrak{p}_{\mathbb{C}}$. This correspondence sends the zero orbit to the zero orbit and the orbit through the nilpositive element of a KS-triple to the one through the nilpositive element of its Cayley transform. Michèle Vergne [16] has proved that there is in fact a diffeomorphism between a G-conjugacy class and the $\mathfrak{k}_{\mathbb{C}}$-conjugacy class asssociated to it by the Kostant-Sekiguchi correspondence.

3 The Modules

In light of the Kostant-Sekiguchi correspondence it is reasonable to study modules associated to $K_{\mathbb{C}}$-nilpotent orbits in the symmetric spaces $\mathfrak{p}_{\mathbb{C}}$ in order to understand real nilpotent orbits. Let e be a nilpotent element in $\mathfrak{p}_{\mathbb{C}}$. Without loss of generality we can embed e in a KS-triple (x, e, f). The action of ad_x determines a grading

$$\mathfrak{g}_{\mathbb{C}} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}^i_{\mathbb{C}}$$

where $\mathfrak{g}^i_{\mathbb{C}} = \{Z \in \mathfrak{g}_{\mathbb{C}} : [x, Z] = iz\}$.

It is a fact that $\mathfrak{g}^0_{\mathbb{C}}$ is a reductive Lie subalgebra of $\mathfrak{g}_{\mathbb{C}}$. Let $G^0_{\mathbb{C}}$ be the connected subgroup of $G_{\mathbb{C}}$ such that $\mathrm{Lie}(G^0_{\mathbb{C}}) = \mathfrak{g}^0_{\mathbb{C}}$. Then for $i \neq 0$ the vector spaces

$\mathfrak{g}^i_{\mathbb{C}} \cap \mathfrak{p}_{\mathbb{C}}$ are $G^0_{\mathbb{C}} \cap K_{\mathbb{C}}$-modules. Moreover, a theorem of Kostant and Rallis [7] asserts that $G^0_{\mathbb{C}} \cap K_{\mathbb{C}}$ admits a Zariski open and dense orbit on $\mathfrak{g}^2_{\mathbb{C}} \cap \mathfrak{p}_{\mathbb{C}}$; that is, the pair $(G^0_{\mathbb{C}} \cap K_{\mathbb{C}}, \mathfrak{g}^2_{\mathbb{C}} \cap \mathfrak{p}_{\mathbb{C}})$ is a prehomogeneous space in the sense of Sato and Kimura [14]. This prehomogeneous space plays an important role in our work, and our effort to better understand it lead us to develop this project. See [11] for more information.

Let us denote $G^0_{\mathbb{C}} \cap K_{\mathbb{C}}$, $\mathfrak{g}^i_{\mathbb{C}} \cap \mathfrak{k}_{\mathbb{C}}$ and $\mathfrak{g}^i_{\mathbb{C}} \cap \mathfrak{p}_{\mathbb{C}}$ by $K^0_{\mathbb{C}}$, $\mathfrak{k}^i_{\mathbb{C}}$ and $\mathfrak{p}^i_{\mathbb{C}}$ respectively. Then we shall show that for $i \neq 0$, $(G^0_{\mathbb{C}}, \mathfrak{g}^i_{\mathbb{C}})$, $(K^0_{\mathbb{C}}, \mathfrak{k}^i_{\mathbb{C}})$ and $(K^0_{\mathbb{C}}, \mathfrak{p}^i_{\mathbb{C}})$ are prehomogeneous spaces. We shall need the following lemma from È. B. Vinberg.

Lemma 1. *Let $H \subseteq GL(U)$ be an linear algebraic group, let L be a closed, connected subgroup of H and let $V \subseteq U$ be a subspace invariant with respect to L. Suppose that for any vector $v \in V$*

$$\mathfrak{h} \cdot v \cap V = \mathfrak{l} \cdot v.$$

Then the intersection of any orbit of H with the subspace V is a smooth manifold, each irreducible component of which is an orbit of L.

Proof. See [18], page 469.

□

Using the previous lemma we prove the following proposition:

Proposition 1. *For $i \neq 0$, $(G^0_{\mathbb{C}}, \mathfrak{g}^i_{\mathbb{C}})$, $(K^0_{\mathbb{C}}, \mathfrak{k}^i_{\mathbb{C}})$ and $(K^0_{\mathbb{C}}, \mathfrak{p}^i_{\mathbb{C}})$ are prehomogeneous spaces. Moreover, $(G^0_{\mathbb{C}}, \mathfrak{g}^2_{\mathbb{C}})$ and $(K^0_{\mathbb{C}}, \mathfrak{p}^2_{\mathbb{C}})$ are regular; that is, the complements of the open dense orbits in $\mathfrak{g}^2_{\mathbb{C}}$ and $\mathfrak{p}^2_{\mathbb{C}}$ (the* singular loci*) are hypersurfaces.*

Proof. To prove that $(G^0_{\mathbb{C}}, \mathfrak{g}^i_{\mathbb{C}})$ is a prehomogeneous space we identify $G_{\mathbb{C}}$, $G^0_{\mathbb{C}}$, $\mathfrak{g}_{\mathbb{C}}$ and $\mathfrak{g}^i_{\mathbb{C}}$ with H, L, U and V respectively in the preceding lemma. Hence we only need to show that for any $v \in \mathfrak{g}^i_{\mathbb{C}}$

$$\mathfrak{g}_{\mathbb{C}} \cdot v \cap \mathfrak{g}^i_{\mathbb{C}} = \mathfrak{g}^0_{\mathbb{C}} \cdot v$$

Clearly $\mathfrak{g}^0_{\mathbb{C}} \cdot v \subseteq \mathfrak{g}_{\mathbb{C}} \cdot v \cap \mathfrak{g}^i_{\mathbb{C}}$. Let $u \in \mathfrak{g}_{\mathbb{C}}$ such that $[u, v] \in \mathfrak{g}^i_{\mathbb{C}}$. Since $u = \sum_j u_j$ with $u_j \in \mathfrak{g}^j_{\mathbb{C}}$, $[u, v] = \sum_j [u_j, v]$ and $[u_j, v] \subseteq \mathfrak{g}^{i+j}_{\mathbb{C}}$. Hence $[u_j, v] = 0$ for $j \neq 0$ and we must have $[u, v] = [u_0, v]$ and $\mathfrak{g}_{\mathbb{C}} \cdot v \cap \mathfrak{g}^i_{\mathbb{C}} \subseteq \mathfrak{g}^0_{\mathbb{C}} \cdot v$. The result follows.

To prove that $(K^0_{\mathbb{C}}, \mathfrak{p}^i_{\mathbb{C}})$ is a prehomogeneous space we identify $G^0_{\mathbb{C}}, K^0_{\mathbb{C}}, \mathfrak{g}^i_{\mathbb{C}}$ and $\mathfrak{p}^i_{\mathbb{C}}$ with H, L, U and V respectively in the preceding lemma. We only need to show that for any $v \in \mathfrak{p}^i_{\mathbb{C}}$

$$\mathfrak{g}^0_{\mathbb{C}} \cdot v \cap \mathfrak{p}^i_{\mathbb{C}} = \mathfrak{k}^0_{\mathbb{C}} \cdot v$$

Clearly $\mathfrak{k}^0_{\mathbb{C}} \cdot v \subseteq \mathfrak{g}^0_{\mathbb{C}} \cdot v \cap \mathfrak{p}^i_{\mathbb{C}}$. Let $u \in \mathfrak{g}^0_{\mathbb{C}}$ be such that $[u, v] \in \mathfrak{p}^i_{\mathbb{C}}$. Since $x = x_k + x_p$ with $x_k \in \mathfrak{k}_{\mathbb{C}}$ and $x_p \in \mathfrak{p}_{\mathbb{C}}$, $[x, v] = [x_k, v] + [x_p, v]$ with $[x_k, v] \in \mathfrak{p}^i_{\mathbb{C}}$ and $[x_p, v] \in \mathfrak{k}^i_{\mathbb{C}}$. Since $[x, v] \in \mathfrak{p}^i_{\mathbb{C}}$ we must have $[x_p, v] = 0$. Hence $\mathfrak{g}^0_{\mathbb{C}} \cdot v \cap \mathfrak{p}^i_{\mathbb{C}} \subseteq \mathfrak{k}^0_{\mathbb{C}} \cdot v$.

To prove that $(K^0_{\mathbb{C}}, \mathsf{k}^i_{\mathbb{C}})$ is a prehomogeneous space we only have to repeat the previous argument replacing $\mathsf{p}^i_{\mathbb{C}}$ by $\mathsf{k}^i_{\mathbb{C}}$.

From theorem 1.4.4 in [13], in order to show that $(G^0_{\mathbb{C}}, \mathsf{g}^2_{\mathbb{C}})$ and $(K^0_{\mathbb{C}}, \mathsf{p}^2_{\mathbb{C}})$ are regular we only need to show that the centralizers $(G^0_{\mathbb{C}})^e$ and $(K^0_{\mathbb{C}})^e$ of e in $G^0_{\mathbb{C}}$ and $K^0_{\mathbb{C}}$ are reductive Lie subgroups. For $G^0_{\mathbb{C}}$ this was done by Springer and Steinberg in [1]. The $K^0_{\mathbb{C}}$ case was settled by Ohta [12]. □

The reader may wonder whether $(K^0_{\mathbb{C}}, \mathsf{k}^2_{\mathbb{C}})$ is regular in general. Here is a counterexample. Let $\mathsf{g} = FI$, the split real form of $F4$. Consider Orbit 20, labeled (204 4) in [3]. Then $\mathsf{k}^2_{\mathbb{C}}$ is a two dimensional representation of $K^0_{\mathbb{C}}$. The singular locus is $\{0\}$ and therefore is not a hypersurface. **Our goal is to describe the irreducible components of the $K^0_{\mathbb{C}}$-modules $\mathsf{p}^i_{\mathbb{C}}$ with $i \neq 0$ for all nilpotent orbits of the Lie group $K_{\mathbb{C}}$ in the symmetric space $\mathsf{p}_{\mathbb{C}}$.**

It is enough to consider the non-compact real forms of all simple complex Lie algebras. We have developed a general theory for classical simple Lie groups by exploiting the fact that their nilpotent orbits are parametrized by certain partitions of an integer related to their rank [2]. Such parametrization is not available for the exceptional groups. As a result our subroutine is needed in order to solve the problem in the cases where G is of exceptional type. As a byproduct we obtain another method of computing these prehomogeneous spaces for any fixed non-compact real form of equal rank. The results for complex Lie groups can be found in [8] and [9]. For classical real Lie groups the theory is currently being developed in [10]. This paper solves the problem for all the exceptional real forms, except EI and EIV. We shall settle these cases in a future paper.

4 Description of the Algorithm

Maintaining the above notations, let $\Delta = \{\alpha_1, \ldots, \alpha_l\}$ be the Bourbaki set of simple roots of $\mathsf{g}_{\mathbb{C}}$. Since g is of inner type, we can find a Cartan subalgebra h of k such that $\mathsf{h}_{\mathbb{C}}$ is a Cartan subalgebra of $\mathsf{k}_{\mathbb{C}}$ and of $\mathsf{g}_{\mathbb{C}}$. Let $\Delta_K = \{\beta_1, \ldots, \beta_l\}$ be the simple root system defined by $\mathsf{h}_{\mathbb{C}}$. The labels of the Dynkin diagram of $K_{\mathbb{C}}$ are obtained by evaluating the β_i's on the neutral element x in the KS-triple (x, e, f). We shall use Djokovic's labeling scheme for the exceptional groups [3, 4]. For the classical cases one can use the simple roots of a Vogan system [6] or one of the $K_{\mathbb{C}}$ root systems given in [5].

There is a one-to-one correspondence between nilpotent orbits of $K_{\mathbb{C}}$ in $\mathsf{p}_{\mathbb{C}}$ and a set of Dynkin diagrams of $K_{\mathbb{C}}$ whose nodes are labeled with numbers in the set of integers between $-\dim \mathsf{g}_{\mathbb{C}}$ and $\dim \mathsf{g}_{\mathbb{C}}$. This last statement means that there are only finitely many orbits of $K_{\mathbb{C}}$ in $\mathsf{p}_{\mathbb{C}}$, which implies that there are only finitely many nilpotent orbits in g by the Kostant-Sekiguchi correspondence.

The roots of $\mathsf{g}_{\mathbb{C}}$ are all imaginary because they all take purely imaginary values on h. We call an imaginary root α *compact* if its root space g^α is included $\mathsf{k}_{\mathbb{C}}$, *non-compact* if $\mathsf{g}^\alpha \subseteq \mathsf{p}_{\mathbb{C}}$.

Definition 1. *Let v be a non-zero element of the root space $\mathfrak{g}_{\mathbb{C}}^{\lambda} \subseteq \mathfrak{p}_{\mathbb{C}}^{i}$ and X_α a non-zero vector in any $\mathfrak{g}^{\alpha} \subseteq [\mathfrak{k}_{\mathbb{C}}^{0}, \mathfrak{k}_{\mathbb{C}}^{0}]$ with $\alpha > 0$. If $[X_\alpha, v] = 0$ for all such X_α's then we say that v is a* highest weight vector *of $K_{\mathbb{C}}^{0}$ on $\mathfrak{g}_{\mathbb{C}}^{i}$ and we call λ a* highest weight *of $\mathfrak{g}_{\mathbb{C}}^{i}$.*

We shall describe an algorithm which computes the highest weights of $\mathfrak{p}_{\mathbb{C}}^{i}$. The algorithm is written in the language LiE. Readers who are not familiar with LiE may consult [17]. The highest weights will be expressed in terms of the fundamental weights. For the reader's convenience all LiE generic functions will be written in boldface characters. The subroutine name is *rirrdmodule()*. The subroutine returns a set of LaTeX statements that generate a table containing all the highest weights of all $\mathfrak{p}_{\mathbb{C}}^{i}$ for $i > 0$. By duality one obtains a decomposition of $\mathfrak{p}_{\mathbb{C}}^{-i}$ also.

4.1 Algorithm

```
# This function computes irreducible components of K^0_C-modules in the
# symmetric space p_C. This function assumes that the real form of the
# algebra g_C is of inner type. We also assume that the Vogan system of
# simple roots contains exactly one non-compact root.
# Authors: Steven Glenn Jackson and Alfred G. Noël. # Purpose: In-
vestigation of prehomogeneous spaces associated with
# real nilpotent orbits.
# Date : August 10, 2004.
# Global Variable
orbit = 0;

rirrdmodule(tex rform; int ncptindex; vec label; vec neutral; mat kroots;
mat krootsadjt ; int krootsdet; grp g ) =
{

# rform: inner type non-compact real form of g_C # ncptindex: indicate
which roots are nompact.
# label : label describing the evaluation of the simple roots of K_C
# on the neutral element.
# neutral: semisimple element of the normal sl_2-triple associated with
# the orbit. It should be expressed in the Bourbaki root system
# of the algebra g_C.
# kroots: a Vogan system of simple roots for K_C.
# krootsadjt: adjoint matrix of kroots
# krootsdet: determinant of kroots

grptype = rform; printflag = 0; setdefault (g); n = n_pos_roots;
l = Lie_rank; alpha = pos_roots; roots = alpha^- alpha;
kcartan = null(l,l);
for i = 1 to l do for j=1to l do kcartan[i,j] =
Cartan(kroots[i],kroots[j]);od;od;
```

Validate the ordering of the K roots
if kcartan != **Cartan**(**Cartan_type**(kroots)) **then** print ("Non compatible ordering of k roots");
else

# Find the number of non-compact roots
ncpt = 0; **for** k = 1 to 2*n **do if** (roots[k][ncptindex] == 1 ‖ roots[k][ncptindex] == -1) **then** ncpt = ncpt +1; **fi**; **od**;

look_for_degree = 0; *mod_counter* = ncpt ;
while *mod_counter* > 0 **do**
spaceofinterest = **null**(ncpt,l);

#Compute degree positive roots
counter = 0;
for k = 1 to 2*n **do if** (roots[k][ncptindex] == 1 ‖ roots[k][ncptindex] == -1) **then** degree = **null**(2*n);
for i = 1 to l **do for** j = 1 to l **do** degree[k] = degree[k] + neutral[j]*roots[k][i]***Cartan**(alpha[i],alpha[j]); **od**; **od**;
if degree[k] == *look_for_degree* **then**
counter = counter+1; spaceofinterest[counter] = roots[k]; **fi**; **fi**; **od**;

space = **null**(counter, l);
for i = 1 to counter **do** space[i] = spaceofinterest[i]; **od**;
dimspace = counter;

keep track of spaces using the dimension of the dual modules.
if *look_for_degree* == 0 **then** *mod_counter* = *mod_counter* - dimspace;
else *mod_counter* = *mod_counter* - 2 dimspace;**fi**;

if *mod_counter* < 0 **then** print(" Negative Modcounter"); break;**fi**;
if dimspace >0 && *look_for_degree* >0 **then**
latexline = " ";
if printflag == 1 **then** latexline = "*cline*3 − 5&"; **fi**; **if** printflag == 0 **then** printflag = 1; orbit = orbit +1;
latexline = "*hline*" + orbit + " & " + grptype; **for** i = 1 to l **do** latexline = latexline+"{"+label[i] + "}" ; **od**; **fi**;
# Pick up Highest weights of $K_{\mathbb{C}}^0$ in space
hw = **all_one**(dimspace);
for i = 1 to l **do if** label[i] == 0 **then for** j = 1 to dimspace **do** candidate = kroots[i] + space[j];

Check if candidate is in space. If yes then it is not a highest weight
for k = 1 to dimspace **do if** candidate == space [k] **then** hw[j] = 0; break;**fi**; **od**; **od**; **fi**; **od**;

```
#Compute highest weight_ts
numhw = 0;
for i = 1 to dimspace do numhw = numhw + hw[i]; od;
rhighestweights = null(numhw,l);
counter = 0; for i = 1 to dimspace do if hw[i] == 1 then counter =
counter+1; rhighestweights[counter] = space[i]; fi; od;

#Convert to fundamental weights of K^0_c
if numhw > 0 then
fundhighestweights = rhighestweights *(krootsadjt) * kcartan;
fundhighestweights = fundhighestweights/krootsdet;

# Generate LaTeX output.
for i = 1 to numhw do latexline = latexline + "("; for j = 1 to l do
latexline = latexline + fundhighestweights[i][j];
if j < l then latexline = latexline + "," fi; od;
latexline = latexline + ") \\"; od;
latexline = latexline + "\ ea \\ "; fi;
latexline = latexline + " & " + look_for_degree + " & " + dimspace +
" & \ ba"; print (latexline);
fi; look_for_degree = look_for_degree+1; od; fi;
}
```

Example 1. Let $\mathfrak{g}$ be the real algebra of type G_2; then $K_{\mathbb{C}}$ is of type $A_1 \oplus A_1$. Let $\Delta = \{\alpha_1, \alpha_2\}$ be the usual Bourbaki system of simple roots of $\mathfrak{g}_{\mathbb{C}}$. Then we can choose $\Delta_k = \{\alpha_1, 3\alpha_1 + 2\alpha_2\}$ to be a system of simple roots for $K_{\mathbb{C}}$. Hence the compact roots of $\mathfrak{g}_{\mathbb{C}}$ are $\pm\alpha_1, \pm(3\alpha_1 + 2\alpha_2)$ and the non-compact roots are $\pm\alpha_2, \pm(\alpha_1 + \alpha_2), \pm(2\alpha_1 + \alpha_2), \pm(3\alpha_1 + \alpha_2)$.

Here is the LiE code that calls the subroutine in order to generate the irreducible modules. This is the content of the file *rgmodule*.

```
g = G2;
rirrdmodule("\typeg",2, [1,1],[1,1], [[1,0],[3,2]], [[2,0],[-3,1]],2,g);
rirrdmodule("\typeg", 2,[1,3],[2,3], [[1,0],[3,2]], [[2,0],[-3,1]],2,g);
rirrdmodule("\typeg", 2,[2,2],[2,2], [[1,0],[3,2]], [[2,0],[-3,1]],2,g);
rirrdmodule("\typeg", 2, [0,4],[2,4], [[1,0],[3,2]], [[2,0],[-3,1]],2,g);
rirrdmodule("\typeg", 2, [4,8],[6,8], [[1,0],[3,2]], [[2,0],[-3,1]],2,g);
```

The name of the file containing the subroutine is *rphmod* which has to be loaded in the LiE environment. Here is the LiE session that produces Table 1:

```
LiE version 2.2 created on Nov 22 1997 at 16:50:29 Authors: Arjeh
M. Cohen, Marc van Leeuwen, Bert Lisser. Mac port by S. Grimm
Public distribution version  type '?help' for help information
type '?' for a list of help entries.

> read rphmod
> read rgmodule
```

Table 1.

Nilpotent orbits in type G				
Orbit	$K_{\mathbb{C}}$-Label	i	$\dim \mathfrak{p}^i_{\mathbb{C}}$	Highest weights of $\mathfrak{p}^i_{\mathbb{C}}$
1	○ ○ 1 1	1	2	$(1,1)$ $(3,-1)$
		2	1	$(3,1)$
2	○ ○ 1 3	1	1	$(-1,1)$
		2	1	$(1,1)$
		3	1	$(3,1)$
3	○ ○ 2 2	2	2	$(1,1)$ $(3,-1)$
		4	1	$(3,1)$
4	○ ○ 0 4	2	4	$(3,1)$
5	○ ○ 4 8	2	2	$(-1,1)$ $(3,-1)$
		6	1	$(1,1)$
		10	1	$(3,1)$

By the "Theorem of the Highest Weight" [6–p. 279], $\mathfrak{p}^i_{\mathbb{C}}$ is completely determined as $\mathfrak{k}^0_{\mathbb{C}}$-module by its highest weights. Moreover, when interpreting the results given in the table, one should be aware that the action of the semisimple part of $\mathfrak{k}^0_{\mathbb{C}}$ on $\mathfrak{p}^i_{\mathbb{C}}$ is completely determined by those coefficients associated with the nodes of the Dynkin diagram of $K_{\mathbb{C}}$ labeled with 0; the other coefficients affect only the action of the center of $\mathfrak{k}^0_{\mathbb{C}}$. Then (disregarding the action of the center) we see that for orbit 4, $\mathfrak{p}^2_{\mathbb{C}} = V^{3\omega_1} \simeq S^3(\mathbb{C}^2)$ which has dimension 4.

The computation of highest weights is accomplished by finding the set of vectors in each $\mathfrak{p}^i_{\mathbb{C}}$ which are annihilated by all positive simple root spaces in $[\mathfrak{k}^0_{\mathbb{C}}, \mathfrak{k}^0_{\mathbb{C}}]$. Such root spaces correspond to the simple roots labeled with 0. Since $\mathfrak{k}^0_{\mathbb{C}}$ acts by the adjoint action, the subroutine returns only the roots β such that $\mathfrak{g}^\beta$ lies in $\mathfrak{p}^i_{\mathbb{C}}$ and $[X_\alpha, X_\beta] = 0$ for all positive α with $\mathfrak{g}^\alpha \subset [\mathfrak{k}^0_{\mathbb{C}}, \mathfrak{k}^0_{\mathbb{C}}]$. This is exactly the set of highest weights of $\mathfrak{p}^i_{\mathbb{C}}$, which is recorded in the array *rhighestweights*. Finally, we express the highest weights in terms of the fundamental weights using the Cartan matrix of $K_{\mathbb{C}}$.

The algorithm is $\mathcal{O}(\text{rank}(\mathfrak{g}) \times (n\_pos\_roots)^3)$. This is due to the fact that the control variable of the outer **while** loop, *mod_counter*, is bounded by the number of positive roots and that the more intensive internal **for** loop is at worst executed $(\text{rank}(\mathfrak{g}) \times (n\_pos\_roots)^2)$ times. This is a worst case analysis. Of course we are assuming that the LiE internal functions are very fast. From our experience we believe that they are optimal. On average the subroutine performs very well. We use it to compute the irreducible modules of the prehomogeneous spaces associated to all the nilpotent orbits of the exceptional real simple Lie groups of inner type. For more information see [10]. The computations were carried on an IMac G4 with speed 1GHz and 1Gb SDRAM of memory.

4.2 Computation of Irreducible Components of $\mathbf{k}_{\mathbf{c}}^{i}$

The above subroutine will return the highest weight vectors of $\mathsf{k}_{\mathbb{C}}^{i}$ if the command

if *(roots[k][ncptindex] == 1 || roots[k][ncptindex] == -1)* **then**
is replaced by
if *(roots[k][ncptindex] != 1 && roots[k][ncptindex] != -1)* **then**.

5 Conclusion and Future Work

We presented a LiE implementation of a simple algorithm for computing the module structures of a large class of prehomogeneous spaces, namely those associated with nilpotent orbits in the adjoint representations of non-compact real reductive Lie groups. In [10] we are developing general methods for the classical types, taking advantage of the parametrization of the nilpotent orbits by partitions. The correctness of the algorithm is a consequence of the well-known "Theorem of the Highest Weight."

This is the foundation of a larger project whose goals are to understand regularity criteria on these prehomogeneous spaces, to find nice formulas for computing their relative invariants and finally to explicitly relate them to Representation Theory.

References

1. Borel A., Carter R., Curtis C. W., Iwahori N., Springer T. A. and Springer R. *Seminar algebraic groups and related finite groups.* Lecture Notes in Mathematics **131**, Springer (1970).
2. Collingwood, D. H. and McGovern, W. M. *Nilpotent orbits in semisimple Lie algebras.* Van Nostrand Reinhold Mathematics Series, New York (1992).
3. Djoković D. *Classification of nilpotent elements in simple exceptional real Lie algebras of inner type and description of their centralizers* J. Alg. **112** (1988) 503-524.
4. Djoković D. *Classification of nilpotent elements in the simple real Lie algebras* $E_{6(6)}$ *and* $E_{6(-26)}$ *and description of their centralizers* J. Alg. **116** (1988) 196-207.
5. King D. R. *Classification of spherical nilpotent orbits in complex symmetric space* J. Lie Theory **14** (2004), no. 2, 339–370.
6. Knapp, A. W. *Lie groups beyond an introduction.* Second edition, Birkhaüser Progress in Mathematics **140** (2002)
7. Kostant B., Rallis S. *Orbits and representations associated with symmetric spaces* Amer. J. Math. **93** (1971) 753-809.
8. Jackson, S. G. and Noël, A. G. *Prehomogeneous spaces associated with complex nilpotent orbits*, to appear in Journal of Algebra
9. Jackson, S. G. and Noël, A. G. *A LiE subroutine for computing prehomogeneous spaces associated with complex nilpotent orbits*, to appear in Lecture Notes in Computer Science, Springer Verlag.
10. Jackson, S. G. and Noël, A. G. *Prehomogeneous spaces associated with real nilpotent orbits* (in preparation).
11. Noël A. G. *Nilpotent Orbits and Theta-stable Parabolic Subalgebras* American Mathematical Society Journal of Representation Theory, **2**, (1998) 1-32.

12. Ohta T., *Classification of admissible nilpotent orbits in the classical real Lie algebras*, J. of Algebra **136**, N0. 1 (1991) 290-333.
13. Rubenthaler, H. *Algèbres de Lie et espaces préhomogènes.* Hermann, Travaux en Cours, Paris, 1992.
14. Sato, M. and Kimura, T. *A classification of irreducible prehomogeneous vector spaces and their relative invariants.* Nagoya Math. J. **65** (1977), 1-155.
15. Sekiguchi J. *Remarks on real nilpotent orbits of a symmetric pair* J. Math. Soc. Japan **39**, No. 1 (1987), 127-138.
16. Vergne M. *Instantons et correspondence de Kostant-Sekiguchi*, R. Acad. Sci. Paris Sér. I Math. **320** (1995), 901-906.
17. Van Leeuwen, M. A. A., Cohen, A. M., and Lisser, B. *LiE: A package for Lie group computations*, Computer Algebra Nederland, Amsterdam, Netherlands (1992)
18. Vinberg, È. B. *The Weyl group of a graded Lie Algebra* IZV. Akad. Nauk SSSR Ser. Mat. Tom **40** (1976) no. 3.

Applications of Graph Coloring

Ünal Ufuktepe and Goksen Bacak

Izmir Institute of Technology, Department of Mathematics,
Urla, Izmir, Turkey
{unalufuktepe, goksenbacak}@iyte.edu.tr

Abstract. A graph G is a mathematical structure consisting of two sets $V(G)$ (vertices of G) and $E(G)$ (edges of G). Proper coloring of a graph is an assignment of colors either to the vertices of the graphs, or to the edges, in such a way that adjacent vertices / edges are colored differently. This paper discusses coloring and operations on graphs with *Mathematica* and web*Mathematica*. We consider many classes of graphs to color with applications. We draw any graph and also try to show whether it has an Eulerian and Hamiltonian cycles by using our package ColorG.

1 Introduction

Graph theory would not be what it is today if there had been no coloring problems. In fact, a major portion of the 20th-century research in graph theory has its origin in the four color problem [1]. A graph G is a mathematical structure consisting of two sets $V(G)$ (vertices of G) and $E(G)$ (edges of G). Proper coloring of a graph is an assignment of colors either to the vertices of the graphs, or to the edges, in such a way that adjacent vertices / edges are colored differently. Vertex coloring is a hard combinatorial optimization problem.

We apply several operations which act on graphs to give different graphs. In addition to apply graph operations, we color vertices of these obtained graphs properly. Also we developed ColorG package to color the vertices and edges of graphs and to find the Eulerian and the Hamiltonian cycles with web*Mathematica*. Many of these graphs are truly beautiful when drawn properly, and they provide a wide range of structures to manipulate and study.

Before concluding this introduction, we recall some basic definitions.

A complete graph is a simple graph such that every pair of vertices is joined by an edge. A nontrivial closed path is called a cycle. A graph which is obtained by joining a new vertex to every vertices of a cycle is called a wheel. A connected acyclic graph is called a tree [4].

2 Graph Coloring with WebMathematica

One of the most exciting new technologies for dynamic mathematics on the World Wide Web is a web*Mathematica*. This new technology developed by Wolfram research enables instructors to create web sites that allows users to compute

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 522–528, 2005.

and visualize results directly from a web browser. web*Mathematica* is based on a standard java technology called servlets. It allows a site to deliver HTML pages that are enhanced by the addition of *Mathematica* commands [5]. When a request is made for one of these pages the *Mathematica* commands are evaluated and the computed result is placed in the page. People who access web*Mathematica* sites do not have to know how to use Mathematica [11].

In this section, we give applications of ColorG package to color the vertices and the edges of the graphs with web*Mathematica*.

2.1 Vertex Coloring

The most applications involving vertex coloring are concerned with determining the minimum number of colors required under the condition that the end points of an edge cannot have the same color. A proper vertex coloring of a graph is an assignment from its vertex set to a color set that the end points of each edge are assigned two different colors. The chromatic number of a graph G, denoted by $\chi(G)$, is the minimum number of different colors required for a proper vertex coloring of G. Applications of vertex coloring include scheduling, assignment of radio frequencies, separating combustible chemical combinations, and computer optimization. We use some commands in the Combinatorica package with Mathematica to color the vertices of graphs and to give web-based examples with webMathematica as in The Fig. 1 [9].

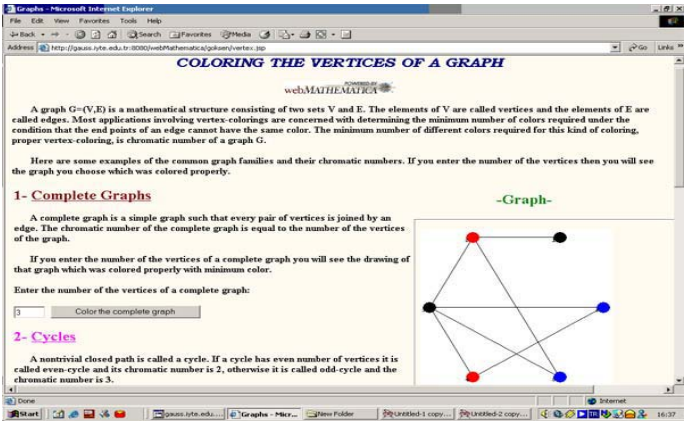


Fig. 1. Vertex Coloring of a Graph with webMathematica

If the users press the CompleteGraph, Cycle, Wheel, Star, RandomTree, and any graph and enter the number of vertices they can get the vertex-colored graph.

2.2 Edge Coloring

Edge coloring is an optimization problem: An edge-coloring of a graph G is an assignment of colors to the edges of G such that edges with a common endpoint

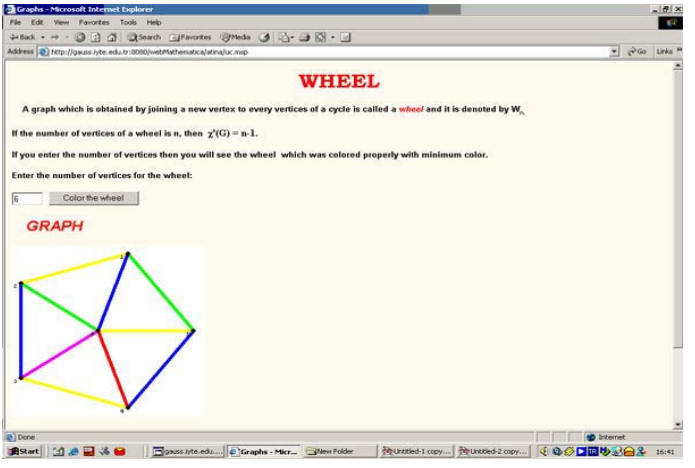


Fig. 2. Edge Coloring of a Graph with webMathematica

have different colors. Let $\chi(G')$ denote the chromatic index of G, that is the minimum number of colors necessary to color the edges of G. Vizing [10] proved that $\chi(G')$ is either $\triangle(G)$ or $\triangle(G)+1$ for each graph G, where $\triangle(G)$ denotes the maximum degree of a vertex in G. Then the graph G belongs to one of two classes; either to class 1 or to class 2. This classification problem is NP-complete, and this implies that there are no polynomial-time algorithms for this problem. We use some commands in the ColorG package with Mathematica to color the edges of graphs and to give web-based examples with webMathematica as in The Fig. 2 [9].

If the user press the Wheel button and enter the number of vertices, he/she can get the edge-colored graph.

3 Generating Graphs with WebMathematica

The most important operations on graphs are sum, union, join, and product of two graphs. We will do these operations with the common graphs; complete graph, random tree, wheel, and cycle.

The join of two graphs is their union with the addition of edges between all pairs of vertices from different graphs. To take the join of two graphs the user should enter the number of the graphs and their vertex numbers into the boxes then he/she sees the join of those graphs and also its proper vertex coloring.

This operation above is extended for the other graph operations: sum, join, and product, also.

The union of two graphs is formed by taking the union of the vertices and edges of the graphs. Thus the union of graphs is always disconnected. The sum operation of two graphs is to take the edges of the second graph and add them to the first graph. The product of two graphs $G \times H$ has a vertex set defined by the Cartesian product of the vertex sets of G and H. There is an edge between

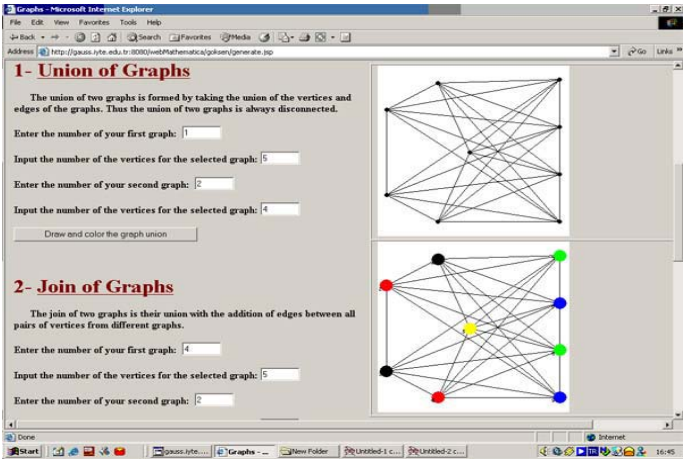


Fig. 3. A Generated Graph with webMathematica

(u, v) and (s, t) if $u = s$ and v is adjacent to t in H or $v = t$ and u is adjacent to s in G.

4 Cycle Structure in Graphs with WebMathematica

A cycle in a graph is a simple closed path. We will represent a cycle in G as a list of vertices $C = v_1, v_2, ..., v_1$ such that there is an edge of G from each vertex to the next in G.

4.1 Eulerian Cycle

Euler initiated the study of graph theory in 1736 with the famous Seven Bridges of Königsberg problem. The town of Königsberg straddled the Pregel River with a total of seven bridges connecting the two shores and two islands. The townsfolk were interested in crossing every bridge exactly once and returning to the starting point. An Eulerian cycle is a complete tour of all the edges of a graph. The term circuit is often used instead of cycle, since each vertex can be visited more than once.

We use ColorG package with Mathematica to find the Eulerian cycle and to give web-based examples with webMathematica. If the number of the vertices is entered, it is possible to see the Eulerian cycle in that graph if there exists.

4.2 Hamiltonian Cycle

A Hamiltonian cycle of a graph G is a cycle which visits every vertex in G exactly once, as opposed to an Eulerian cycle which visits each edge exactly once. A Hamiltonian path is like a Hamiltonian cycle, except that it is a path. The problem of computing a Hamiltonian cycle or a Hamiltonian path is fundamen-

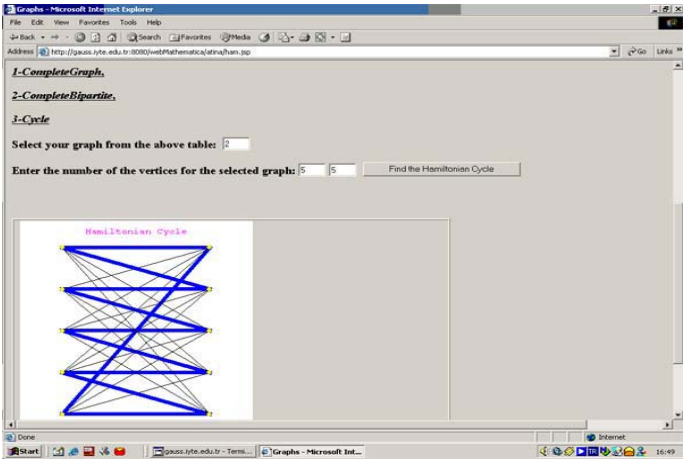


Fig. 4. A Hamiltonian Cycle in a Complete Bipartite Graph with webMathematica

tally different from the problem of computing an Eulerian cycle, because testing whether a graph is Hamiltonian is NP-complete.

We use ColorG package with Mathematica to find the Hamiltonian cycle and to give web-based examples with web*Mathematica*. If the number of the vertices is entered, it is possible to see the Hamiltonian cycle in that graph if there exists. The Fig. 4. shows the Hamiltonian Cycle for the Complete Bipartite Graph.

4.3 An Application

Some scheduling problems induce a graph coloring, i.e., an assignment of positive integers (colors) to vertices of a graph. We discuss a simple example for coloring the vertices of a graph with a small number k of colors and present computational results for calculating the chromatic number, i.e., the minimal possible value of such a k. Draw up an examination schedule involving the minimum number of days for the following problem.

Example: Set of students: $S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9$ Examination subjects for each group: {algebra, real analysis, and topology}, {algebra, operations research, and complex analysis}, {real analysis, functional analysis, and topology}, {algebra, graph theory, and combinatorics}, {combinatorics, topology, and functional analysis}, {operations research, graph theory, and coding theory}, {operations research, graph theory, and number theory}, {algebra, number theory, and coding theory}, {algebra, operations research, and real analysis}.

Let S be a set of students, $P = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ be the set of examinations respectively algebra,real analysis, topology, operational research, complex analysis, functional analysis, graph theory, combinatorics, coding theory, and number theory. $S(p)$ be the set of students who will take the examination $p \in P$. Form a graph $G = G(P, E)$, where $a, b \in P$ are adjacent

if and only if $S(a) \bigcap S(b) \neq \emptyset$. Then each proper vertex coloring of G yields an examination schedule with the vertices in any color class representing the schedule on a particular day. Thus $\chi(G)$ gives the minimum number of days required for the examination schedule. The Mathematica commands for this solution are as follows:

```
<< DiscreteMath`ColorG`
k = Input["Input the number of the students"];
S = Table[Input["Input number of the lessons which the student
will choose"],  k];
b = Union[Flatten[Table[KSubsets[S[[i]], 2], i, k], 1]];
ColorVertices[t = DrawG[b]];
h = VertexColoring[t]; d=ChromaticNumber[t];
Print[d"days are required and you can see below the lessons in the
same  parenthesis which are on the same day"]
Table[Flatten[Position[h, i], 2], i, Max[h]]
```

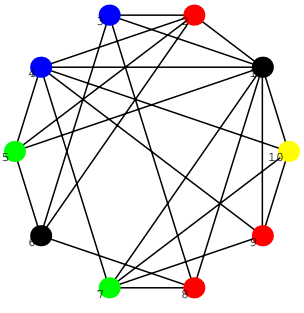


Fig. 5. The colored graph of the example

5 days are required and you can see below the lessons in the same parenthesis which are on the same day

$$\{\{1,6\},\{2,8,9\},\{3,4\},\{5,7\},\{10\}\}$$

It was very exciting to take 100-year old ideas, simple as they are, and implement them in *Mathematica* and web*Mathematica* for anybody. But, there is more work to be done, both of a theoretical and practical nature. If we consider a coloring problem posed as a two-person game, with one person (Alice) trying to color the graph, and the other (Bob) trying to prevent this from happening. Alice and Bob alternate turns, with Alice having the first move. A move consisting of selecting an uncolored vertex x and assigning it a color from the color set X distinct from the colors assigned previously to neighbors of x. If after $n = |V(G)|$ moves, the graph G is colored, Alice is the winner. Bob wins if an impass is reached before all vertices in the graph are colored. The game

chromatic number of a graph G, denoted by $\chi_g(G)$, is the least cardinality of a color set X for which Alice has a winning strategy [3]. We believe that it is possible to play this game online with web*Mathematica*. Of course, this leads to the rich area of game coloring and the many difficult and intriguing questions there.

The following URL address shows our works on graph coloring: http://gauss.iyte.edu.tr:8080/webMathematica/goksen/

References

1. Appel,K.I., Haken,W., and Koch,J.: Every planar map is four colorable I: Discharing, Illinois J.Math, (1977)
2. Birkhoff,G.D. and Lewis,D.C.: Chromatic polinomials, Trans.Amer.Math.Soc., 60:355-451, (1946)
3. Bodlaender, H.L.: On the complexity of some coloring games, Lecture Notes in Computer Science, (1991)
4. Jonathan,G and Jay,Y.: Graph Theory and Its Applications,CRC Press, (1999)
5. Pemmaraju,S. and Skiena,S.: Computational Discrete Mathematics, Cambrigde Univ. Press, (2003)
6. Skiena,S.: Implementing Discrete Mathematics-Combinatorics and Graph Theory with Mathematica, Addison-Wesley Publishing Company,(1990)
7. Soaty,T.L., and Kainen,P.C.: The four color problem, Dover, New York, (1986)
8. Thulasiraman, K. and Swamy,M.N.S.: Graphs: Theory and Algorithms, John Wiley and Sons, Inc., (1992)
9. Ufuktepe, U., Bacak, G., and Beseri, T. : Graph Coloring with webMathematica, Lecture Notes in Computer Science, Springer-Verlag, (2003)
10. Vizing, V.G.: On an estimate of the chromatic class of a p-graph, Discret. Analiz, (1964)
11. Wickham,T.:webMathematica A User Guide,Wolfram Research, Inc., (2002)
12. Wolfram,S.:The Mathematica Book, Cambrigde Univ. Press, (1996)

Mathematica Applications on Time Scales

Ahmet Yantır and Ünal Ufuktepe

Izmir Institute of Technology, Department of Mathematics,
Urla, Izmir, Turkey
{ahmetyantir, unalufuktepe}@iyte.edu.tr

Abstract. Stefan Hilger introduced the calculus on time scales in order to unify continuous and discrete analysis in 1988. The study of dynamic equations is an active area of research since time scales unifies both discrete and continuous processes, besides many others. In this paper we give many examples on derivative and integration on time scales calculus with *Mathematica*. We conclude with solving the first order linear dynamic equation $N^{\Delta}(t) = N(t)$, and show that the solution is a generalized exponential function with *Mathematica*.

1 Introduction

Stefan Hilger introduced the calculus on time scales in order to unify continuous and discrete analysis in 1988. Some works in this direction are [1,3,6,7]. There are many applications on time scales in [2].

In this paper, first we give the definitions of time scale, forward and backward jump operators which are the required tools of time scale calculus. Time scale calculus as calculus of integers can be examined in two parts: Δ-derivative and Δ-integration, ∇-derivative and ∇-integration. In this study we only deal with Δ- derivative and Δ-integration. We give some applications on Δ-derivative and Δ-integration with *Mathematica*. We create TimeScale package to calculate the backward and forward jump operators and derivative of the functions defined a time scale. We give the Mathematica codes for Δ-integration. Similarly, ∇ integration can be done with *Mathematica*. We also give a brief introduction to linear dynamic equation and solve an application due to [2].

2 Basic Definitions and Mathematica Codes

A time scale (measure chain) $\mathbf{T}$, is a nonempty, closed subset of real numbers. The examples of time scales defined with *Mathematica* can be found in [9,10].

Let $\mathbf{T}$ be a time scale. The forward jump operator $\sigma : \mathbf{T} \to \mathbf{T}$ and the backward jump operator $\rho : \mathbf{T} \to \mathbf{T}$ are defined by $\sigma(t) = \inf\{s : s \in \mathbf{T}, s > t\}$ and $\rho(t) = \sup\{s : s \in \mathbf{T}, s < t\}$ for all $t \in \mathbf{T}$ respectively.

If $\mathbf{T} = \mathbf{R}$, then $\sigma(t) = t$, and $\rho(t) = t$ and if $\mathbf{T} = \mathbf{Z}$, then $\sigma(t) = t + 1$, and $\rho(t) = t - 1$. Also $\sigma(\max \mathbf{T}) = \max \mathbf{T}$ and $\rho(\min \mathbf{T}) = \min \mathbf{T}$.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 529–537, 2005.

t is called right dense if $\sigma(t) = t$, left dense if $\rho(t) = t$ and right scattered if $\sigma(t) > t$, left scattered if $\rho(t) < t$. If t is both left and right dense then t is called dense point, t is isolated point if it is both left and right scattered. The graininess function μ: $\mathbf{T} \to [0, \infty)$ is defined by $\mu(t) = \sigma(t) - t$.

We define the forward operator in two different ways, one for purely discrete sets and the other one for combination of discrete points and closed intervals as follows:

In[1]:= << TimeScale`
In[2]:= sigma[4, i/2]

In this command the first component stands for the point, and the second component stands for the rule of time scale. The values of σ for the same point are different for different time scales:

In[3]:= sigma[4, i^2**]**

In order to remove the deficiency we form another command to compute the value of σ for more general time scales as follows:

In[4]:= T = 5 ≤ x ≤ 7 || x = 15/2 || 9 ≤ x ≤ 11 || x = 12 ||
x = 18
In[5]:=sigma1[9]

Similarly, we can compute backward jump operator, the dual of forward jump operator, with as follows:

In[6]:= rho[4, i^2**]**
In[7]:= rho[$\frac{15}{2}$**]**

3 Derivative on Time Scales

In order to define the derivative on time scale we need the following set $\mathbf{T}^k$ which is derived from the time scale $\mathbf{T}$ as follows:

$$\mathbf{T}^k = \begin{cases} \mathbf{T} - \{\max \mathbf{T}\}, & \text{if } \max \mathbf{T} < \infty \text{ and } \max \mathbf{T} \text{ is right scattered;} \\ \mathbf{T}, & \text{otherwise.} \end{cases}$$

Definition 1. *Let* $f : \boldsymbol{T} \to \boldsymbol{R}$ *and* $t \in \boldsymbol{T}^k$. *If there exists a neighborhood* U_t *such that*

$$|f(\sigma(t)) - f(s) - a[\sigma(t) - s]| \leq \epsilon|\sigma(t) - s| \tag{1}$$

is satisfied for all t, $a \in \boldsymbol{R}$ *and* $s \in U_t$. *Then* f *is* Δ*-differentiable at the point* t *and* a *is called* Δ*-derivative of* f *at the point* t.

$$a = f^{\Delta}(t) = \begin{cases} \lim\limits_{s \to t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s}, & \text{if } \mu(t) = 0; \\ \frac{f(\sigma(t)) - f(t)}{\mu(t)}, & \text{if } \mu(t) > 0. \end{cases} \tag{2}$$

Δ-derivative is defined on $\mathbf{T}^k = \mathbf{T} - \{\max \mathbf{T}\}$, not on the whole time scale. If $t = \max \mathbf{T}$ then a neighborhood U_t of t contains only t. So the inequality (1) can be written only for $s = t$,

$$|f(\sigma(t)) - f(t) - a[\sigma(t) - t]| \leq \epsilon|\sigma(t) - t|$$
$$|f(t) - f(t) - a[t - t]| \leq \epsilon|t - t|$$
$$0 \leq 0$$

Thus the definition of Δ-derivative is satisfied for every value of a. Then a can not be determined uniquely.

Let $f : \mathbf{T} \to \mathbf{R}$ and $t \in \mathbf{T}^k$. The following two results can be obtained from formula (2):
If f is continuous at t and t is right scattered then f has a Δ-derivative at t and

$$f^{\Delta}(t) = \frac{f(\sigma(t)) - f(t)}{\mu(t)}. \tag{3}$$

If t is right dense then f has a Δ-derivative at t if and only if

$$\lim_{s \to t} \frac{f(t) - f(s)}{t - s}$$

is finite. Then

$$f^{\Delta}(t) = \lim_{s \to t} \frac{f(t) - f(s)}{t - s}. \tag{4}$$

If we unify the above results we get the useful formula, summarizing the Δ-derivative:

$$f(\sigma(t)) = f(t) + f^{\Delta}(t)\mu(t). \tag{5}$$

We can formulate the properties of Δ derivative of two Δ differentiable functions such as linearity, the product rule, and the quotient rule as follows:

$$(f + g)^{\Delta}(t) = f^{\Delta}(t) + g^{\Delta}(t) \tag{6}$$

$$(fg)^{\Delta}(t) = f^{\Delta}(t)g(t) + f(\sigma(t))g^{\Delta}(t) = f(t)g^{\Delta}(t) + f^{\Delta}(t)g(\sigma(t)). \tag{7}$$

$$(\frac{f}{g})^{\Delta}(t) = \frac{f^{\Delta}(t)g(t) - f(t)g^{\Delta}(t)}{g(t)g(\sigma(t))}. \tag{8}$$

Clearly $(1)^{\Delta} = 0$ and $(t)^{\Delta} = 1$ from (2). So by using formula (7) we can find $(t^2)^{\Delta} = (t \cdot t)^{\Delta} = t + \sigma(t)$. We can use (8) to compute

$$f^{\Delta}(t) = \frac{-2(t + \sigma(t))}{(t^2 - 1)((\sigma(t))^2 - 1)}$$

where $f(t) = \dfrac{t^2 + 1}{t^2 - 1}$.
For $\mathbf{T} = \mathbf{N}_0^{\frac{1}{2}} = \{\sqrt{n} : n \in \mathbf{N}_0\}$, if $t \in \mathbf{T}$ then $\sigma(t) = \sqrt{t^2 + 1}$. Thus
$f^{\Delta}(t) = \dfrac{-2(t + \sqrt{t^2 + 1})}{(t^2 - 1)t^2}$.

For $\mathbf{T} = \{\frac{n}{2} : n \in \mathbf{N}_0\}$, if $t \in \mathbf{T}$ then and $\sigma(t) = \frac{2t+1}{2}$. Thus $f^{\Delta}(t) = -\dfrac{4t + 1}{t^2 + t - \frac{3}{4}}$.

Further results and applications are included in [2,4].

We define this new derivative operator with *Mathematica* in two different ways, one for time scales which are the combination of discrete and closed intervals and the other for time scales including only isolated points as follows:

In[8]:= Clear[sigma, T]
In[9]:= f[t_]:= $\frac{t^2+1}{t^2-1}$
In[10]:=T=5 ≤ x ≤ 7 || x= $\frac{15}{2}$ **|| 9 ≤ x ≤ 11 || x=12 || x=18;**
In[11]:= Tderivative1[f, 7]

Tderivative2 determines the derivative of a function on a discrete time scale.

In[12]:= Clear[f, T, n, t, sigma, i, deriv]
In[13]:=Tderivative2[t^3**, 4,** 2^i**]**

4 Integration on Time Scale

Definition 2. *A function $f : \boldsymbol{T} \to \boldsymbol{R}$ is called regulated if it has finite right-sided limits at all right dense points in $\boldsymbol{T}$ and it has finite left-sided limits at all left dense points in $\boldsymbol{T}$.*

A function $f : \mathbf{T} \to \mathbf{R}$ is called rd-continuous if it is continuous at all right-dense points in $\mathbf{T}$ and its left-sided limits exist (finite) at all left-dense points in $\mathbf{T}$. The set of rd-continuous functions in $\mathbf{T}$ is denoted by C_{rd}.

Theorem 1. *(Existence of antiderivative) Every rd-continuous function has an antiderivative. In particular if $t_0 \in \boldsymbol{T}$, then F defined by*

$$F(t) := \int_{t_0}^{t} f(\tau)\Delta\tau \quad for \quad t \in \boldsymbol{T}$$

is an antiderivative of f.

Proof (4).

Theorem 2. *If $f \in C_{rd}$ and $t \in \boldsymbol{T}^k$, then*

$$\int_{t}^{\sigma(t)} f(\tau)\Delta\tau = (\sigma(t) - t)\mu(t)f(t).$$

Proof (4).

Theorem 3. *Let $a, b \in \boldsymbol{T}$ and $f \in C_{rd}$. If $[a, b]$ consists of only isolated points, then*

$$\int_a^b f(t)\Delta t = \begin{cases} \sum\limits_{t\in[a,b)} \mu(t)f(t) \ , & a < b \ ; \\ 0 & a = b; \\ -\sum\limits_{t\in[b,a)} \mu(t)f(t) \ , & a > b. \end{cases}$$

Proof. Assume that $a < b$ and let $[a, b] = \{t_0, t_1, t_2, ..., t_n\}$ where

$$a = t_0 < t_1 < t_2 < ... < t_n = b.$$

$$\begin{aligned}\int_a^b f(t)\Delta t &= \int_{t_0}^{t_1} f(t)\Delta t + \int_{t_1}^{t_2} f(t)\Delta t + ... + \int_{t_{n-1}}^{t_n} f(t)\Delta t \\ &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} f(t)\Delta t \\ &= \sum_{i=0}^{n-1} \int_{t_i}^{\sigma(t_i)} f(t)\Delta t \\ &= \sum_{i=0}^{n-1} (\sigma(t_i) - t_i) f(t_i) \\ &= \sum_{t\in[a,b)} (\sigma(t) - t) f(t)\end{aligned}$$

If $a > b$, by using the fact

$$\int_a^b f(t)\Delta t = -\int_b^a f(t)\Delta t$$

we obtain

$$\int_a^b f(t)\Delta t = -\sum_{t\in[b,a)} (\sigma(t) - t) f(t)$$

which is the desired result.

If f is defined on a time scale which consists of only isolated points then by using Theorem 3 it is possible to evaluate the integral as follows:

```
In[14]:= Clear[f, g, l, u];
In[15]:= Tintegrate1[f_ , g_ , l_ , u_ ] := Module[{T, s},
T= Table[g, {i, l, u}];
Print[" The time scale is= ", T];
s = Sum[(f /. x -> T[[i]])*(T[[i + 1]] - T[[i]]), {i, 1, u - l}];
Print[" The value of the integral= ", s]]
In[16]:= Tintegrate1[x^3,i/2, 1, 10]
Out[16]:= The time scale is = {1/2, 1, 3/2, 2, 5/2, 3, 7/2, 4, 9/2, 5}
The value of the integral = 2025/16
```

In this module f denotes the integrant, g denotes the rule of discrete time scale in terms of i, l denotes the lower bound, and u denotes the upper bound of the integral.

If f is defined on an arbitrary time scale which is the combination of closed intervals and discrete sets then we can evaluate the value of the integral:

```
In[17]:= A=Interval[{1,3},{5,5},{6,8},{12,13}]
In[18]:=Tintegrate2[f_, a_, b_]:=Module[{B, i, j, Sum1, Sum2},
B=IntervalIntersection[A,Interval[{a,b}]]; i=1;
Print[" The time scale is = " ,A];
Print[" The Integral Interval= " ,B];
While[B[[i]][[2]]≠ Max[B],i=i+1];Sum1=0;Sum2=0;
j=1;While[j≤i, Sum1=Sum1+Integrate[f,{x,B[[j]][[1]],B[[j]][[2]]}];
Print[" Ordinary integration from ",B[[j]][[1]]," to " ,B[[j]][[2]]];
j++;]; Print[" The value of ordinary integral=",Sum1]; j=1;
While[j<i, Sum2=Sum2+(f/.x->B[[j]][[2]])*(B[[j+1]][[1]]-B[[j]][[2]]);
Print[" Discrete integral from=",B[[j]][[2]]," to ",B[[j+1]][[1]]];
j++;]; Print[" The value of discrete integral=",Sum2];
Print[" The value of Time Scale integral=", Sum1+ Sum2 ];]
In[19]:=Tintegrate2[x^2, 2, 7]//N
```

Here f denotes the integrant, a denotes the lower bound, and b denotes the upper bound of the integral.

5 Linear Dynamic Equations and Applications

In this section we try to solve first order linear dynamic equations with *Mathematica*.

Definition 3. *A function $p : \boldsymbol{T} \to \boldsymbol{R}$ is called regressive if $1 + \mu(t)p(t) \neq 0$ is satisfied for all $t \in \boldsymbol{T}$.*

Concerning the initial value problems

$$y^{\Delta} = p(t)y, \quad y(t_0) = 1 \tag{9}$$

Hilger [6] proved the existence and uniqueness of (9) under the conditions p is rd-continuous and regressive. And (9) is called regressive if p is regressive.
The solution of (9) is called *the exponential function* and denoted by $e_p(\cdot, t_0)$. The explicit formula for $e_p(t, s)$ is given by

$$e_p(t, s) = exp\Big\{ \int_s^t \xi_{\mu(\tau)}(p(\tau))\Delta\tau \Big\} \tag{10}$$

where $\xi_h(z)$, the cylindrical transformation,

$$\xi_h(z) = \begin{cases} \dfrac{Log(1 + hz)}{h}, & \text{if } h \neq 0; \\ z, & \text{if } h = 0. \end{cases}$$

The properties of the exponential functions are given by Bohner and Peterson [4] and summarized in [3].

By using the exponential function the solutions of many dynamic equations can be found in a similar way as in ordinary differential equations. As an example by considering the dynamic equation

$$y^{\Delta^3} - 2y^{\Delta^2} + y^{\Delta} + 2y = 0$$

where $y^{\Delta^i},\ i = 2, 3$ denotes the i^{th} delta derivative of y and assuming that $y(t) = e_\lambda(t, t_0)$ with constant λ is the solution, we can find that the linear combination of $e_{-1}(t, t_0),\ e_1(t, t_0),\ e_2(t, t_0)$ is the solution of the dynamic equation above.

As another example we consider the initial value problem

$$y^{\Delta^2} = a^2 y,\ y(t_0) = 1,\ y^{\Delta}(t_0) = 0$$

where a is a regressive constant. By using the same technique we find that both $e_a(t, t_0),\ e_{-a}(t, t_0)$ and therefore $\alpha e_a(t, t_0) + \beta e_{-a}(t, t_0)$ solve the dynamic equation. By using initial conditions we find $\alpha = \beta = \frac{1}{2}$. Thus

$$y(t) = \frac{e_a(t, t_0) + e_{-a}(t, t_0)}{2}$$

is the solution.

If p and $-p$ are regressive and rd-continuous then the hyperbolic functions are defined by

$$\cosh_p = \frac{e_p + e_{-p}}{2},\ \sinh_p = \frac{e_p - e_{-p}}{2}.$$

Similarly the trigonometric functions are defined by the means of the solution of the initial value problem

$$y^{\Delta^2} = -a^2 y,\ y(t_0) = 1,\ y^{\Delta}(t_0) = 0.$$

The following example is due to Agarwal, Bohner, O'Regan and Peterson [3] and is an application of linear dynamic equation.

Example 1. Let $N(t)$ be the amount of plants of one particular kind at time t in a certain area. By experiments we know that according to $N' = N$ during the months of April until September. At the beginning of the October, all plants suddenly die, but seeds remain in the ground and start growing again at the beginning of April with N being doubled. We model this situation using the time scale

$$\mathbf{T} = \bigcup_{k=0}^{\infty} [2k,\ 2k+1]$$

where $t = 0$ is April 1 of the current year, $t = 1$ is the October 1 of the current year, $t = 2$ is the April 1 of the next year, $t = 3$ is October 1 of the next year, and so on. We have

$$\mu(t) = \begin{cases} 0, \text{ if } 2k \le t < 2k+1; \\ 1, \text{ if } t = 2k+1. \end{cases}$$

On $[2k,\ 2k+1)$ we have $N' = N$, i.e. $N^{\Delta} = N$. However we also know that $N(2k+2) = 2N(k+1)$, i.e. $N(2k+1) = N(2k+1)$, i.e. $N^{\Delta} = N$ at $2k+1$. As a result N is the solution of the dynamic equation $N^{\Delta} = N$. Thus, if N(0)=1 is given, N is exactly $e_1(t,0)$ on the time scale.

In order to solve this problem we give the following *Mathematica* codes:

```
In[20]:= int[i_]:=Interval[{2*i,2*i+1}];

In[21]:= A = IntervalUnion [int[0], int[1], int[2], int[3], int[4], int[5],
int[6], int[7], int[8], int[9]];
In[22]:= DynamicSolve[p_ , a_ , b_ ]:= Module[{B, i, j, Sum1,
Sum2 }, K = IntervalIntersection[A, Interval[{a, b}]]; i = 1;
B = Union[K, Interval[{b, b}]];
While[B[[i]][[2]] ≠ Max[B], i=i+1]; Sum1 = 1;
j=1; While[j ≤ i,
Sum1=Sum1*Exp[Integrate[p, {x, B[[j]][[1]], B[[j]][[2]]}]]; j++;];
j=1 ; Sum2=1;While[j<i, Sum2 = Sum2*
Exp[((p*Log[2]) /. x → B[[j]][[2]]) * (B[[j+1]][[1]]-B[[j]][[2]])]; j++;] ;
Print["The Solution of the dynamic equation =", Sum1*Sum2]; ]

In[23]:= DynamicSolve[1, 0, 10]
```

Here the first component of the command DynamicSolve is the function p of the dynamic equation $y^{\Delta} = p(t)y$, second is the initial point, and the third is the time at which we want to find the number of plants.
Another way of computing the number of plants without using the time scale integral is the following loop:

```
In[24]:= Clear[yo,f,g,a,b]
In[25]:= yo=1; a=0; b=1;
In[26]:= For[i=1,i<6,s=DSolve[{y'[x]=y[x],y[a]=yo},y[x],x] ;
f[x_]=y[x]/.s[[1]]; a=a+2; fd=RSolve[{z[x+1]=2 z[x],z[b]=f[a-1]}, z[x],
x]; g[x_]=z[x]/.fd[[1]]; b=b+2; yo=g[b-1]; i++]

In[27]:=yo
```

6 Conclusion

In this paper, we have given a very basic introduction to the time scales. Through a series of example, our goals were to present a variety of time scales, give the derivative, integral and dynamic equations on time scale. We compute some examples of each by *Mathematica*, and compare and contrast discrete and continuous calculus. But the symbolic computation in general for $\Delta-$derivative and the graphs of functions on time scales haven't done by *Mathematica*. It will be our future work.

References

1. Agarwal, R.P.& Bohner M., Basic calculus on time scale and some of its applications, Results Math. 35, (1-2) 3-22, (1999).
2. Agarwal, R.P., Bohner M., O'Regan, D.& Peterson, A, Dynamic equations on time Scales: A survey, Journal of Computational and Applied Mathematics, Vol.141, (1-2), 1-26, (2002).
3. Aulbach B.& Hilger, S., Linear dynamic process with in homogenous time scale in Nonlinear Dynamics and Quantum Dynamical Systems, Mathematical Research, Vol.59, Akademic Verlag Berlin, 9-20, (1999).
4. Bohner,M.& Peterson,A., Dynamic Equations on Time Scales, Birkhäuser Boston, (2001).
5. Gray,J.W., Mastering Mathematica, Academic Press, (1997).
6. Hilger, S.: Analysis on measure chains-a unified approach to continuous and discrete calculus, Results Math. 18, 18-56, (1990).
7. Hilger, S., Differential and difference calculus-unified!, Nonlinear Anal. Theory Methods Appl. 30 (5), 2683-2694, (1997).
8. Wolfram,S., The Mathematica, Book, Cambrigde Univ. Press, (1996).
9. Yantir, A. Ufuktepe Ü., Basic Calculus on Time Scale with Mathematica, Lecture Notes in Computer Science, 2567, 821-827, (2003).
10. Yantir, A.Derivative and Integration on Time Scale with Mathematica, Challenging The Boundaries of Symbolic Computation: Proceeding of 5 th International Mathematica Symposium, 325-331., (2003).

A Discrete Mathematics Package for Computer Science and Engineering Students

Mustafa Murat Inceoglu

Ege University, Department of Computer Engineering,
Bornova 35100 Izmir, Turkey
inceoglu@bornova.ege.edu.tr

Abstract. Discrete mathematics is one of the very basic mathematics courses in computer engineering (CE) and/or computer science (CS) departments. This course covers almost all of the basic concepts for many other courses in the curriculum and requires active learning of students. For this purpose, especially "propositions" concept, which cannot be understood well in prior years, should be covered with every detail. Previous studies show that learning by entertaining activities and competition has positive effect on student motivation. In this study, an Internet-based Discrete Mathematics Package (DMP) for "propositions" that can work in mobile devices and encourages competitive learning between students has been developed within related literature.

1 Introduction

It is clear that motivation has positive effects on learning. Especially courses with mathematical context require much more motivation because of the nature of the course. The competency between students is an important factor for student motivation and successful understanding of concepts. The students who are in a competitive environment should develop their skills more and work harder. An interactive and Internet-based discrete structures package that can increase student motivation has been developed in this study. The paper continues with the background of the study, then the application developed will be introduced and finally the results of the study will be presented and discussed.

2 Background

Discrete mathematics is very important in computer science and computer engineering branches. Discrete mathematics covers the basic introduction concepts for computer hardware working principles, programming languages, automata and finite state machines, data structures and databases, encryption and decryption methods. Logical propositions are very important in discrete mathematics science. Logical propositions are especially important for understanding mathematical inference, understanding the working principles and design of computer circuits, and validation of computer programs. Considering the importance of above explanations,

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 538 – 546, 2005.

the Discrete Mathematics Package (DMP) is very important for students to understand the operations over logical propositions.

In this study, propositions are taught by a real-time competitive tournament technique placed in a symbolic mathematics package. Tournament structure has been discussed in previous studies [1, 2, 3]. The importance of competition in different learning methods has been emphasized in [4].

DMP is a package that has been developed for students to make them understand logical propositions better, evaluate themselves in quizzes and increase their knowledge within a competitive environment. The package supports 4 logical parameters that are named as *p*, *q*, *r* and *s*; 8 logical operators that are named as AND, OR, NOT, XOR, EQU, IMP, NAND and NOR; and 2 logical constants named as TRUE (T) and FALSE (F). Parenthesis can also be used to alter the priorities of logical expressions.

The research on literature implies five applications developed for creating truth tables and other logical functions.

The first application has been realized by [5] using applets of Java programming language. This program allows the usage of maximum 26 logical parameters, 8 logical operators AND, OR, NOT, CONDITIONAL, BI-CONDITIONAL, XOR, NOT, NAND and NOR and 2 logical constants. The program can run on any Internet browser. As explained in [5] no controller has been used for the interpretation of logical expressions and it has been assumed that the user writes the correct expressions. The usage of non-standard symbols and the obligation to write the correct logical expressions are the weak sides of this software.

The second program that has been analyzed also works as a Java applet. In addition to its truth table creation option, the program also has logical inference capability. The program consists of 5 logical operators AND, OR, NOT, IMPLICATION and EQUIVALENCE, 2 logical constants; and 3 logical parameters p, q and r. The user again has to enter the correct expression in this software [6].

The third program, which is also another Java applet, is both an integer calculator and a logical computer [7]. The program has 5 logical operators NOT, AND, OR, IMPLICATION and BIJECTION, 2 logical constants 0 (False) and 1 (True). The program accepts logical expression inputs from the user and returns the truth table of the related logical expression to the user. The user is responsible for checking the text of the logical expression.

The fourth program [8] has been written in JavaScript language and uses 5 logical operators NOT, AND, OR, IMPLICATION and EQUIVALENCE, and a maximum of 10 logical parameters. The program accepts pre-controlled text inputs as logical expressions and returns a truth table corresponding to related logical expression. This program also depends on the user for text control.

The fifth application that is written in C++ programming language has been developed for personal digital assistant (PDA) devices and supports a maximum of 4 parameters. It is a Boolean function simplifier [9].

Logic Works [12], Pspice Mixed Mode Simulator [13], Micro-Cap [14], Win-Breadboard [15] and Electronic Workbench [16] packages are both used for the reduction of logical expressions and for simulating the logical circuits. However, these programs are primarily developed to be used in digital electronics courses.

3 The Package

When DMP starts, the client establishes connection to *USER* database on server, so all of the actions of the students can be tracked. The database stores the actions of students and every successful or unsuccessful operation in order to extract statistical information from the system. The statistical information gives clues for the enhancements of the program and gives valuable information about the errors and weak sides of the students.

Discrete Mathematics program Package has two main modules that are text processing engine and realization engine.

Text Processing Engine: It is a program segment that has been developed to control the logical expressions written in INPUT-1 and/or INPUT-2 lines for whether they are correct or not. The number of parenthesis in logical expressions; the correct use of logical operators, logical propositions, logical constants, the number of inputs in one or two input lines and finally the usage errors are being controlled. If there is an error in the text, the cause of this error is displayed on RESULT/ERROR line. If there is no error in text, nothing is displayed on RESULT/ERROR line.

Realization Engine: When error-free logical expression, logical parameters, operators and constants passes successfully through the text processing engine, the expression steps are parsed and loaded on a realization tree on this program segment. The realization tree is based on inorder algorithm [17].

The introduction of command buttons in the DMP is given as follows.

SYNTAX CHECK button controls the text in the INPUT-1 line via text processing engine.

TAUTOLOGY button first controls the text in INPUT-1 line via text processing engine. If the logical expression is correct, the expression is passed through the realization engine and the result of the expression is checked whether it is a TAUTOLOGY or not. If the result is tautology, then TAUTOLOGY output is displayed on the RESULT/ERROR line and the truth table is displayed.

EQUIVALENT button checks the expression in INPUT-1 and INPUT-2 lines via text processing engine and both expressions are passed through the realization engine and tested for logically equality. If the truth tables of both expressions are the same as the output of the realization engine, EQUAL is displayed on the RESULT/ERROR line. If they are not equal, then NOT EQUAL is displayed.

CONTRADICTION button checks the logical expression in INPUT-1 line via text processing engine and if the logical expression is correct, this expression is passed through the realization engine and checked whether the result of the expression is a contradiction or not. If the result is contradiction, CONTRADICTION is displayed on the RESULT/ERROR line and the truth table of the expression is displayed at the same time.

ENGLISH (or TURKISH) button converts every Turkish word in the program screen to corresponding English words and the ENGLISH button itself turns to TURKISH.

QUIZ button displays the screen given in Figure-1 and three questions are asked to the user. The first question is about finding the result of a logical expression drawn from Q_BANK database. TAUTOLOGY, CONTRADICTION and OTHER choices are available for student depending on the result of the logical expression (If the result of the expression is not a tautology or contradiction, OTHER button can be selected)

The second question is about the simplification of a logical expression drawn from the *Q_BANK* database. All of the logical expressions and answer choices are drawn from the *Q_BANK* database.

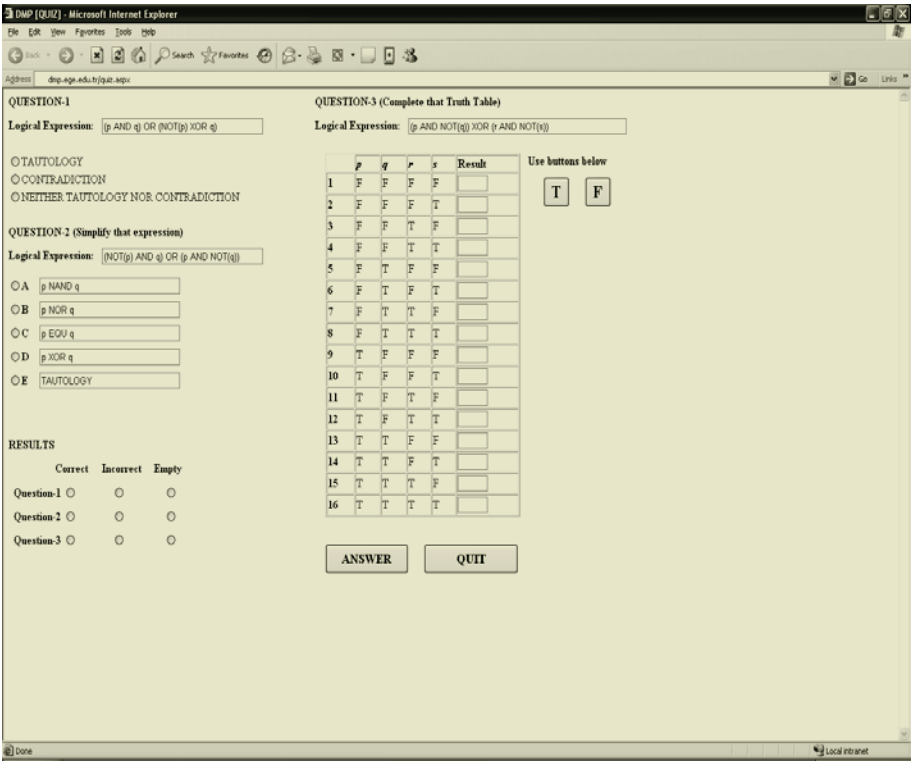


Fig. 1. QUIZ window

The last question is about the truth table of a logical expression drawn from *Q_BANK* database. A maximum of 4 logical parameters are allowed in this part of package and a maximum of 16 slots that are labeled from 1 to 16 expect inputs from the user as logical combination results by pressing the **T** and **F** buttons. If the number of logical parameters is fewer than 4 (2 combinations for single logical parameter, 4

combinations for 2 logical parameters, and 8 combinations for 3 logical parameters) no input is allowed on unused slots.

When the user in quiz stage, (s)he cannot use any functions of the main program. It is important the questions should be replied correctly in the quiz stage, so unless the **ANSWER** button is pressed, the user can change his/her answers at anytime. When the user presses the **ANSWER** button, the three answers are compared with the answers in the *Q_BANK* database and the choice button in **RESULTS** window is marked for every correct answer. If the user does not want to answer any questions or (s)he just presses the **ANSWER** button without using the keyboard, this answer will be evaluated as blank.

TRUTH TABLE button controls the expression in INPUT-1 line via text processing engine and the corresponding truth table is displayed at TRUTH TABLE section by passing through the realization engine. The synonyms like A and B are listed in synonyms section. There can be a maximum of 6 synonyms (from **A** to **F**). An example about the truth table application is given in Figure-2.

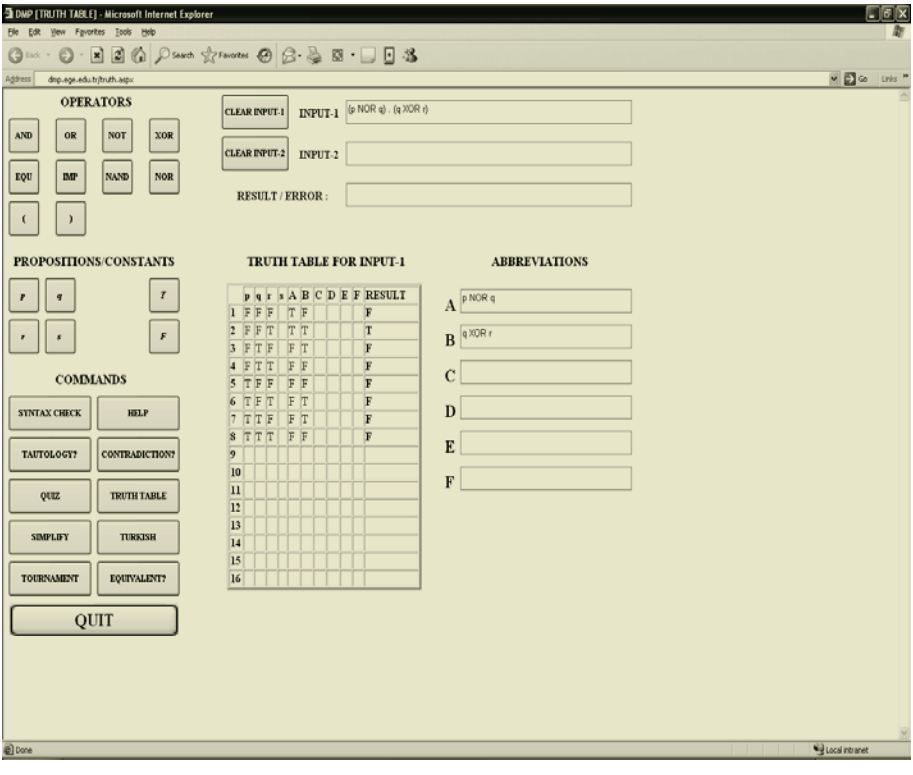


Fig. 2. TRUTH TABLE window

SIMPLIFY button checks the logical expression in INPUT-1 line via text processing engine and makes an induction for a maximum of 4 logical parameters based on

Karnaugh Map [18] method. The simplified result is then displayed on the RESULT/ERROR line. The result of the simplification is achieved by summation of multiplications as minterm expressions [10, 11, 18] and only AND, OR, and NOT logical operators are used in the result of this operation.

Discrete Mathematics program Package helps user to reach the truth tables of logical operators via **HELP** button.

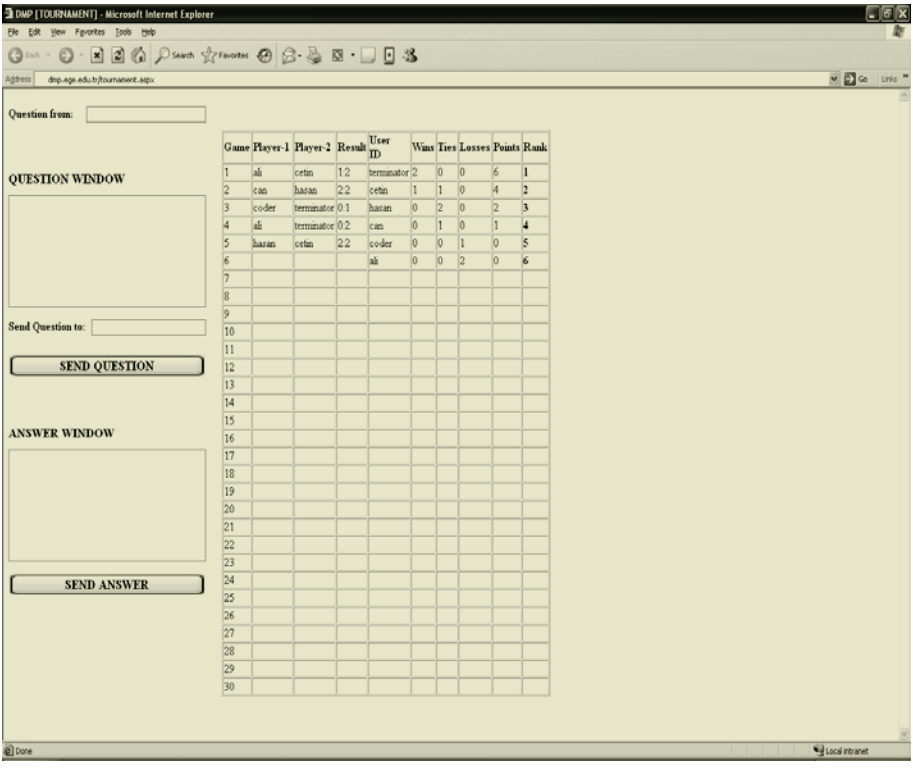
The package defined above is the first version of DMP developed at the end of 2003 nad it is named as DMP1. The pedagogic results and other details about the DMP1 package is available on [19]. With the evaluation of feedback from students and the results from the *USER* database, whole package has been revisioned in July 2004 and named as DMP2. DMP2 package covers all of the functions of DMP1 package. The difference between DMP1 and DMP2 packages are explained below.

DMP1 is developed as a desktop application using Microsoft Visual Basic and Access Database. In order to make design and programming easier, DMP2 package is planned and developed with Microsoft Visual Basic.NET, ASP.NET and SQL Server in 2-tier software architecture (first tier is database server, and second tier is web server). DMP2 can be used via Internet and can work as a mobile application on mobile phones.

While DMP1 uses Karnaugh Map technique for simplification, Quinn-McClusky technique [18] has been used in DMP2 to increase simplification performance. However, in addition to these improvements, the 2-bit ranged combinations, which cannot be simplified by Karnaugh Map or Quinn-McClusky methods that are based on 1-bit ranged binary gray code, are simplifiend by XOR and EQU operators. For example;

$F(p,q,r,s) = \overline{p}.\overline{q}.\overline{r}.\overline{s} + \overline{p}.q.\overline{r}.s. + \overline{p}.\overline{q}.r.s + \overline{p}.q.r.\overline{s}$ function cannot be simplified by Karnaugh Map and Quinn-McClusky methods, but it can be turned into $F(p,q,r,s) = \overline{p}.\overline{q}.(r \otimes s) + \overline{p}.q.(r \oplus s)$ form with our algorithm. In this example, '.' has been used for AND, '$\overline{p}$' for NOT, '+' for OR, '$\oplus$' for XOR and '$\otimes$' for EQU operators. As a result, while there are 25 logical operators in the first function, there are 10 logical operators in the simplified function.

There is an enhancement to the DMP1 package in the DMP2 package which is based on interaction between students, and it is named as TOURNAMENT. TOURNAMENT has been designed to make students learn better by means of competency. The aim of tournament is to make students write correct logical expression, simplify logical expressions, find results of logical expressions by means of give propositions or calculate the result of logical expression as tautology or contradiction. In this system, two or more students can work in pairs and send each other questions. The question-answer relation between the *sender* of the question and the *receiver* of the question is named as match and both the *sender* and *receiver* can achieve points as a result of the match. TOURNAMENT software also requires the *sender* of the question to answer the same question. The software controls that maximum 3 questions can be sent from a user to another user within a specific time period. The match results are stored in database as the question sent and the answers of the *sender* and *receiver*, then the related data is evaluated by the teaching



Game	Player-1	Player-2	Result	User ID	Wins	Ties	Losses	Points	Rank
1	ali	cetin	1:2	terminator	2	0	0	6	1
2	can	hasan	2:2	cetin	1	1	0	4	2
3	coder	terminator	0:1	hasan	0	2	0	2	3
4	ali	terminator	0:2	can	0	1	0	1	4
5	hasan	cetin	2:2	coder	0	0	1	0	5
6				ali	0	0	2	0	6
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									

Fig. 3. TOURNAMENT window

assistant. If the *sender* has the right answer for his/her question and the *receiver* has a wrong answer, the *sender* receives 3 points. If the *sender* has a wrong answer and the *receiver* has the right answer, the *receiver* receives 3 points. If they both have the correct answer, they both get 1 point. If they both have wrong answers, they get no points from that question. The final points table can be used for grading of Discrete Mathematics Laboratory Work. Considering that students may have different number of matches, DMP2 creates the order of the points table by ranking the student with the highest points/match ratio on top (Figure-3).

4 Conclusions

DMP1 has been tested on high school students in fall semesters of 2003 and 2004, and it has been proved that it has positive effects on overall course grades [19].

After the development of DMP2, it has been tested on 20 first-year computer engineering students on October 2004, and positive feedbacks are received from students who use this system. DMP2 not only tested as a desktop and Internet application but as a pocket PC, tablet PC, mobile phone application.

The effectiveness of DMP2 was evaluated with a questionnaire. The users of DMP2 gave a score between one and five to each of five questions, with one being the lowest, and five being the highest. The average of score was 4.4. Table 1 shows the results of the questionnaire. According to question (1), some of the users felt that the questions were difficult. Question (2) shows that DMP2 system is so easy to use. In terms of discrete mathematics learning, question (3) shows that DMP2 was quite useful for it. The questions (4) and (5) show the users were very interested in this system, and that they would like to keep using it.

Table 1. The results of questionnaries

Question No	Questionnarie	Average	St.Dev.
1	Were the questions provided by DMP2 difficult?	3.8	0.53
2	Do you think DMP2 easy to use?	4.5	0.45
3	Do you think DMP2 useful for discrete mathematics learning?	4.5	0.38
4	Do you think DMP2 interesting?	4.7	0.53
5	Do you want to keep using this system?	4.3	0.89

DMP2 is an easy-to-use package (Table-1, Question 2, 3) that has been developed for deep understanding of the propositions concept which is an important topic in discrete mathematics course and computer science. DMP2 has much more enhanced features over DMP1 because of its Internet and mobile devices support.

Considering the feedbacks from the students, the next version of DMP2 is planned to be implemented in Java platform in order to have an open-source coding of the system and its source code can be shared over Internet.

Using the properties of Java platform, DMP2 is also planned be enchanced to cooperate with widely used Mathematica and Matlab packages, and a question bank and evaluation system for graph structures will be developed.

References

[1] Adams, J.: Change-It: An Object-Oriented Capstone Project for CS-1. Proc. 29th ACM Special Interest Group on Computer Science Education (SIGCSE) Technical Symposium. Computer Science Education (1998) 10-14

[2] Reese, D.: Using Multiplayer Games to Teach Interprocess Communication Mechanism. ACM Special Interest Group on Computer Science Education (SIGCSE) Bull. 32 (2000) 45-57

[3] Lawrance R.: Teaching Data Structures Using Competitive Games. IEEE Transaction on Education. 47 (2004) 459-466

[4] Felder R, Silverman, L.: Learning and Teaching Styles in Engineering Education. Engineering Education. 78 (1988) 674-681

[5] Triscari, J. (2004): Truth Table Generator Applet, [Online]. Available at: www.cs.uwm.edu/~triscari/Truth/Truth.html
[6] Oursland, M. (2004): Propositional Logic Applet, [Online]. Available at: www.oursland.net/Aima/Propositionapplet.html
[7] McPhail, B. (2004): Calculators, [Online]. Available at: www.reed.edu/~mcphailb/applets/calc
[8] Siverberg, E. (2004): Hey Kids! It's the Magic Truth Table Generator, [Online]. Available at: xenon.stanford.edu/~silver/truth
[9] Bitincka, L, Antoniu, GE, PDA-based Boolean Function Simplification: A Useful Educational Tool. Informatica. 15 (2004) 329-336
[10] Rosen, K.: Discrete Mathematics and Its Applications. 3rd edn. McGraw Hill, Singapore (1994)
[11] Johnsonbaugh, R.: Discrete Mathematics, 4th edn. Prentice Hall, New Jersey (1997)
[12] Capilano Computing Systems, Logic Works v3.0. Benjamin Cummings, Redwood City, CA, 1995
[13] Orcad Corporation (2004): WinPspice Mixed Mode Simulator, [Online]. Available at: www.orcad.com
[14] Thompson, A, O'Brien, T, Steele, B (2004): Micro-Cap V, [Online]. Available at: www.spectrum-soft.com
[15] Yoeric Software (2004): WinBreadboard, [Online]. Available at: www.yoeric.com
[16] Interactive Image Technologies (2004): Electronics Workbench, [Online]. Available at: www.interactiv.com
[17] Tremblay, A, Sorenson, K.: An Introduction to Data Structures with Applications. McGraw Hill, Japan (1976)
[18] Mano, M.: Digital Logic and Computer Design. Prentice Hall, New York (1979)
[19] Inceoglu, M, Discrete Structures Program Package (DSPP) and An Application, Turkish Online Journal of Educational Technology 3 (2004) Article 12, [Online]. Available at: www.tojet.net

Circle Inversion of Two-Dimensional Objects with Mathematica

R.T. Urbina[1] and A. Iglesias[2,3]

[1] Department of Mathematics, National University of Piura, Piura, Perú
teourg@hotmail.com
[2] Department of Computer Science, University of Tsukuba,
Laboratory of Advanced Research, Building B, Room # 1025,
Kaede Dori, 305-8573, Tsukuba, Japan
[3] Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n E-39005, Santander, Spain
iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. One of the most interesting and less known two-dimensional transformations is the so-called *circle inversion*. This paper presents a new Mathematica package, `CircleInversion`, for computing and displaying the images of two-dimensional objects (such as curves, polygons, etc.) by the circle inversion transformation. Those objects can be described symbolically in either parametric or implicit form. The output obtained is consistent with Mathematica's notation and results. The performance of the package is discussed by means of several illustrative and interesting examples.

1 Introduction

One of the most interesting and less known two-dimensional transformations is the so-called *circle inversion*. This transformation is quite reminescent of reflection with respect to a line. Under this last transformation, two points P and P' are symmetric with respect to a straight line if they lie at equal distances to the line on a perpendicular to that line. Therefore, any circle with center on the line that passes through one of the points passes necessarily through the other. The circle inversion is a very similar transformation, obtained by replacing the word "line" with the word "circle". This sort of inversion was first systematically investigated by the famous mathematician Jakob Steiner sometime about 1830. The interested reader is also referred to [1, 6, 8, 12, 13] for historical works on this issue, and to [2, 3, 4, 5, 7, 9, 10, 11, 14, 16] for further information on circle inversion.

This paper presents a new Mathematica package, `CircleInversion`, for computing and displaying the images of two-dimensional objects (such as curves, polygons, etc.) by the circle inversion transformation. Those objects can be described symbolically in either parametric or implicit form. The output obtained

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 547–555, 2005.

is consistent with Mathematica's notation and results. The performance of the package is discussed by means of several illustrative and interesting examples.

The structure of this paper is as follows: Section 2 introduced some mathematical definitions about the circle inversion as well as some interesting properties to be used throughout the paper. Then, Section 3 introduces the new Mathematica package, `CircleInversion`, and describes the commands implemented within. The performance of the package is also discussed in this section by using some illustrative examples for parametric and implicit surfaces. Finally, Section 4 closes with the main conclusions of this paper and some future work.

2 Mathematical Preliminaries

Let there be a circle C with center at O and radius R. And let there be a point P that could be located anywhere in the plane, except at the center O. We then apply the circle inversion transformation with respect to circle C, denoted as $CI(C)$, on this point P as follows (see Figure 1): consider the line through O and P. $CI(C)$ maps P to another point, P', on this line with the following property:

$$d(O,P).d(O,P') = R^2 \tag{1}$$

where d denotes the Euclidean distance on the plane. That is, the distance from the origin to P times the distance from the origin to P' is equal to the radius squared. And O, P and P' are colinear. Note that the point P' defined as above is unique, so we call it the *inverse* of point P. Similarly, the point P' maps to P. So these points are said to be *symmetric* in the circle, or to be *reflections* of each other. Also, if P and P' are inverse points each other, then the line L through P and perpendicular to OP is sometimes called a *polar* with respect to point P, known as the *inversion pole*. In addition, the curve to which a given curve is transformed under inversion is called its *inverse curve*. All these definitions will be used throughout the paper.

Note also that this transformation CI is defined in the entire Euclidean plane. Except of the center of C, which is known as the *center of inversion*, all points P in the plane have indeed inverse images P'. The center of inversion is often

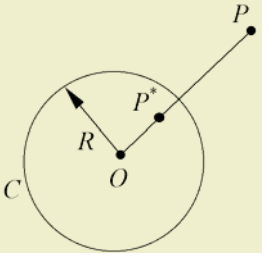


Fig. 1. Graphical interpretation of the circle inversion transformation

left over as a point with no inverse image, but sometimes is said to be mapped to the point at infinity [3, 4, 5]. Caution must be exercised as this is not the same point at infinity where, for example, all parallel lines meet.

The inversion $CI(C)$ maps the points inside C to the points in its exterior, and vice versa. The points on C are fixed under that transformation, and so are the circles perpendicular to it. Lines through the center of inversion (with the exception of the center itself) also have that property. The curves that coincide with their own inverse are called *anallagmatic*. On the other hand, observe that CI conserves angles. That is, $\angle POQ = \angle P'OQ'$. Such a map is called *conformal*.

2.1 Some Relevant Properties

The following properties regarding the inversion of lines and circles hold (see, for instance, [5, 8, 14] for details):

1. A line will always invert to either another line or a circle. In particular, a line through the center of the circle of inversion inverts to itself. On the contrary, a line that does not go through the center of the circle of inversion inverts to a circle which does go through the center of the circle of inversion.
2. A line that is completely outside the circle of inversion inverts to a circle (that goes through the center of the circle of inversion) that is completely inside the circle of inversion. Furthermore, the circle will have a radius which is strictly less than one-half the radius of the circle of inversion.
3. A line that is tangent to the circle of inversion inverts to a circle (that goes through the center of the circle of inversion) that is also tangent to the circle of inversion at the same point of tangency as the original line. Furthermore, the circle will have a radius which is exactly one-half the radius of the circle of inversion.
4. A line that intersects the circle of inversion in two points, but does not go through the center of the circle of inversion, inverts to a circle (that goes through the center of the circle of inversion) that intersects the circle of inversion at the same two points as the original line. Furthermore, the circle will have a radius which strictly greater than one-half the radius of the circle of inversion.
5. A line that is the perpendicular bisector of the radius of the circle of inversion inverts to a circle that is congruent to the circle of inversion.
6. Assume that the original line does not go through the center of the circle of inversion. A line parallel to the original line that does go through the center of the circle of inversion will be tangent to the circle created by inverting the original line.
7. A circle will always invert to either another circle or a line. In particular, a circle that is orthogonal (the tangent lines are perpendicular) to the circle of inversion will invert to itself.
8. A circle whose center is also the center of the circle of inversion inverts to a circle that also has the same center. Furthermore, if the original circle is also congruent to the circle of inversion, then it will invert to itself.
9. A circle that contains the center of the circle of inversion inverts to a line that does not contain the center of the circle of inversion.

10. A circle that is completely inside the circle of inversion inverts to either a circle or a line that is completely outside the circle of inversion. Conversely, a circle that is completely outside the circle of inversion inverts to a circle that is completely inside the circle of inversion.

3 The Package `CircleInversion`: Some Illustrative Examples

This section describes some examples of the application of this package to the cases of two-dimensional curves defined in either parametric or implicit form. Firstly, we need to load the package:

```
In[1]:=<<CircleInversion‘
```

3.1 Inversion of Parametric Curves

In this section we consider two-dimensional objects given in parametric form. The first example is given by the ellipse of parametric equation $x = cos(t)\ \ , y = 2\ sin(t)$ with $t \in [0, 2\pi]$:

```
In[2]:= ellipse= {Cos[t],2*Sin[t]};
```

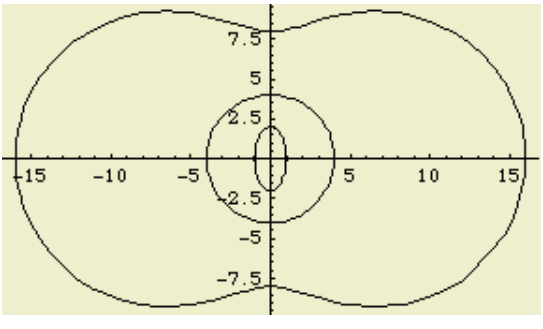


Fig. 2. Example of circle inversion: original and inverted ellipses and cicle of inversion. Note that the original ellipse is inside the circle whereas its inverted image is outside the circle

The `ParametricCircleInversion` command allows us to calculate the circle inversion of the ellipse given in parametric form. In this example, we employ the circle centered at the origin of radius $R = 4$:

```
In[3]:= ellipseinverted=ParametricCircleInversion[%,Radius->4];
```

Figure 2 displays the original and inverted ellipses as well as the circle used for the inversion:

```
In[4]:= ParametricPlot[{%,%%,CircleInversion[t,CircleRadius->4]}//
            Evaluate,{t,0,2 Pi},AspectRatio->Automatic]
```

$Out[4] :=$ *See Figure* 2

As the reader can see, the circle inversion exchanges the interior of the circle with the exterior of the circle. In other words, if P is a point inside of the circle C, then the inverse point P' is outside of circle C. The intuitive proof is that if P is inside circle C, then $d(O,P) < R$. Thus, in order for the definition of inversion (1) to be satisfied, $d(O,P') > R$. Thus, P' is outside of circle C.

One interesting feature is that, in addition to obtaining the graphical representation of the inverted ellipse, its parametric equations can also be obtained:

```
In[5]:= ellipseinverted
```

$$Out[5] := \left\{ \frac{16\ Cos[t]}{Cos^2[t] + 4\ Sin^2[t]}, \frac{32\ Sin[t]}{Cos^2[t] + 4\ Sin^2[t]} \right\}$$

Another example is given by a polygon of vertices:

```
In[6]:= p={{-1,0},{0,1},{1,0},{0,-1}};
```

The equations of its sides are given by:

```
In[7]:= Rb = Polygon[p,t]
```

$$Out[7] := \{\{-1+t, t\}, \{t, 1-t\}, \{1-t, -t\}, \{-t, -1+t\}\}$$

Applying the circle inversion transformation with respect to the unit circle centered at the origin we get:

```
In[8]:= RbInv = PolygonInversion[p,t]
```

$$Out[8] := \left\{ \left\{ \frac{-1+t}{(-1+t)^2 + t^2}, \frac{t}{(-1+t)^2 + t^2} \right\}, \left\{ \frac{t}{(1-t)^2 + t^2}, \frac{1-t}{(1-t)^2 + t^2} \right\}, \right.$$
$$\left. \left\{ \frac{1-t}{(1-t)^2 + t^2}, -\left(\frac{t}{(1-t)^2 + t^2} \right) \right\}, \left\{ -\left(\frac{t}{(-1+t)^2 + t^2} \right), \frac{-1+t}{(-1+t)^2 + t^2} \right\} \right\}$$

Finally, we display the original polygon and its inverse:

```
In[9]:= Show[ParametricPlot[# // Evaluate,{t,0,1}]& /@ {Rb,RbInv},
AspectRatio->Automatic];
```

$Out[9] :=$ *See Figure* 3

This example illustrates properties 1 and 4 (see Section 2.1 for details).

3.2 Inversion of Implicit Curves

The `ImplicitCircleInversion` command allows us to compute the circle inversion of any two-dimensional object given in implicit form. For example, the Lamé curve of fourth degree:

```
In[10]:=ImplicitPlot[x^ 4+y^4==1,{x,-1,1},{y,-1,1},AxesOrigin->{0,0}];
```

$Out[10] :=$ *See Figure* 4

Now, we compute the circle inverse with respect to the circle centered at the origin and of radius $\sqrt{2}$:

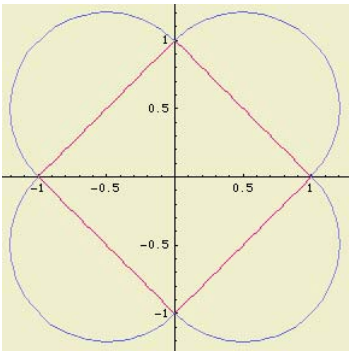


Fig. 3. Original four-sided polygon and its inverse by the circle inversion transformation. Note that the lines are mapped onto curves that pass through the intersection points of the lines with the unit circle (the polygon corners in this example)

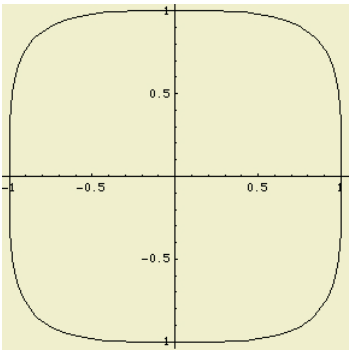


Fig. 4. Lamé curve of fourth degree

```
In[11]:= invlame1=ImplicitCircleInversion[x^4+y^4==1,{x,y},
                        CircleCenter->{0,0},CircleRadius->Sqrt[2]];
```

$$Out[11] := \frac{16\ x^4}{\left(x^2+y^2\right)^4} + \frac{16\ y^4}{\left(x^2+y^2\right)^4} == 1$$

as well as with respect to the circle of center at $(2, 2)$ and radius 4:

```
In[12]:= invlame2=ImplicitCircleInversion[x^4+y^4==1,{x,y},
                        CircleCenter->{2,2},CircleRadius->4];
```

$$Out[12] := \left(2 + \frac{16\ (-2+x)}{(-2+x)^2+(-2+y)^2}\right)^4 + \left(2 + \frac{16\ (-2+y)}{(-2+x)^2+(-2+y)^2}\right)^4 == 1$$

Figure 5 shows the graphical representation of such inverse for both cases.

Another nice example is given by the lemniscata curve, displayed in Figure 6(left). The inverses of that curve with respect to the unit circles of center at $(1, 0)$ and $\left(0, -\frac{1}{3}\right)$ are displayed. The corresponding equations are given by:

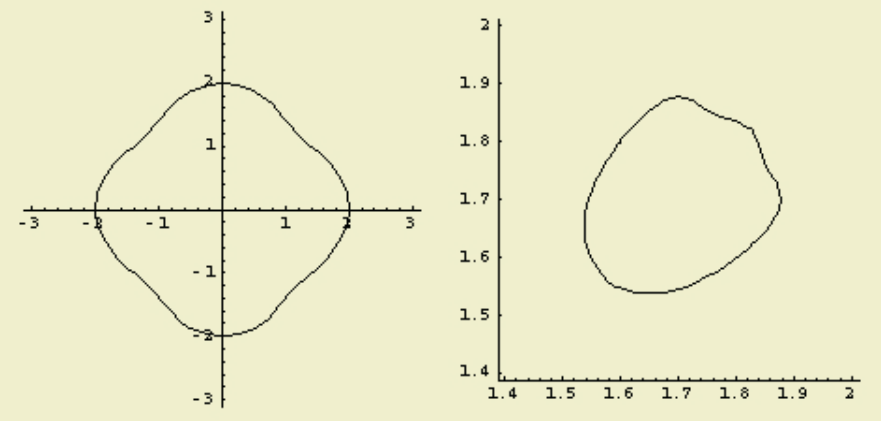


Fig. 5. Circle inversion of the Lamé curve of fourth degree displayed in Figure 4 with respect to the circle: (left) centered at the origin and of radius $\sqrt{2}$; (right) of center at $(2, 2)$ and radius 4

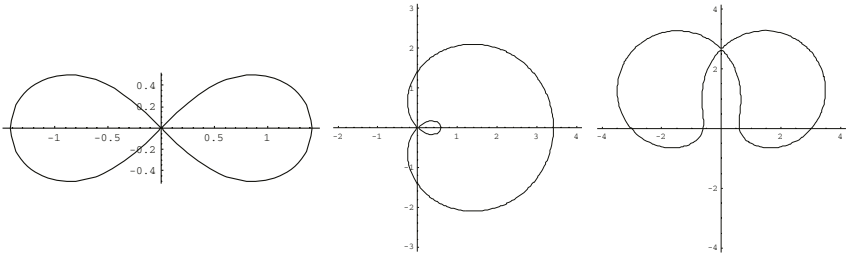


Fig. 6. (left) the lemniscata curve; (middle) inverse with respect to the unit circle centered at $(1, 0)$; (right) inverse with respect to the unit circle centered at $\left(0, -\frac{1}{3}\right)$

```
In[13]:= lemniscata= ((x+1)^2+y^2)((x-1)^2+y^2)==1;
In[14]:= Simplify[ImplicitCircleInversion[lemniscata,{x,y},
           CircleCenter->#,CircleRadius->1]]& /@ {{1,0},{0,-1/3}}
```

$$Out[14] := \left\{ \frac{1 - 4\,x + 4\,x^2 + 4\,y^2}{\left(1 - 2\,x + x^2 + y^2\right)^2} == 1, \right.$$

$$\left. \frac{\left(73 - 162\,x + 90\,x^2 + 6\,y + 90\,y^2\right)\,\left(73 + 162\,x + 90\,x^2 + 6\,y + 90\,y^2\right)}{81\left(9\,x^2 + (1 + 3\,y)^2\right)^2} == 1 \right\}$$

Note that the first equation corresponds to the well-known Cassini's oval curve. At its turn, the inverse of such a curve is obviously the Lemniscata curve:

```
In[15]:=ImplicitCircleInversion[(1-4 x+4 x^2+4 y^2)/(1-2x+x^2+y^2)^2==1,
                {x,y},CircleCenter->{1,0},CircleRadius->1]
```

$$Out[15] := \left\{\left(1 - 2\,x + x^2 + y^2\right)\,\left(1 + 2\,x + x^2 + y^2\right) = 1\right\}$$

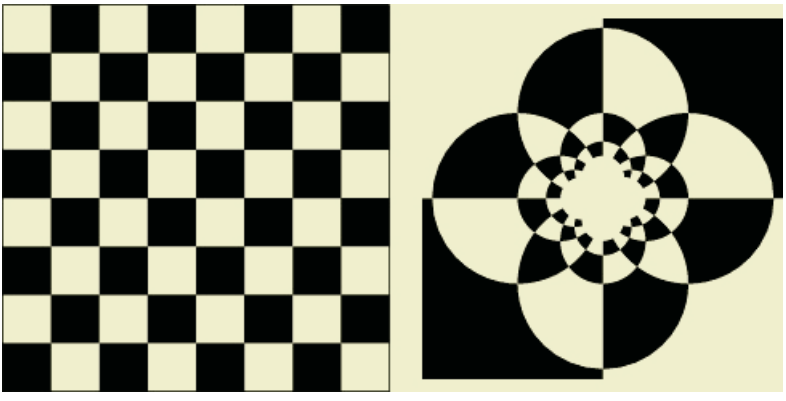


Fig. 7. Example of circle inversion transformation: (left) original chessboard image; (right) its inverse

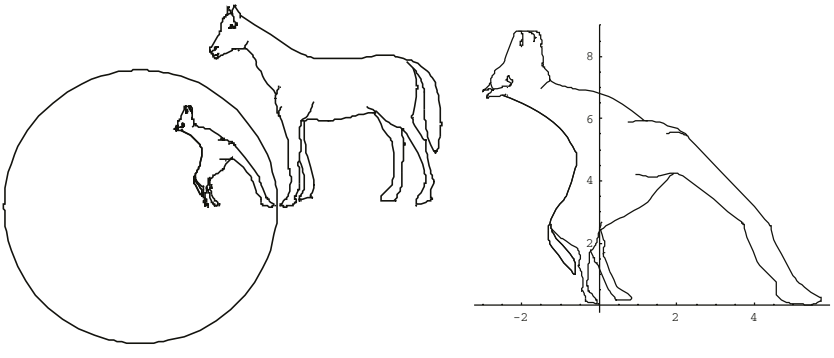


Fig. 8. Example of circle inversion transformation: (left) the original horse outside the circle is mapped onto that inside the circle ; (right) close up of the inverted image

3.3 Other Examples

In this section other examples of the application of our package are described. The first example is displayed in Figure 7: the figure on the right has been obtained by applying the circle inversion transformation onto the chessboard depicted on the left.

The second example is depicted in Figure 8: the horse outside the circle on the left figure is inverted. As remarked above, the mapped image lies inside the circle. On the right, we show a close up of such mapped image.

4 Conclusions and Future Work

In this paper, a new Mathematica package, `CircleInversion`, for computing and displaying the images of two-dimensional objects (such as curves, polygons, etc.) by the circle inversion transformation is introduced. The objects can be de-

scribed symbolically in either parametric or implicit form. The output obtained is consistent with Mathematica's notation and results. The performance of the package is discussed by means of several illustrative and interesting examples.

All the commands have been implemented in Mathematica version 5 [17]. The symbolic capabilities and the powerful Mathematica functional programming [15] features have been extensively used to make the program shorter and more efficient. Future lines of research include the improvement of the present package and the extension of all commands to the three-dimensional case.

References

1. Casey, J.: Theory of Inversion. In: A Sequel to the First Six Books of the Elements of Euclid, Containing an Easy Introduction to Modern Geometry with Numerous Examples, 5th ed., Hodges, Figgis & Co., Dublin (1888) 95-112
2. Coolidge, J.L.: Inversion. In: A Treatise on the Geometry of the Circle and Sphere. Chelsea, New York (1971) 21-30
3. Courant, R., Robbins, H.: Geometrical Transformations. Inversion. In: What Is Mathematics?: An Elementary Approach to Ideas and Methods, 2nd ed. Oxford University Press, Oxford (1996) 140-146
4. Coxeter, H.S.M., Greitzer, S.L.: An Introduction to Inversive Geometry. In: Geometry Revisited. Math. Assoc. Amer., Washington (1967) 103-131
5. Coxeter, H.S.M.: Inversion in a Circle and Inversion of Lines and Circles. In: Introduction to Geometry, 2nd ed. John Wiley and Sons, New York (1969) 77-83
6. Darboux, G.: Lecons sur les systemes orthogonaux et les coordonnées curvilignes. Gauthier-Villars, Paris (1910)
7. Dixon, R.: Inverse Points and Mid-Circles. In: Mathographics. Dover, New York (1991) 62-73
8. Durell, C.V.: Inversion. In: Modern Geometry: The Straight Line and Circle. Macmillan, London (1928) 105-120
9. Fukagawa, H., Pedoe, D.: Problems Soluble by Inversion. In: Japanese Temple Geometry Problems. Charles Babbage Research Foundation, Winnipeg, Canada (1989) 17-22, 93-99
10. Gardner, M.: The Sixth Book of Mathematical Games from Scientific American. University of Chicago Press, Chicago, IL (1984)
11. Greenberg, M.: Euclidean and Non-Euclidean Geometries: Development and History. Third Edition. W.H. Freeman and Company, New York (1993)
12. Johnson, R.A.: Modern Geometry: An Elementary Treatise on the Geometry of the Triangle and the Circle. Houghton Mifflin, Boston, MA (1929) 43-57
13. Lachlan, R.: The Theory of Inversion. In: An Elementary Treatise on Modern Pure Geometry. Macmillian, London (1893) 218-236
14. Lockwood, E.H.: Inversion. In: A Book of Curves. Cambridge University Press, Cambridge, England (1967) 176-181
15. Maeder, R.: Programming in Mathematica, Second Edition, Addison-Wesley, Redwood City, CA (1991)
16. Wells, D.: The Penguin Dictionary of Curious and Interesting Geometry. Penguin, London (1991) 119-121
17. Wolfram, S.: The Mathematica Book, Fourth Edition, Wolfram Media, Champaign, IL & Cambridge University Press, Cambridge (1999)

Specific Aspects of Training IT Students for Modeling Pulses in Physics

Adrian Podoleanu[1], Cristian Toma[2], Cristian Morarescu[3], Alexandru Toma[4], and Theodora Toma[5]

[1] University of Kent, Canterburry, UK
[2] Titu Maiorescu University, Department of Applied Informatics, Bucharest, Romania
[3] Politehnica University, Department of Computers, Bucharest, Romania
[4] Marin Preda College, Department of Mathematics and Informatics, Bucharest, Romania
[5] Nicolae Iorga College, Department of Mathematics, Bucharest, Romania

Abstract. This study presents a method for teaching fundamental aspects in mathematics and physics for students in informatics and information technology, so as these concepts to be available for modeling suddenly emerging phenomena at any moment of time during the last years of study. By teaching basic aspects using the language of their specialty and by addressing questions connected to the mathematical or physical model adequate for their projects at unexpected moment of time (using a computer connected at a local network or wireless e-mail connection of mobile phones) the students are able to improve their abilities of qualitative analysis of phenomena. Moreover, their technical judgment patterns can be established in an objective manner. Finally students in Informatics and Information Technology are able not only to understand basic phenomena in physics and electronics, but also to simulate complex phenomena as suddenly emerging pulses in a proper manner, the importance of quick and proper answers being underlined in their mind.

1 Introduction

This study presents specific learning methods established at Kent University, United Kingdom, and adapted for students at Departments of Informatics or Information Technology. During the first and second year of study, they must achieve knowledge in scientific areas as mathematics and physics, so as to understand the basis of microelectronics and electronic devices, but they are willing to pass quickly through these scientific areas, being much more interested in the scientific domain of informatics (considered to be more important for their future job). Conceptual organization of the training domain knowledge based on learning stages is well known [1], but students in Information Technology are selected from the best pupils, and they can't simply pass from one stage to another without a deeper understanding of the use of each stage for their future job. This requires a strong connection between learning stages, which can't

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 556–562, 2005.

be limited just to a few words presented at the beginning of a new matter of study. By the other hand, graduates in Information Technology (IT) must give quick and correct answers to potential customers, and this requires also a specific training method similar to mobile embedded control systems for robots (see [2]), but adapted to human entities.

So as to fulfill these two requirements, a teaching method based on intuitive models and computer simulation of tasks (transmitted to students by mobile phones) has been developed at Titu Maiorescu University, Bucharest. The general aspects of the method were established at Politehnica University, Bucharest, taking into account previous observations upon teaching physics for engineers. Next it has been tested at Marin Preda College and at Nicolae Iorga College, and the final version has been applied to students at Titu Maiorescu University from the first year on study, when some basic matters (mathematics, physics) must be taught.

Compulsory matters of study as mathematics and physics are considered as unpleasant duties. During the communist system, these tasks could be performed by imposing a certain rhythm for study in all technical universities, under the circumstances of compulsory attendance for students to all school activities and of government distribution of graduates at the end of their studies, according to the average value of all their marks. But such a method is not suitable in the new circumstances of former communist countries, when the hiring procedure is based on an interview with questions connected with their specific area of study.

During the first years after the falling of the communist system, Departments of Informatics and Information Technology at Universities were tempted to limit the study of mathematics and physics at just a few basic aspects. These alternative proved to be not a good idea, because students were not able any more to understand the basis of electronic devices used in information technology, with major consequences on using their all capabilities. In order to avoid such difficulties, the new created Faculty of Informatics and Information Technology at University Titu Maiorescu, Bucharest, decided to hire academic staff consisting only of graduates of Departments of Computers, Informatics, Control Engineering or Electronics. May be it looks strange that graduates in computers could teach mathematics, or graduates in control engineering or electronics could teach physics, but we must not forget that in former communist countries all studies at state universities are still free of taxes, and thus many eminent researchers could complete their studies by attending lectures in scientific areas related to their specialization (as Bachelor Students, but more often as Master Students) and obtaining the final diploma. Such professors could not present basic notion in mathematics or physics in the rigorous manner, according to all postulates, but they had an enormous advantage: *all notions were presented using comparisons and correlations with informatics and information technology*, while these was their basic scientific language (see also aspects presented in [3]). Thus students could understand the use of basic notions in mathematics and physics for their future specialty, and all new notions were translated automatically in terms connected with informatics or information technology.

Such a procedure is connected with aspects in psychology, were the term *gestalt* is well-known. It shows that every man possess a certain psychic or intellectual model in his mind, all new concepts transmitted to him having to be attached to this internal model. In first years of university's studies, students do not possess yet this internal intellectual model, but they are trying to build it according to their wish of becoming informaticians or engineers in information technology. If new concepts transmitted at lectures in mathematics or physics do not fit this requirement, sooner or later they will be rejected and any new attempt of recalling them (at lectures in electronics) would need a lot of time (practically the concepts have to be taught again, in a very short time, by a professor who is not prepared for this and who is in a hurry to finish this task so as to be able to presents concepts in his specialty, electronics in our case).

To avoid such problems, the staff of University Titu Maiorescu (a ten years old university, specialized in psychology, sociology and law) decided to use this basic matters of study (which must be presented to the students at the beginning of their academic studies) so as to create the basic structure of their intellectual model, the assimilation of future aspects connected with their future job being much more easier under these circumstances. Moreover, starting from this internal intellectual model, the students can perform the learning way in the opposite direction (as related to time); using their specific concepts (in informatics or information technology), they can easy remember the way used by the professor of mathematics or physics for associating mathematics or physics concepts with them and so they can understand very well phenomena taking place in electronic devices. This represents an important achievement for students or graduates in information technology, because they become able to know all capabilities of electronics devices used, the limits and causes of undesirable phenomena, or even to improve the use of such devices in acquisition, storage or transmission of information. It is also possible for the students to use mathematical models in physics for trying to model or simulate phenomena in psychology or sociology, an important goal for the university staff, which is interested in obtaining software contracts for complex psychological or sociological applications.

2 Intuitive Models in Physics Adapted for Informatics and Information Technology

First step in translating the mathematics and physics language consisted in adding intuitive models to all theoretical principles in physics. For example, according to the romanian programme, studies in physics begin with the principle of least action. This principle is considered as a basic principle in nature, so it is quite easy to add examples from behavior of simple mechanisms and of everyday life, when all individuals tends to a state characterized by least effort. Then potential energy is presented, in analogy with external fields or influences acting upon particles (in physics) and individuals (in everyday life). Further it is quite easy for students to understand the equations of analytical mechanics. After finishing the chapter of mechanics,basic notions from statistical physics are

presented, the temperature being put into correspondence with internal energy of an ensemble of individuals. Thus it is explained the fact that in even in an ensemble seeming to be homogeneous one can find individuals with different levels of energy. Then considerations about entropy (the second principle of thermodynamics) are presented, being shown how external influences can change the degree of self-organization inside a certain ensemble.

After these basic notions from physics were connected with *aspects from information theory (degree of organization, probability)*, basic notions from electrodynamics were presented (the Maxwell laws) in connection with aspects from *information technology (the need of transmission of information at far distances)*. First it is shown the utility of using sources of radiation (charged particles in rest or movement) for generating electromagnetic waves moving with light speed. Then properties of electromagnetic waves are presented, being shown how amplitude and frequency can be established by the source of radiation (where the transmitter acts). Further, wave phenomena as diffraction and interference are put into correspondence with the necessity of changing the property of the signal according to certain aspects from the environment (displacement of bodies, for example), considered as informations having to be transmitted (see also [4]for details about these intuitive models). This way the whole matter of study for the first semester of physics was taught, with the result that the students could analyze phenomena in classical physics (mechanics, statistical physics and optics) from the point of view of an engineer in informatics and information technology.

Next semester aspects from quantum physics and solid state physics were presented, being underlined the strong connection between uncertainty appearing for estimation on frequency using multiplication with harmonic signals and averaging procedures acting upon a certain signal on limited time intervals. This aspects is used for presenting not only Heisenberg principle of uncertainty in physics, but also uncertainty appearing for estimation of wave parameters in signal theory (for heterodyne or homodyne procedures). Thus students can understand the necessity of establishing well-defined decision criteria, an important aspect in the theory of information's transmission. Then main aspects from theory of solid state are presented, with strong emphasis on the fact that the existence of valence band and of conduction band allows a quick transmission of changes in electrons or holes density from one part of a metal or a semiconductor to another; thus an information represented by an electric current can be transmitted through a set of electronic devices for processing it according to a certain algorithm.

3 Applications of Intuitive Models

However, activity in mathematics and physics for students in informatics must not be stopped at the end of the teaching activity. A basic purpose of presenting these matters was to create intuitive models for students, which must be prove to be useful in their work as engineers in informatics and information technology. Thus the next step has been represented by unexpected questions from

mathematics and physics addressed to the students during the teaching hours at technical matters of study (theory of algorithms, electronics and electrical circuits theory), so as the utility of all basic notions from mathematics and physics to be out of the question. Due to the fact that professors teaching technical matters have the same basic studies as professors teaching mathematics and physics -*all being graduates of technical departments of universities* - no problems appeared in establishing the set of questions by the academic staff involved. By this procedure, a main purpose from sociology was reached: *the pass from an imposed behavior to a normal behavior*, the intuitive models from basic areas of study being present in the mind of students all the time.

But the utility of basic notions can't be restricted to a simple understanding of their utility. As engineers, the graduates of technical departments must faced unpredictable problems, so they must be able to apply their knowledge in different circumstances. Their creative capabilities could be tested in an extremely correct manner using the facilities of computers inside a LAN (local area network). The application activities at technical matters consist also in projects. During the communist regime, projects at technical universities consisted in certain themes having to be finished by the end of the school semester, with no connections between the themes of different students. Each phase of the project was checked at two weeks, with about fifteen minutes allowed for each student. Such a procedure is no more suitable nowadays, when the demand of increasing the speed of work at the technical and economical staff inside a company. The solution chosen by Titu Maiorescu University was to adapt capabilities offered by computers inside the local area network (LAN) in order to correct from the very beginning all mistakes made by students in selecting their mathematical or physical model for their project. For a correct evaluation of their capabilities, all students receive by e-mail not only a theme for a project, but also a set of suggested set of equations or physical models, being required to select the adequate set or model in a very short time. After a time interval (limited to a couple of days) each student has to be present in front of the computer again, so as to send a message according to the set or model selected.

It would have been quite easy to impose the presence of the student in front of the computer till he selects the answer, so as to be sure that he wasn't advised by anyone else. Yet this would have been a stress environment for the student, being not recommended when he has to build his first projects. So the professors from mathematics and physics and professors of technical matters prepared together a set of arguments for each possible answer. As soon as the student types the solution selected by him at the personal computer, he receives the set of possible arguments for his answer, having to underline *immediately* the correct arguments for the answer selected. Then the student is evaluated by the correctness of the selected answer and of the arguments supporting his decision, all informations being recoded in an electronic format (the method was tested for the first time at Marin Preda College, and after some improvements of the correspondent software it has been used by Titu Maiorescu University). Thus students learn to choose the adequate mathematical or physical model for their problem; due

to the LAN, major mistakes in building their project are avoided from the very beginning. Further, by working at next phases of the project, students learn to build the whole algorithm or electronic device starting from the initial model. After two semesters with about four projects, students are prepared to build their own mathematical model for any possible application. The achievement of this stage can be checked by suggesting to the best students to select parts of different sets of equations and of physical models and to join them together, so as to build a new model. For checking the correctness of the model, the same procedure (based on a selection of arguments in front of the computer) can be used.

4 Permanent Training of Students

The final step consists in adapting all previous aspects to the real work of an engineer in informatics and information technology. This task is scheduled for the last school semester, so as the knowledge of the future engineer to be available at each moment of time with informations and qualitative analysis. For this purpose, the set of questions connected with the choice of a certain set of mathematical equations or of a certain physical model adequate to the application are not transmitted toward a student situated in front of a personal computer. During a certain time interval established for individual study and for working at school projects, the e-mail connection of a mobile phone is used for sending short and clear questions connected with the advantages or disadvantages of the mathematical or physical model chosen by the student. As related to the previous step, a major difference consists in the fact that questions are more difficult and are send to the student one by one, at unknown time intervals. The answer is required within a few moments (using the same e-mail connection of the mobile phone), while it implies only a qualitative analysis, and it is also stored by the central computer (the server) in an electronic format. This way the students were permanently connected to major contracts of the university, such as the design of acquisition main boards for X-rays cameras (for dental surgery) and the building of the corresponding software. By sending messages presenting the hierarchy of students with good ideas and answers, an atmosphere of competition was created, their working efficiency being much improved. Moreover, their answers were already recorded so as the specialists in human judgment to establish the patterns of technical judgment for each young researcher, which can be send to the company intending to hire the researcher after graduating (aspects similar to those presented in [5]). Moreover, the students in informatics and information technology were also able of a deeper understanding and of computer simulation for complex phenomena in physics and optoelectronic. A major test consisted in modeling suddenly emerging phenomena. These phenomena can't be understood without a basic training in mathematics and physics. Moreover, there are similarities between suddenly emerging phenomena in physics or electronics and the necessity of giving a proper answer in a short time (the answer being also a phenomenon based on an instant qualitative change); thus all psychic and intel

5 Conclusions

This study has presented a method for teaching fundamental aspects in mathematics and physics for students in informatics and information technology, so as these concepts to be available at any moment of time after finishing study at these basic matters and even after graduating. First all basic concepts are presented from the point of view of an engineer in informatics or information technology, so as their utility to be quite clear for the students. Then some questions connected with these basic questions are addressed to the students during the teaching hours at technical matters of study, so as the connection between mathematics, physics and technical matters of study to be presented. Finally some questions related to the choice of a certain mathematical set of equations or of a certain physical model are send to the students using LAN (for students in the last school semester the use of wireless connection to e-mail by mobile phone is added) so as to test the capability of students to select the adequate model for their project. Moreover, taking into account similarities between suddenly emerging phenomena in physics and electronics and the necessity of giving proper answers in a short time), the requirement of modeling suddenly emerging phenomena was used for stimulating all psychic and intellectual abilities of students (the importance of qualitative changes and speed being underlined in their mind). At present time there are attempts to extend the method at training the technical economical personnel at electrical companies.

References

1. Pecheanu, E., Dumitriu, L., Segal, C. : Domain knowledge modeling for inteligent instructional systems.- Lecture Notes in Computer Science **3038** (2004) 497 – 504
2. Srovnal, V., Horak, B., Bernadik, R. : Strategy extractive for mobile embedded control systems. Apply to multi-agent technology.- Lecture Notes in Computer Science **3038** (2004) 631–637
3. Baader, F. : The instance problem and the most specific concept in the description logic w.r.t. terminological cycles with descriptive semantics.- Lecture Notes in Computer Science **2821** (2003) 64–78
4. Toma, C. : The use of cuadri-dimensional interval - the main possibility for improving the Lorentz formulae interpretation, - Proceedings of ECIT-97 Symposium **2** (1997) 202-206
5. Takeda, M., Inenaga, Sh., Bannai, H. : Discovering most classificatory patterns for very expressive pattern classes. - Lecture Notes in Computer Science **2843** (2003) 486–493

Filtering Aspects of Practical Test-Functions and the Ergodic Hypothesis

Flavia Doboga[1], Ghiocel Toma[2], Stefan Pusca[2], Mihaela Ghelmez[2], and Cristian Morarescu[3]

[1] ITT Industries, Washington, U.S.A.
[2] Politehnica University, Department of Physics, Bucharest, Romania
[3] Poliethnica University, Department of Computers, Bucharest, Romania

Abstract. This papers presents properties of dynamical systems able to generate practical test-functions (defined as functions which differ to zero on a certain interval and possess only a finite number of continuous derivatives on the whole real axis) when the free-term of the differential equation (corresponding to the received input signal) is represented by alternating functions. The shape of the output signal (obtained by numerical simulations in Matlab based on Runge-Kutta functions) is analyzed, being shown that for high-frequency inputs an external observer could notice (in certain condition) the generation of two different pulses corresponding to two distinct envelopes. Such as aspect differs to the oscillations of unstable type second order systems studied using difference equations.

1 Introduction

In the ideal mathematical case, suddenly emerging pulses should be simulated using test-functions (functions which differ to zero only on a limited time interval and possessing an infinite number of continuous derivatives on the whole real axis. However, such test functions, similar to the Dirac functions, can't be generated by a differential equation. The existence of such an equation of evolution, beginning to act at an initial moment of time, would imply the necessity for a derivative of certain order to make a jump at this initial moment of time from the zero value to a nonzero value. But this aspect is in contradiction with the property of test-functions to have continuous derivatives of any order on the whole real axis, represented in this case by the time axis. So it results that an ideal test-function can't be generated by a differential equation. For this reason, the analysis must be restricted at practical test-functions [1], defined as functions which differ to zero on a certain interval and possess only a finite number of continuous derivatives on the whole real axis. Mathematical methods based on difference equations are well known [2], but for a higher accuracy of the computer simulation we shall use Runge-Kutta methods in Matlab. The properties of dynamical systems able to generate such practical test-functions will be studied, for the case when the free-term of the differential equation (corresponding to the

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 563–568, 2005.

received input signal) is represented by alternating functions. The shape of the output signal (obtained by numerical simulations in Matlab based on Runge-Kutta functions) will be analyzed, being shown that for high-frequency inputs an external observer could notice (in certain condition) the generation of two different pulses corresponding to two distinct envelopes. Such as aspect differs to the oscillations of unstable type second order systems studied using difference equations [3].

2 Equations Suitable for Generating Symmetrical Pulses

As it is known, a test-function on $[a, b]$ is a function which is nonzero on this interval and which possess an infinite number of continuous derivatives on the whole real axis. For example, the function

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases}$$

is a test-function on $[-1, 1]$. If the graph of the test-function is similar to the rectangular pulse (a unity-pulse), it is considered to be an ideal test-function. An example is the case of the function

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{0.1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases}$$

is close to being an ideal test-function.

Using the expression of $\varphi(\tau)$ and of its derivatives of first and second order, a differential equation which admits as solution the function φ can be obtained. However, a test-function can't be the solution of a differential equation. Such an equation of evolution implies a jump at the initial moment of time for a derivative of certain order, and test-function must possess continuous derivatives of any order on the whole real axis. So it results that a differential equation which admits a test-function φ as solution can generate only a practical test-function f similar to φ, but having a finite number of continuous derivatives on the whole real axis. In order to do this, we must add initial conditions for the function f (generated by the differential equations) and for some of its derivatives $f^{(1)}$, and/or $f^{(2)}$ etc. equal to the values of the test-function φ and of some of its derivatives $\varphi^{(1)}$, and/or $\varphi^{(2)}$ etc. at an initial moment of time t_{in} very close to the beginning of the working interval. This can be written under the form

$$f_{t_{in}} = \varphi_{t_{in}}, \;\; f^{(1)}_{t_{in}} = \varphi^{(1)}_{t_{in}} \text{ and/or } f^{(2)}_{t_{in}} = \varphi^{(2)}_{t_{in}} \text{ etc.} \tag{1}$$

If we want to generate practical test-functions f which are symmetrical as related to the middle of the working interval, we can choose as origin the middle of this interval, and so it results that the function f should be invariant under the transformation

$$\tau \to -\tau$$

Functions invariant under this transformation can be written in the form $f(\tau^2)$, and so the form of a general second order differential equation generating such functions must be

$$a_2\left(\tau^2\right)\frac{d^2f}{d\left(\tau^2\right)^2}+a_1\left(\tau^2\right)\frac{df}{d\tau^2}+a_0\left(\tau^2\right)f=0 \tag{2}$$

However, for studying the filtering properties of practical test-functions we must add a free-term, corresponding to the received signal (the input of the system). Thus, a model for generating a practical test-function using a received signal $u=u(\tau),\ \ \tau\in[-1,1]$, is

$$a_2\left(\tau^2\right)\frac{d^2f}{d\left(\tau^2\right)^2}+a_1\left(\tau^2\right)\frac{df}{d\tau^2}+a_0\left(\tau^2\right)f=u \tag{3}$$

subject to

$$\lim_{\tau\to\pm1}f^k(\tau)=0 \text{ for } k=0,1,\ldots,n. \tag{4}$$

which are the boundary conditions of a practical test-function.

The previous equation is linear as related to the input function u. So we can study independently the output of the system for an input represented by an alternating signal and for an input represented by a constant signal (a step-input). The two outputs signals can be joined together for obtaining the output signal for the case when the input is represented by a mix of a constant and of an alternating function.

3 Filtering Aspects of Practical Test-Functions

When coefficients a_k in (3) are set to

$$a_2=0, a_1=1 \text{ and } a_0=-1, \tag{5}$$

a first order system is obtained under the form

$$\frac{df}{d\left(\tau^2\right)}=f+u \tag{6}$$

which converts to

$$\frac{df}{d\tau}=2\tau f+2\tau u \tag{7}$$

representing a damped first order dynamical system. For the an alternating input $u=\sin 10\tau$, numerical simulations performed using Runge-Kutta functions in Matlab show an attenuation of about $A=3$. In figure 1 is represented the output f of this system for $u=\sin(10\tau)$, and in figure 2 is represented the output f of this system for $u=\cos(10\tau)$. It can be noticed that the mean value of the output oscillations generated in these circumstances is a function of the phase of the input signal. Similar aspects have been noticed for an alternating input

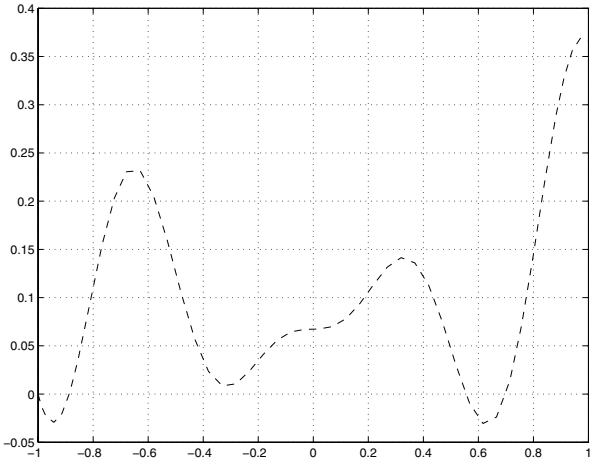


Fig. 1. f versus time for first order damped system, input $u = \sin(10\tau)$

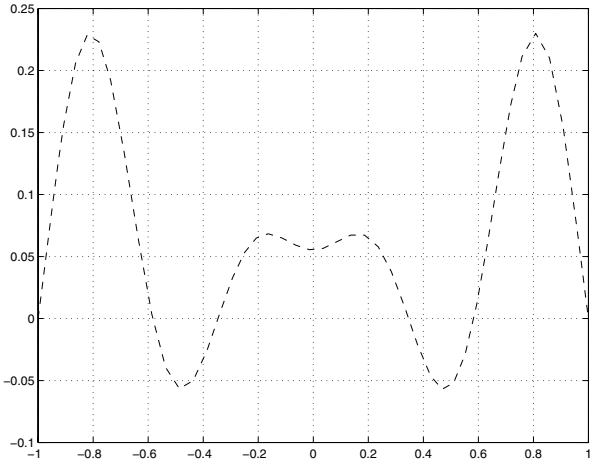


Fig. 2. f versus time for first order damped system, input $u = \cos(10\tau)$

of the form $u = \sin(100\tau)$ the output f of the dynamical system for this case being represented in figure 3, or of the form $u = \cos(100\tau)$ the output f of the dynamical system for this case being represented in figure 4.

When $a_2 = 0, a_1 = 1$ and $a_0 = 0$, another first order model is

$$\frac{df}{d(\tau^2)} = u \tag{8}$$

which converts to

$$\frac{df}{d\tau} = 2\tau u \tag{9}$$

representing an undamped dynamical system. The outputs f of this undamped system are quite similar to the outputs of the previous damped system , for the same inputs u (the differences are less than 15 %).

4 Connection with the Ergodic Hypothesis

Studying graphics presented in figure 3 and figure 4, we can notice the presence of two distinct envelopes. Their shape depends on the phase of the input alternating component. At first sight, an external observer could notice two different pulses generated by the dynamical system (each one corresponding to an envelope). In a more rigorous manner, we can consider that that at a certain moment of time can be detected, with equal probability, one of the two branches of evolution (corresponding to certain intervals around the two envelopes). Thus the mean value of the output f on a small time interval can be considered not as a mean value in time, but also as a mean value for two distinct internal states of the system which *exist together on this time interval*. This is an aspect similar to the ergodic hypothesis used in thermodynamics. By replacing the sinusoidal alternating input u with rectangular alternating functions, the existence of two different branches of evolution would become more obvious (the transition time from one branch to the other would become very short, and so the probability of measuring values different to the two envelopes for the output f .would decrease). This aspect can be put in correspondence with aspects in quantum mechanics, were distinct states can be measured at a certain moment of time, for certain external interactions. The statistical aspects of measuring different values can be studied using the dependence of the envelopes on the phase of the input alternating component.

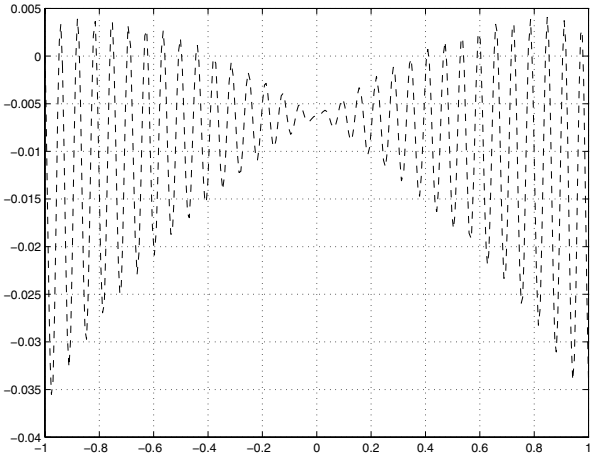


Fig. 3. f versus time for first order damped system, input $u = \sin(100\tau)$

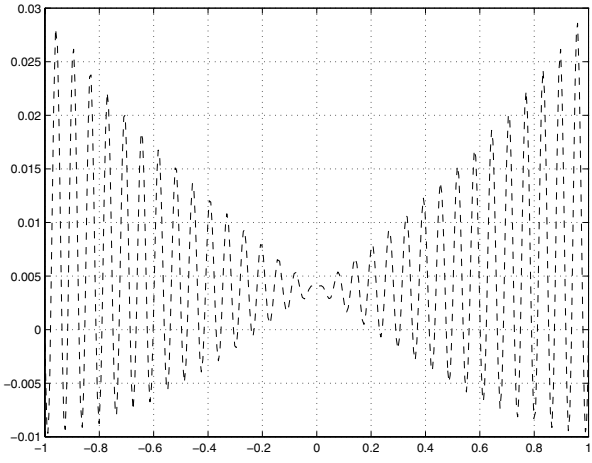


Fig. 4. f versus time for first order damped system, input $u = \cos(100\tau)$

5 Conclusions

This study has presented filtering properties of practical test-functions, the input being represented by alternating sinusoidal functions. The shape of the output signal (obtained by numerical simulations in Matlab based on Runge-Kutta functions) has been analyzed, being shown that for high-frequency inputs an external observer could notice (in certain condition) the generation of two different pulses corresponding to two distinct envelopes. This aspect has been put in correspondence with aspects in quantum mechanics, were distinct states can be measured at a certain moment of time, for certain external interactions. The statistical aspects of measuring different values can be studied using the dependence of the envelopes on the phase of the input alternating component.

References

1. Toma, C. : An extension of the notion of observability at filtering and sampling devices, Proceedings of the International Symposium on Signals, Circuits and Systems Iasi SCS 2001, Romania 233–236
2. Dzurina, J. : Oscillation of second order differential equations with advanced argument, Math. Slovaca, **45** 3 (1995) 263–268
3. Zhang, Zh., Ping, B., Dong, W. : Oscillation of unstable type second order nonlinear difference equation, Korean J. Computer and Appl. Math. **9** 1 (2002) 87–99

Definition of Wave-Corpuscle Interaction Suitable for Simulating Sequences of Physical Pulses

Minas Simeonidis[1], Stefan Pusca[2], Ghiocel Toma[2], Alexandru Toma[3], and Theodora Toma[4]

[1] University of Tuebingen, Germany
[2] Politehnica University, Department of Physics, Bucharest, Romania
[3] Marin Preda College, Department of Mathematics and Informatics, Bucharest, Romania
[4] Nicolae Iorga College, Department of Mathematics, Bucharest, Romania

Abstract. This study presents a logic definition for the interaction between waves and corpuscles suitable for simulating the action of a sequence of electromagnetic waves upon corpuscles. First are defined the classes of measuring methods based on the wave aspect of matter and on the corpuscular aspect of matter, using considerations about a possible memory of previous measurements (operators). A suitable algorithm associated to this formalization is applied on adjoining space-time intervals, so as the space-time validity of certain assertions to be proved. The results are applied for defining the wave-corpuscle interaction in a logic manner.

1 Introduction

As it is known, basic concepts in physics connected with interaction are the wave and corpuscle concepts. In classical physics the corpuscle term describes the existence of certain bodies subjected to external forces or fields, and the wave concept describes the propagation of oscillations and fields. In quantum physics, these terms are closely interconnected, the wave train associated to a certain particle describes the probability of a quantum corpuscle (an electron or a photon) to appear; the results of certain measurements performed upon the quantum particle are described by the proper value of the operators corresponding to the physical quantity to be measured. However, certain intuitive problems connected with measurement procedures on closed-loop trajectories in special relativity and non-commutative properties of operators in quantum physics imply a more rigorous definition of measurement method and of the interaction phenomena, classified from the wave and from the corpuscular aspect of matter, so as to avoid contradiction generated by terminological cycles [1]. This study presents consequences of logic definition for the class of measuring methods based on the wave aspect of matter and for the class of measuring methods based on the corpuscular aspect of matter upon interaction phenomena, using

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 569–575, 2005.

considerations about a possible memory of previous measurements (operators) in case of a sequence of received pulses; it continues in a rigorous manner intuitive aspects presented in [2]. It is shown that measurements methods based on transient phenomena (waves) do not imply a memory of previous actions, while methods based on non-transient phenomena imply the existence of certain quantity which keeps its previous value after the end of a measuring procedure. Then a suitable algorithm associated to this expressive pattern classes (similar to those presented in [3]) is applied on adjoining space-time intervals, so as the space-time validity of certain assertions to be proved and to define in a rigorous manner the interaction between a set of pulses and a corpuscle.

2 Aspects Connected with Measurements on a Set of Pulses Received on Adjoining Space-Time Intervals

As it is known, the special relativity theory considers that the Lorentz formulae describe the transformation of the space-time coordinates corresponding to an event when the inertial reference system is changed. These formulae are considered to be valid at any moment of time after a certain synchronization moment (the zero moment) irrespective to the measuring method used. However, there are some problems connected to the use of mechanical measurements on closed-loop trajectories. For example, let us consider that at the zero moment of time, in a medium with a gravitational field which can be neglected (the use of the galileean form of the tensor g_{ik} being allowed) two observers are beginning a movement from the same point of space, in opposite directions, on circular trajectories having a very great radius of curvature. After a certain time interval, the observers are meeting again in the same point of space. For very great radii of curvature, the movements on very small time intervals can be considered as approximative inertial (as in the case of the transverse Doppler effect, where the time dilation phenomenon was noticed in the earth reference system which is approximative inertial on small time intervals). The Lorentz formulae can be applied on a small time interval $\Delta t(1)$ measured by one of the observers inside his reference system, and it results (using the Lorentz formula for time) that this interval corresponds to a time interval

$$\Delta t'(1) = \frac{\Delta t(1)}{\sqrt{1 - \frac{v(1)^2}{c^2}}} \tag{1}$$

in the reference system S_2 of the other observer, which moves with speed $v(1)$ as related to the reference system S_1 on this time interval. So the time dilation phenomenon appears. If each observer considers the end of this time interval ($\Delta t(1)$ or $\Delta t'(1)$) as a new zero moment (using a resynchronization procedure), the end of the second time interval $\Delta t(2)$ (with the new zero moment considered as origin) will correspond to a time moment

$$\Delta t'(2) = \frac{\Delta t(1)}{\sqrt{1 - \frac{v(2)^2}{c^2}}} \tag{2}$$

measured in the other reference system S_2 which moves with speed $v(2)$ as related to system S_1 on the time interval $\Delta t'(2)$ (with the new zero moment considered as origin). As related to the first zero moment (when the circular movement has started) the end of the second time interval appears at the time moment

$$t_2 = \Delta t(1) + \Delta t(2) \tag{3}$$

for the observers situated in reference system S_1, and at the time moment

$$t'(2) = \Delta t'(1) + \Delta t'(2) = \frac{\Delta t(1)}{\sqrt{1 - \frac{v(1)^2}{c^2}}} + \frac{\Delta t(2)}{\sqrt{1 - \frac{v(2)^2}{c^2}}} \tag{4}$$

for the other observer.

Due to the fact that

$$\Delta t'(1) > \Delta t(1) \tag{5}$$

and

$$\Delta t'(2) > \Delta t(2) \tag{6}$$

it results that

$$t'(2) = \Delta t'(1) + \Delta t'(2) > \Delta t(1) + \Delta t(2) = t(2) \tag{7}$$

and thus a global time dilation for the time interval $\Delta t(1) + \Delta t(2)$ appears. The procedure can continue, by considering the end of each time interval

$$\Delta t(1) + \Delta t(2) + \ldots + \Delta t(i)$$

as a new zero moment, and so it results that on all the circular movement period, a time moment

$$t(k) = \sum_{i=0}^{k} \Delta t(i) \tag{8}$$

(measured by the observer in reference system S_1) corresponds to a time moment

$$t'(k) = \sum_{i=0}^{k} \Delta t'(i) = \sum_{i=0}^{k} \frac{\Delta t(i)}{\sqrt{(1 - \frac{v_i^2}{c^2}}} \tag{9}$$

(measured by the observer situated in reference system S_2), which implies

$$t'(k) > t(k) \tag{10}$$

By joining together all these time intervals $\Delta t(i)$ we obtain the period of the whole circular movement T. While the end of this movement is represented by the end of the time interval $\Delta t(N)$ in the reference system S_1, it results that T can be written under the form

$$T = t(N) = \sum_{i=0}^{N} \Delta t(i) \tag{11}$$

(considered in the reference system S_1), and it results also that this time moment (the end of the circular movement) corresponds to a time moment

$$T' = t'(N) = \sum_{i=0}^{N} \Delta t'(i) \tag{12}$$

measured in the reference system $S_{@}$. While

$$\Delta t'(i) = \frac{\Delta t(i)}{\sqrt{1 - \frac{v(i)^2}{c^2}}} > \Delta t(i) \tag{13}$$

it results

$$T' > T \tag{14}$$

If the time is measured using the age of two twin children, it results that the twin in reference system S_2 is older than the other in reference system S_1, (having a less mechanical resistance of bones) and it can be destroyed by it after both observers stop their circular movements. However, the same analysis can be made starting from another set of small time intervals $\Delta_n t'(i)$ considered in the reference system S_2 which corresponds to a new set of time intervals $\Delta_n t(i)$ considered in the reference system S_2 (established using the same Lorentz relation) and finally it would result that the period of the circular movement T' measured in system S_2 corresponds to a period T greater than T' considered in reference system S_1. If the time is measured using the age of two twin children, it results that the twin in reference system S_1 is older than the other in reference system S_2, (having a less mechanical resistance of bones) and it can be destroyed by it after both observers stop their circular movements. But this result is in logic contradiction with the previous conclusion, because a man can not destroy and in the same time be destroyed by another man.

As a first attempt of solving this contradiction, one can suppose that Lorentz formulae are valid only for electromagnetic phenomena (as in the case of the transversal Doppler effect) and not in case of mechanical phenomena. But such a classification is not a rigorous classification, being not suitable for formal logic. In next section we will present a more rigorous classification of phenomena used in space-time measurements, which can be used for *gedanken* experiments using artificial intelligence based on formal logic.

3 A Rigorous Definition of Wave and Corpuscle Concepts and of Wave-Corpuscle Interaction

The logical contradiction presented in previous section appeared due to the fact that an element with internal memory has been used. The indication of this element has'not been affected by the resynchronization procedure. In modern physics such an element with internal memory is connected with the corpuscular aspect of matter, with a body. On the contrary, a measuring procedure

based on an electromagnetic or optic wave-train is a transient phenomenon. The synchronization of clocks is possible only after the wave-train arrives at the observer. Excepting a short time interval after the reception the received wave-train doesn't exist inside the observer's medium, so there isn't any space area where a physical quantity which characterizes the wave to cumulate. That's the reason why a correct solution of the twins paradox must be based not on the association of electromagnetic (or optic) phenomena with the Lorentz formulae, but on the association of the Lorentz formulae with wave phenomena describing the propagation of a wave inside the observers reference systems. The wave class is more general than the class of electromagnetic and optic waves (we can mention the wave associated with particles in quantum mechanics). Besides, in the most general case, the interaction between two reference systems appears under the form of a field, not under the form of a material body. Moreover, this aspect implies an intuitive interpretation for the dependence of the mass of a body inside a reference system.

Using the formal logic, all we have shown can be presented in a rigorous manner.

A) We define the notion of "propagation"phenomenon in two inertial reference systems (the system where the event takes place and the system where a signal generated by the event is noticed)

Definition 1. *It exists a set of adjoining space intervals* $\{S_0, S_1, \ldots, S_n\}$*, a set of adjoining time intervals* $\{T_0, T_1, \ldots, T_n\}$ *in a certain reference system; it exists a set of physical quantities* $F_u = \{F_{u1}, F_{u2}, \ldots, F_{um}\}$ *and a set of relations* $R_{10}, R_{21}, \ldots,$ *so as*

$$F_u(S_1, T_1) = R_{10}F_u(S_0, T_0),\ F_u(S_2, T_2) = R_{21}F_u(S_1, T_1), \ldots$$

and

$$\{F_u(S_0, T_0) \neq 0,\ F_u(S_0, t) = 0 \text{ for } t \notin T_0\} \Longrightarrow$$

$$\{F_u(S_1, T_1) \neq 0,\ F_u(S_1, t) = 0 \text{ for } t \notin T_1\} \ldots \Longrightarrow$$
$$\{F_u(S_n, T_n) \neq 0,\ F_u(S_n, t) = 0 \text{ for } t \notin T_n\}$$

It can be noticed that we described a propagation phenomenon having a finite existence inside the reference system, the number of intervals being finite.

B) We define the notion of corpuscle inside a certain reference system

Definition 2. *It exists a set of adjoining space intervals* $\{S_0, S_1, \ldots, S_n, \ldots\}$*, and a set of adjoining time intervals* $\{T_0, T_1, \ldots, T_n, \ldots\}$ *in a certain reference system; it exists a set of physical quantities* $F_c = \{F_{c1}, F_{c2}, \ldots, F_{cm}\}$ *and a set of relations* $R_{10}, R_{21}, \ldots,$ *so as*

$$F_c(S_1, T_1) = R_{10}F_c(S_0, T_0),\ F_c(S_2, T_2) = R_{21}F_c(S_1, T_1), \ldots$$

and

$$\{F_c(S_0, T_0) \neq 0,\ F_c(S_0, t) = 0 \text{ for } t \notin T_0\} \Longrightarrow$$

$$\{F_c(S_1, T_1) \neq 0,\ F_c(S_1, t) = 0\ for\ t \notin T_1\} \ldots \Longrightarrow$$

$$\{F_c(S_n, T_n) \neq 0,\ F_c(S_n, t) = 0\ for\ t \notin T_n\} \Longrightarrow \ldots$$

It can be noticed that these relations are describing a phenomenon which can possess an unlimited evolution in time and space inside the reference system; it can be also said that the phenomenon has its own existence, it exists by itself.

C) We define the emission of a wave-train U_e in a reference system and its transformation in another train when it interacts with the observers's medium

Definition 3. *It exists an area S_{0e} and a time interval T_{0e} in the reference system where the emission takes place so that*

$$F_{ue}(S_{0e}, T_{0e}) \neq 0,\ F_{ue}(S_{0e}, t) = 0\ for\ t \notin T_{0e}$$

It exists a space area S_{0r} and a time interval T_{0r} in the observer's reference system, and a relation Tr so that

$$F_{ur}(S_{0r}, T_{0r}) = Tr\left[F_{ue}(S_{0e}, T_{0e})\right],$$

$$F_{ur}(S_{0r}, T_{0r}) \neq 0,\ F_{ur}(S_{0r}, t) = 0\ for\ t \notin T_{0r}$$

So it exists a certain physical quantity characterizing the body which is influenced by the received wave train even after this wave train has disappeared (it exists a memory of the previous measurements).

D) We define the transformation of a sequence of received pulses $\Sigma_k Ue_k$ in a sequence $\Sigma_k Ur_k$, k = 1...n after interaction with the observers'reference system, by considering that each pulse (wave-train) is transformed in an independent manner by the material medium of the observer's reference system, according to its specific Lorentz transformation

Definition 4.

$$Ur_k = L_k\left[Ue\right]_k$$

$$\Sigma_k Ue_k = \Sigma_k Ur_k$$

where L_k represents the Lorentz transformation performed upon the Ue_k wave by the system, with the interaction moment of this wave with the material medium of the observer considered as zero moment of time (synchronization moment) for the Lorentz transformation L_k.

E) We define the interaction between a sequence of pulses and the material body of the observer's reference system (a corpuscle) as an interaction function Int between the material medium and each transformed pulse Ur_k corresponding to a received pulse Ue_k, the mass m of the body measuring the influence of the received wave-train Ue_k upon the body.

Definition 5.

$$\frac{1}{m} = Int\left[Ur_k\right] = Int\left[L_k\left(Ue\right)_k\right]$$

When Lorentz transformation L_k doesn't generate a pulse Ur_k (for example when the relative speed between the material body and the wave is equal to c, the speed of light in vacuum), the mass m is equal to ∞ , which means that no interaction due to the received pulse Ue_k exists (an idea appeared at Marin Preda College, which connects the notion on infinite mass with the absence of interaction). So $m = \infty$ for a body inside a reference system S shows that we can't act upon the material body using wave pulses emitted in system S; however, changes in the movement of the body (considered in system S) due to other external forces seem to be allowed.

All previous definitions implies the necessity of using distinct memory areas for each pulse, if we intend to simulate sequences of optical pulses. By interaction with a certain material medium, each pulse is transformed according to Lorentz formulae, and the modified parameters of each pulse must replace the previous informations in the memory cells. For wave trains considered inside the material medium, a method to simplify the use of the memory cells (appeared at Nicolae Iorga College) would consist in considering the wave as a mixture of two certain states (similar to a rectangular wave), each state corresponding to a certain set of parameters stored in a memory cell; thus a small number of coefficients (for multiplying each state before adding them) would be able to describe the wave evolution with a very good approximation. Such an aspect is similar to Heisenberg representation in quantum theory (where state of a particle is always the same and the operators change in time) and it will be studied in the future.

4 Conclusions

This study has presented a logic definition for the class of measuring methods based on the wave aspect of matter and for the class of measuring methods based on the corpuscular aspect of matter, using considerations about a possible memory of previous measurements (operators). It has been shown that measurements methods based on transient phenomena (waves) do not imply a memory of previous actions, while methods based on non-transient phenomena imply the existence of certain quantity which keeps its previous value after the end of a measuring procedure.

References

1. Baader, F. : The instance problem and the most specific concept in the description logic w.r.t. terminological cycles with descriptive semantics, Lecture Notes in Computer Science **2821** (2003) 64–78
2. Toma, C. : A connection between special relativity and quantum theory based on non-commutative properties, Proceedings of ECIT-97 Symposium **1** (1997), 138–142
3. Takeda, M., Inenaga, Sh., Bannai, H.: Discovering most classificatory patterns for very expressive pattern classes, Lecture Notes in Computer Science **2843** (2003) 486–493

Practical Test-Functions Generated by Computer Algorithms

Ghiocel Toma

Politehnica University, Department of Physics, Bucharest, Romania

Abstract. As it is known, Runge-Kutta methods are widely used for numerical simulations [1]. This paper presents an application of such methods (performed using MATLAB procedures) for generating practical test functions. First it is shown that differential equations can generate only functions similar to test functions (defined as practical test functions); then invariance properties of these practical test functions are used for obtaining a standard form for a differential equation able to generate such a function. Further this standard form is used for computer aided generation of practical test-functions; a heuristic algorithm (based on MATLAB simulations) is used so as to establish the most simple and robust expression for the differential equation. Finally it is shown that we obtain an oscillating system (a system working at the limit of stability, from initial null conditions, on limited time intervals) which can be built as an analog circuit using standard electrical components and amplifiers, in an easy manner.

1 Introduction

Many times the analysis of signals requires an integration on a limited time interval, which can't be performed in a robust manner (with sampling procedures) without using functions similar to test-functions (for multiplying the received signal before the integration, so as the result of the integration to be practically constant at the end of the integration period, at the sampling moment of time). Usually this operation is performed by an integration of the signal on this time interval, using an electric current charging a capacitor - the result of the integration being proportional to the mean value of the signal. However, such structures are very sensitive at random variations of the integration period. Even when devices with higher accuracy are used for establishing this time interval some random variations will appear due to the stochastic switching phenomena - when the electric current charging the capacitor is interrupted. For this reason, a multiplication of the received signal with a test-function - a function which differs to zero only on this time interval and with continuous derivatives of any order on the whole real axis - is recommended. In the ideal case, such a test-function should have a form similar to a rectangular pulse - a unity pulse - considered on this time interval. However, such test functions, similar to the Dirac functions, can't be generated by a differential equation. The existence of such an equation

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 576–584, 2005.

of evolution, beginning to act at an initial moment of time, would imply the necessity for a derivative of certain order to make a jump at this initial moment of time from the zero value to a nonzero value. But this aspect is in contradiction with the property of test-functions to have continuous derivatives of any order on the whole real axis, represented in this case by the time axis. So it results that an ideal test-function can't be generated by a differential equation. For this reason, we must restrict our analysis at the possibilities of generating practical test-functions. This practical or truncated test-functions differ to zero only on a certain interval and possess only a finite number of continuous derivatives on the whole real axis. We must find out what properties should be satisfied by a differential equation of evolution, so as starting from certain initial conditions such a practical test-function to be generated.

2 Preliminaries

In previous section has been shown that an ideal test function can't be generated by an equation of evolution (see also [2]). Besides, the problem of generating truncated test functions can't be solved by studying aspects connected with solitary waves [3] or by studying period doubling and chaos generated by thermal instability [4], because we must restrict restrictt restrict our analysis at a certain time interval and we must study only differential equations. So we must study equations of evolution able to generate pulses available for our task - the multiplication with the received signal - so as the average procedure to be insensitive at random variations of the integration period. The function which is integrated must be as possible zero at the end of the integration period; this result can be obtained only when the function which multiplies the received signal is a practical test function, how it has been shown. Finally the advantage of using such practical test function for wavelets processing is presented.

3 Differential Equations Able to Generate Practical Test-Functions

As it is known, a *test-function* on [a, b] is a C^{∞} function on $\mathbf{R}$ which is nonzero on (a, b) and zero elsewhere. For example, the bump-like function

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

is a test-function on $[-1, 1]$.

Definition 1. *An ideal test-function is a test-function that has a graph similar to a rectangular pulse (a unity-pulse) which is 1 on (a, b) and 0 elsewhere.*

For example, the bump-like function

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{0.1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

is close to being an ideal test-function. Such ideal test functions are recommended for multiplying the received signal.

Definition 2. *A practical test-function on [a, b] is a C^n function f on R (for a finite n) such that*
a) f is nonzero on (a, b)
b) f satisfies the boundary conditions $f^{(k)}(a) = f^{(k)}(b) = 0$ for $k = 0, 1, ..., n$ and
c) f restricted to (a, b) is the solution of an initial value problem (i.e. an ordinary differential equation on (a, b) with initial conditions given at some point in this interval).

The primary task of artificial intelligence consists in generating practical test-functions by numerical integration, using the expressions of a certain test function and of some of its derivatives. An initial value problem will be established, and the solution will be find using a Runge-Kutta method of order 4 or 5 in MATLAB.

Subroutine 1. Identifying an initial value problem to generate practical test-functions on $[-1, 1]$ begins with considering differential equations satisfied by the bump function φ; the first and second derivatives of φ are obtained (using standard program for derivatives) under the form
Step A

$$\varphi^{(1)}(\tau) = \frac{-2\tau}{\left(\tau^2 - 1\right)^2} \exp\left(\frac{1}{\tau^2 - 1}\right) \tag{3}$$

$$\varphi^{(2)}(\tau) = \frac{6\tau^4 - 2}{\left(\tau^2 - 1\right)^4} \exp\left(\frac{1}{\tau^2 - 1}\right) \tag{4}$$

Step B
A special algorithm tries to obtain a correspondence between the expressions of test function and of its derivatives, by replacing the exponential function. By simply dividing the function $\varphi(\tau)$ at $\varphi^{(1)}(\tau)$ we obtain the correspondence between $\varphi(\tau)$ and $\varphi^{(1)}(\tau)$ under the form

$$\varphi^{(1)} = \frac{-2\tau}{\left(\tau^2 - 1\right)^2}\varphi \tag{5}$$

Then the special algorithm replaces the functions $\varphi(\tau)$ and $\varphi^{(1)}(\tau)$ with functions $f(\tau)$ and $f^{(1)}(\tau)$; as initial conditions, it considers the values of $\varphi(\tau)$ at a moment of time $\tau = -0.99$ (close to the moment of time $\tau = -1$). Thus it results for generating a practical test function f the first order initial value problem

$$f^{(1)} = \frac{-2\tau}{\left(\tau^2 - 1\right)^2} f, \quad f(-0.99) = \varphi(-0.99) \tag{6}$$

The initial condition is roughly 1.5×10^{-22}, which means approximately zero. Numerical integration gives a solution having the form of φ but with a very

small amplitude of 10^{-12}. The same way this special algorithm obtains a correspondence between the expressions of $\varphi, \varphi^{(2)}$; this results under the form

$$\varphi^{(2)} = \frac{6\tau^4 - 2}{(\tau^2 - 1)^4}\varphi \tag{7}$$

The replacement of functions $\varphi, \varphi^{(2)}$ with functions $f, f^{(2)}$ and the initial condition under the form $f(-0.99) = \varphi(-0.99)$, $f^{(1)}(-0.99) = \varphi^{(1)}(-0.99)$ leads to the second order initial value problem for generating a practical test function f

$$f^{(2)} = \frac{6\tau^4 - 2}{(\tau^2 - 1)^4} f, \quad f(-0.99) = \varphi(-0.99),\ f^{(1)}(-0.99) = \varphi^{(1)}(-0.99) \tag{8}$$

Numerically integration gives a solution similar to φ, but with an amplitude that is only four times greater than that obtained from the first order initial value problem.

Step C

The algorithm analyzes the possibilities of generating a practical test function similar to an ideal unitary pulse. For this purpose, it replaces the bump-like function $\varphi(\tau)$ with the almost ideal test function

$$\varphi_a(\tau) = \begin{cases} \exp\left(\frac{0.1}{\tau^2 - 1}\right) & \text{if} \quad \tau \in (-1, 1) \\ 0 \text{ otherwise} \end{cases} \tag{9}$$

In the same way used for studying possibilities of generating practical test-functions similar to test-function $\varphi(\tau)$, the algorithm analyzes possibilities of generating practical test-functions similar to test-function $\varphi_a(\tau)$. Taking into account the expressions of $\varphi_a, \varphi_a^{(2)}$ (obtained using standard algorithms for derivatives), the correspondence between φ_a and $\varphi_a^{(2)}$ results now under the form

$$\varphi_a^{(2)} = \frac{0.6\tau^4 - 0.36\tau^2 - 0.2}{(\tau^2 - 1)^4}\varphi_a \tag{10}$$

By replacing the functions φ_a and $\varphi_a^{(2)}$ with f and $f^{(2)}$ and considering similar initial conditions $f(-0.99) = \varphi_a(-0.99), f^{(1)}(-0.99) = \varphi_a^{(1)}(-0.99)$ it results the second order initial value problem under the form

$$f^{(2)} = \frac{0.6\tau^4 - 0.36\tau^2 - 0.2}{(\tau^2 - 1)^4} f, f(-0.99) = \varphi_a(-0.99), f^{(1)}(-0.99) = \varphi_a^{(1)}(-0.99) \tag{11}$$

Numerical integration (performed using MATLAB functions) gives a solution that is nearly ideal; its amplitude is close to 1 for more than 2/3 of the interval $[-1, 1]$.

The whole heuristic program will continue by trying to design the most simple differential equation able to generate a practical test function. The algorithm is based on the fact that all practical test-functions numerically generated by the

initial value problems considered so far are symmetric about $\tau = 0$, which means that they are invariant under the transformation

$$\tau \to -\tau.$$

It results

Subroutine 2

Step A: Functions invariant under this transformation can be written in the form $f(\tau^2)$ and the second order differential equations generating such functions must necessarily have the form

$$a_2\left(\tau^2\right)\frac{d^2f}{d\left(\tau^2\right)^2} + a_1\left(\tau^2\right)\frac{df}{d\tau^2} + a_0\left(\tau^2\right)f = 0 \tag{12}$$

Step B: Using standard algorithms, all derivatives presented in the previous relation are replaced by derivatives having the form

$$b_1(\tau)\frac{df}{d\tau}, \quad b_2(\tau)\frac{d^2f}{d\tau^2}$$

Because

$$\frac{df}{d\tau} = 2\tau\frac{df}{d\left(\tau^2\right)} \quad \text{and} \quad \frac{d^2f}{d\tau^2} = 4\tau^2\frac{d^2f}{d\left(\tau^2\right)^2} + 2\frac{df}{d\left(\tau^2\right)} \tag{13}$$

the previous differential equation results now under the form

$$\frac{a_2\left(\tau^2\right)}{4\tau^2}\frac{d^2f}{d\tau^2} + \left(\frac{a_1\left(\tau^2\right)}{2\tau} - \frac{a_2\left(\tau^2\right)}{4\tau^3}\right)\frac{df}{d\tau} + a_0\left(\tau^2\right)f = 0 \tag{14}$$

Step C: Adding a possible free term in previous differential equation, it results a model for generating a practical test-function using a received signal $u = u(\tau), \tau \in [-1, 1]$, under the form

$$\frac{a_2\left(\tau^2\right)}{4\tau^2}\frac{d^2f}{d\tau^2} + \left(\frac{a_1\left(\tau^2\right)}{2\tau} - \frac{a_2\left(\tau^2\right)}{4\tau^3}\right)\frac{df}{d\tau} + a_0\left(\tau^2\right)f = u \tag{15}$$

subject to

$$\lim_{\tau\to\pm1} f^k(\tau) = 0 \text{ for } k = 0, 1, \ldots, n,$$

which are the boundary conditions of a practical test-function. While we are looking at the most simple solutions, the free term u is set by the algorithm to a constant value.

Step D: The coefficient

$$\frac{a_2\left(\tau^2\right)}{4\tau^2}$$

which multiplies

$$\frac{d^2f}{d\tau^2}$$

is analyzed so as to result a constant expression for it. While the denominator is $4\tau^2$, for the constant set to unity results $a_2 = 4\tau^2$.

Step E: The coefficient

$$\left(\frac{a_1\left(\tau^2\right)}{2\tau} - \frac{a_2\left(\tau^2\right)}{4\tau^3} \right)$$

which multiplies

$$\frac{df}{d\tau}$$

is then analyzed so as to result a constant expression for it. While $a_2 = 4\tau^2$ (from step D), starting an algorithm for polynomial expressions so as to obtain a null coefficient (because we try to obtain the most simple differential equation), it results that $a_1 = 2$. Under these circumstances, the second term in the differential equation vanishes.

Step F: For obtaining the coefficient $a_0\left(\tau^2\right)$ which multiplies the term f in the differential equation the polynomial algorithm tries first to set a_0 to zero. But numerical simulation shows that the response does not satisfy the boundary conditions. So the polynomial algorithm will set the coefficient a_0 to unity, and the differential equations results under the form

$$4\tau^2 \frac{d^2 f}{d\left(\tau^2\right)^2} + 2\frac{df}{d\left(\tau^2\right)} + f = u \tag{16}$$

which converts to

$$\frac{d^2 f}{d\tau^2} + f = u \tag{17}$$

This is an autonomous differential equation; the form being invariant at time translation, so the point $\tau = -1$ is translated to $\tau = 0$. Thus we obtain the differential equation of an oscillating second order system, described by the transfer function

$$H(s) = \frac{1}{\left(T_0\right)^2 s^2 + 1} \tag{18}$$

where $T_0 = 1$ in our case.

Subroutine 3. We continue our study by analyzing the behavior of oscillating systems in the most general case (when $T_0 \neq 1$). This implies a differential equation under the form

$$\left(T_0\right)^2 \frac{d^2 f}{d\tau^2} + f = u \tag{19}$$

Step A. The behavior of this oscillating system can be obtained using standard algorithms (being a linear second order system). When u is represented by a constant, the output of the system consists of oscillations around a constant value. Analyzing the coefficients of the linear differential equation obtained, the existence of the oscillating linear system can be easy noticed, and the working

interval can be set to the time interval corresponding to an oscillation, this means the time interval $[0, 2\pi T_0]$. When $u = 1$, the model generates the practical test-function

$$f(\tau) = 1 - \cos(\tau/T_0) \tag{20}$$

Step B. The algorithm generates the output of the oscillating system under the influence of a continuous useful signal u_1 (supposed to be constant) with an alternating noise u_2 of angular frequency ω added. Using the property of linearity, the equations

$$a_2\left(\tau^2\right)\frac{d^2 f_1}{d\left(\tau^2\right)^2} + a_1\left(\tau^2\right)\frac{df_1}{d\tau^2} + a_0\left(\tau^2\right) f_1 = u_1 \tag{21}$$

and

$$a_2\left(\tau^2\right)\frac{d^2 f_2}{d\left(\tau^2\right)^2} + a_1\left(\tau^2\right)\frac{df_2}{d\tau^2} + a_0\left(\tau^2\right) f = u_2 \tag{22}$$

imply that $f = f_1 + f_2$ is a solution of

$$a_2\left(\tau^2\right)\frac{d^2 f}{d\left(\tau^2\right)^2} + a_1\left(\tau^2\right)\frac{df}{d\tau^2} + a_0\left(\tau^2\right) f = u_1 + u_2 \tag{23}$$

(this aspect can be noticed by an algorithm by identifying constant values for the coefficients of the differential equation used). This reduces the study of the model when the input u is a mix of continuous useful signal and noise (an alternating input for example) to two cases: that of a continuous useful signal, and that of the noise. Then we can add the results to obtain the output when the noise overlaps the useful signal.

Step C. By checking the performances of this system under the influence of an external constant input $u = 1$, an averaging procedure on the working time interval $[0, 2\pi T_0]$ shows that it recovers the mean value of the useful signal $u = 1$ over this interval:

$$\frac{1}{2\pi T_0}\int_0^{2\pi T_0}\left(1 - \cos(\tau/T_0)\right)d\tau = 1 \tag{24}$$

The human user can also notice that the integration of this practical test-function on $[0, 2\pi T_0]$ is practically insensitive to the switching phenomena appearing at the sampling moment of time $2\pi T_0$ because

$$f(2\pi T_0) = 0 \text{ and } f^{(1)}(2\pi T_0) = 0 \tag{25}$$

Step D. By the other hand, an analysis of the oscillating system for an alternating input of $u = \sin\omega\tau$ with frequency $\omega > \sqrt{2}/T_0$ shows that the system attenuates this input:

$$\left|\frac{\text{input}}{\text{output}}\right| = (T_0\omega)^2 - 1 > 1 \tag{26}$$

The human observer can notice that oscillations of the form

$$a \sin(t/T_0) + b \cos(t/T_0) \tag{27}$$

generated by the alternating component of the input account for all the solutions of the associated homogeneous system

$$(T_0)^2 \frac{d^2 f}{d\tau^2} + f = 0 \tag{28}$$

These oscillations give null result by an integration over the working interval $[0, 2\pi T_0]$.

Final Conclusion. By comparing the fact that the mean value of the received useful signal (supposed to be constant) can be obtained using a linear differential equation where the external signal appears as a free term (this being a task performed by artificial intelligence), the human user can replace the use of a practical test function in a multiplying procedure (when it multiplies the received signal) with an use of a practical test function represented by a linear oscillating system in a generating procedure (the external signal representing the free term in the differential equation corresponding to the oscillating system).

As a conclusion, the simplest model for generating practical test-functions on $[0, 2\pi T_0]$ when the continuous signal is $u = 1$ designed using computer generated practical test functions based on MATLAB procedures and standard algorithms and verified using the same MATLAB procedures consists of the second order oscillating system

$$(T_0)^2 \frac{d^2 f}{d\tau^2} + f = u \tag{29}$$

over the interval $[0, 2\pi T_0]$, in which test-functions are subject to the boundary conditions

$$f(0) = 0, \quad f(2\pi T_0) = 0 \tag{30}$$

these implying also

$$f^{(1)}(2\pi T_0) = 0, \quad f^{(1)}(2\pi T_0) = 0 \tag{31}$$

4 Conclusions

This paper has presented a heuristic algorithm for generating practical test functions using MATLAB procedures. First it has been shown that ideal test functions can't be generated by differential equations, being underlined the fact that differential equations can generate only functions similar to test functions (defined as practical test functions). Then a step by step algorithm for designing the most simple differential equation able to generate a practical test function is presented, base on the invariance properties of the differential equation and on standard MATLAB procedures. The result of this algorithm is represented by a system working at the stability limit from initial null conditions, on limited time intervals, the external signal representing the free term in the differential equation corresponding to the oscillating system. Such a system can be built using standard components and operational amplifiers, in an easy manner.

References

1. Kulikov, G. : An advanced version of the local-global step-size control for Runge-Kutta methods applied to index 1 differential algebraic systems, Lecture Notes in Computer Science **3037** (2004) 565–569
2. Toma, C. : Acausal pulses in physics-numerical simulations, Bulgarian Journal of Physics (to appear)
3. D'Avenia, P., Fortunato, D., Pisani, L. : Topological solitary waves with arbitrary charge and the electromagnetic field, Differential Integral Equations **16** (2003) 587–604
4. Frankel, M., Roytburd, V. : Finite-dimensional model of thermal instability, Appl. Math. Lett. **8** (1995) 39–44

Possibilities for Obtaining the Derivative of a Received Signal Using Computer-Driven Second Order Oscillators

Andreea Sterian and Ghiocel Toma*

Politehnica University, Department of Physics, Bucharest, Romania

Abstract. As it is known, a first step in modeling dynamic phenomena consists in measuring with higher accuracy some physical quantities corresponding to the dynamic system. However, for suddenly-emerging phenomena ,the data acquisition can't be restricted at sampling procedures for a received signal (corresponding to a certain physical quantity). A significant quantity is represented by the derivative (the slope) of the received signals, because all dynamical models must take it into consideration. Usually the derivative of a received signal is obtained by filtering the received signal and by dividing the difference between the filtered values of the signal at two different moments of time at the time difference between these time moments. Many times these filtering and sampling devices consists of low-pass filters represented by asymptotically stable systems, sometimes an integration of the filter output over a certain time interval being added. However, such a structure is very sensitive at random variations of the integration period, and so it is recommended the signal which is integrated to be approximately equal to zero at the end of the integration period. It will be shown that the simplest structure with such properties is represented by an oscillating second order computer-driven system working on a time period.

1 Introduction

It is known that the derivative of a received signal is usually obtained by filtering the received signal (using low-pass filters) and by dividing the difference between the filtered values of the signal at two different moments of time at the time difference between these time moments. The time difference Δt is very small and it is usually set by oscillators having a higher accuracy, and so it can be considered as constant.

Usually the filtering device consists of low-pass filters represented by asymptotically stable systems, sometimes an integration of the filter output over a certain time interval being added. However, such a structure is very sensitive at random variations of the integration period, and so it is recommended the signal

* This work is partly supported by the European Commision under contract EVG1-CT-2002-0062 (OPTSDET).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 585–591, 2005.

which is integrated to be approximately equal to zero at the end of the integration period. So we must try to use oscillating systems for filtering the received signal (just in this case the filtered signal and its slope are approximately zero at the end of a certain time interval). However, for avoiding instability of such oscillating systems we must add certain electronic devices (gates) controlled by computer commands, so as to restore initial null conditions for the oscillating system. Before designing such a structure, we must notice that filtering and sampling devices consisting of low-pass filters of first or second order have the transfer function

$$H(s) = \frac{1}{T_0 s + 1} \tag{1}$$

(for a first order system) and

$$H(s) = \frac{1}{T_0^2 s^2 + 2bT_0 s + 1} \tag{2}$$

(for a second order system). They attenuate an alternating signal of angular frequency $\omega >> \omega_0 = 1/T_0$ about ω/ω_0 times (for a first order system) or about $(\omega/\omega_0)^2$ times (for a second order system). The response time of such systems at a continuous useful signal is about $4 - 6T_0$ ($5T_0$ for the first order system and $4T_0/b$ for the second order system). If the signal given by the first or second order system is integrated over such a period, a supplementary attenuation for the alternating signal of about $4 - 6\omega/\omega_0$ can be obtained.

However, such structures are very sensitive at the random variations of the integration period (for unity-step input, the signal, which is integrated, is equal to unity at the sampling moment of time). Even if we use oscillators with a very high accuracy, such random variations will appear due to the fact that an electric current charging a capacitor usually performs the integration. This capacitor must be charged at a certain electric charge Q necessary for further conversions; this electric charge can't be smaller than a certain value Q_{lim}, while it has to supply a minimum value I_{min} for the electric current necessary for conversions on the time period t_{conv} required by these conversions, the relation

$$Q_{lim} = I_{min} t_{conv} \tag{3}$$

being valid. So the minimum value $I_{int}(min)$ for the electric current charging the capacitor in the integrator system is determined by the relation

$$I_{int}(min) = \frac{Q_{lim}}{t_{int}} \tag{4}$$

where t_{int} is the integration period required by the application (knowing the sampling frequency f_s, we can approximately establish t_{int} using the relation $t_{int} = 1/f_s$). So the current charging the capacitor can't be less than a certain value. thus random variations of the integration period will appear due to the fact that the random phenomena are generated when a nonzero electric current is switched off.

These random variations can't be avoided if we use asymptotically stable filters. By the other hand, an improvement in an electrical scheme used for integrators in analog signal processing (see [1], [2]) can't lead to a significant increasing in accuracy, as long as such electronic devices perform the same task (the system has the same transfer function). There are also known techniques for reducing the switching noise in digital systems, but such procedures can be applied only after the analog signal is filtered and sampled, so as to be prepared for further processing. So we must give attention to some other kind of transfer functions and to analyze their properties in case of filtering and sampling procedures.

Mathematically, an ideal solution consists in using an extended Dirac function for multiplying the received signal before the integration (see [3]), but is very hard to generate thus extended Dirac functions (a kind of acausal pulses) using nonlinear differential equations (see [4] for more details). So we must use some simple functions for solving our problem.

2 The Necessity of Using Oscillating Systems for Filtering the Received Signal

As it has been shown, first or second order stable systems are not suitable for filtering the received signal in case of integration and sampling procedures. They do not have the accuracy required by the operation

$$\frac{u(t_2) - u(t_1)}{t_2 - t_1} = \frac{u(t_2) - u(t_1)}{\Delta t} \tag{5}$$

We need a system having the following property: starting to work from initial null conditions, for a unity step input it must generate an output and a derivative of this output equal to zero at a certain moment of time (the condition for the derivative of the output to be equal to zero has been added so as the slope and the first derivative of the slope of the signal which is integrated to be equal to zero at the sampling moment of time, when the integration is interrupted). It is quite obvious that the single second order system possessing such properties is the oscillating second order system having the transfer function

$$H_{osc} = \frac{1}{T_0^2 s^2 + 1} \tag{6}$$

receiving a step input and working on the time interval $[0, 2\pi T_0]$. For initial conditions equal to zero, the response of the oscillating system at a step input with amplitude A will have the form

$$y(t) = A\left(1 - \cos\left(\frac{t}{T_0}\right)\right) \tag{7}$$

By integrating this result on the time interval $[0, 2\pi T_0]$, we obtain the result $2\pi A T_0$, and we can also notice that the quantity which is integrated and its

slope are equal to zero at the end of the integration period. Thus the influence of the random variations of the integration period (generated by the switching phenomena) is practically rejected.

Analyzing the influence of the oscillating system upon an alternating input, we can observe that the oscillating system attenuates about $(\omega/\omega_0)^2$ times such an input.

The use of the integrator leads to a supplementary attenuation of about $[(1/(2\pi)(\omega/\omega_0)]$ times. The oscillations having the form

$$y_{osc} = a\sin(\omega_0 t) + b\cos(\omega_0 t) \tag{8}$$

generated by the input alternating component have a lower amplitude and give a null result after an integration over the time interval $[0, 2\pi T_0]$.

As a conclusion, such a structure provides practically the same performances as a structure consisting of an asymptotically stable second order system and an integrator (response time of about $6T_0$, an attenuation of about $(1/6)(\omega/\omega_0)^3$ times for an alternating component having frequency ω) moreover being less sensitive at the random variations of the integration period. It is the most suitable for the operation

$$\frac{u(t_2) - u(t_1)}{\Delta t} \tag{9}$$

where $\Delta t = t_2 - t_1$. For restoring the initial null conditions after the sampling procedure (at the end of the working period) some electronic devices must be added. In the next section it will be shown that these devices must be represented by computer-driven electronic gates which must discharge certain capacitors at the end of each period of the oscillating system (a period corresponding to a working time).

3 The Necessity of Comparing the Value of the Derivative over Two Working Periods

The most simple structure having the transfer function

$$H = \frac{1}{T_0^2 s^2 + 1} \tag{10}$$

consists of some operational amplifier for lower frequency, with resistors R_0 connected at the (-) input and capacitors C_0 connected between the output and the (-) input (the well-known negative feedback) together with computer-driven electronic gates (for discharging these capacitors at the end of each working period); no resistors and capacitors are connected between the (+) connection and the "earth" (as required by the necessity of compensating the influence of the polarizing currents at the input of the amplifiers), so as to avoid instabilities of operational amplifiers at higher frequencies. In figure 1 is represented such a structure, having the period of oscillation of about $15\mu s$ (the capacity of C_1 and C_2, corresponding to C_0, being set to $69pF$). However, tests have shown that,

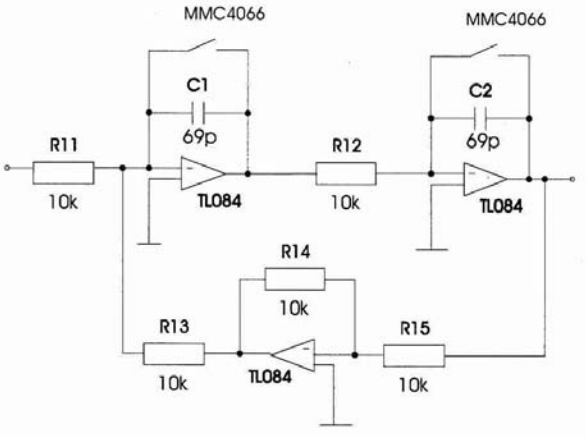


Fig. 1. Circuit for medium frequency

for a constant input A, the output is about $0.35A$ at the end of a period)instead of zero). So the capacitors were replaced by some others, having a capacity 10 times greater than the capacity of C_1 and C_2 in figure 1. Thus the time period became equal to $150\mu s$. For a constant input A, the output of the oscillating system is represented in figure 2. It can be noticed that the output is about $0.1A$ at the end of a complete oscillation. The output of the oscillating system can be integrated over a period using a similar device (based on an operational amplifier with a resistor R_i connected at the (-) input and a capacitor C_i connected on the negative feedback loop), at the end of the period the integrated signal being sampled. For a robust integration, we must chose as sampling moment of time the moment when the output is equal to zero (thus the working time interval presents a small difference as related to a period of the oscillating system, a scan be noticed studying figure 2). The time constants T_i - for the integrating system - and T_0 -for the oscillating system - have the form

$$T_i = R_i C_i, \;\; T_0 = R_0 C_0 \tag{11}$$

If the resistors R_0, R_i and the capacitors C_0, C_i are made of the same material, the coefficient for temperature variation will be the same for resistors and will be also the same for capacitors. Thus the ratio

$$A\frac{2\pi T_0}{T_i} = A\frac{2\pi R_0 C_0}{R_i C_i} = 2\pi A\left(\frac{R_0}{R_i}\right)\left(\frac{C_0}{C_i}\right) \tag{12}$$

(the result of the integration) is insensitive at temperature variations (for more details, see [4]).

However, for determining the derivative of the received signal we can't simply use the ratio

$$\frac{u(t_2) - u(t_1)}{t_2 - t_1} = \frac{u(t_2) - u(t_1)}{\Delta t} \tag{13}$$

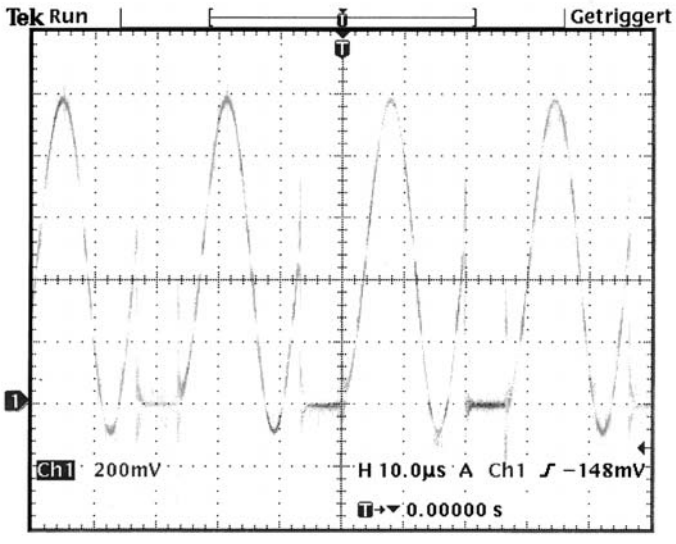


Fig. 2. Output of circuit for working period of 150 μs

while it is quite possible for the received signal to begin to change its value, with a constant slope, at a time moment within the working period $[0, 2\pi T_0]$ of the oscillating system. Thus we can't just consider the result obtained over two successive working periods (presented above) as the value of the derivative. We have to wait another working period, and then we must compare the values

$$\frac{u(t_2) - u(t_1)}{t_2 - t_1} = \frac{u(t_2) - u(t_1)}{\Delta t}, \quad \frac{u(t_3) - u(t_2)}{t_3 - t_2} = \frac{u(t_3) - u(t_2)}{\Delta t} \tag{14}$$

and only when the result of these two operations are almost equal we can assign their result to the value of the derivative of the received signal.

4 An Extension of the Notion of Observability

As it has been shown in the previous paragraph, in the conditions of a step input the output of an oscillating second order system possesses two components: a step component and an alternating component of angular frequency ω_0. The fact that this output and its derivative are equal to zero at the sampling moment of time can be connected with the notion of observability in systems theory. At the sampling moment of time, both the state variables $y(t)$ and dy/dt are equal to zero (are unobservable) and thus the signal which is integrated and its slope are equal to zero at this moment of time (the whole sampling structure is practically insensitive at the random variations of the integration period). However, there is a major difference between the notion of observability in this case and the usual notion of stability (considered for analog linear systems): in our case the

state variables are analyzed from the observability point of view only at certain moments of time (at the sampling moments of time).

5 Conclusions

This paper has presented a possibility of obtaining the derivative of the received electrical signal using a filtering device consisting of an oscillating second order system and an integrator. The oscillating systems is working on a time period for filtering a received electrical signal, with initial null conditions. The output of this oscillating system is integrated over this time period (at the end of this period the integrated signal being sampled). In the conditions of a unity-step input, the output of the oscillating system (the quantity which is integrated) is practically equal to zero at the sampling moment of time (when the integration is interrupted). The necessity of using two such oscillating systems if we intend to process the received signal in a continuous manner has been also presented. The method can be used for obtaining the derivative of the optoelectronic signal in case of phase detection for vibration measurements (see [5]).

References

1. Sun, Y., Fidler, J. K.: Synthesis and Performance Analysis of Universla Minimum -Component Integrator Based IFLF-OTA Grounded Capacitor Filter, IEE Proc.Circ.Dev.Syst., **143** 4 (1996) 107–114
2. Smith, S.L., Simenco, E.: Low Voltage Integrators for High-Frequency CMOS Filters Using Current-Mode Techniques, -IEEE Trans.Circ. Syst. II, An. and Dig. Signal Processing, **43**.1 (1996) 39–48
3. Toma, C. : An extension of the notion of observability at filtering and sampling devices, Proceedings of the International Symposium on Signals, Circuits and Systems Iasi SCS 2001, Romania 233–236
4. Toma, C. : Device and method for filtering high-frequency electrical signals, Romanian License of Invention 105124 (1989)
5. Sterian, A., Toma, C.: Phase detection for vibration measureemnts based on test-functions, SPIE Proceedings **5503**, 6th International Conference on Vibration Measurements based on Laser Technoques - Advances and Applications AIVELA 2004, Italy, 164–168

Simulating Laser Pulses by Practical Test Functions and Progressive Waves

Rodica Sterian[1] and Cristian Toma[2]

[1] Hyperion University, Department of Electronics and Computer Science, Bucharest, Romania
[2] Titu Maiorescu University, Department of Applied Informatics, Bucharest, Romania

Abstract. This study presents simulations performed in Matlab for the change of dielectric properties of mixtures of fatty acids under the influence of external optical or electrical pulses. These simulations are based on the use of practical test-functions. It is shown that, for an external observer, the behavior of the material medium for an alternating input of high-frequency (corresponding to an optical pulse) presents a single oscillation on the whole working interval, similar to the dynamics of ϵ noticed during the experiments. For the case of suddenly-applied electric fields, it is shown that the use of a similar dynamical model, in conjunction with the hypothesis of progressive waves generated inside mixtures of fatty acids by suddenly-applied electric fields, can simulate the dynamics of electrical properties of such mixtures in a correct manner.

1 Introduction

For studying dynamics of electrical properties for a mixture of fatty acids under the influence of external optical or electrical pulses, differential equations able to generate practical test-functions [1] are recommended. They has the advantage of generating functions which differ to zero only on a finite time interval, unlike methods based on unstable resonant orbits [2] or bifurcations [3]. This study presents simulations performed in Matlab for the change of dielectric properties of mixtures of fatty acids under the influence of external optical or electrical pulses. The behavior of the material medium is studied using differential equations corresponding to practical test-functions, with a free-term added (corresponding to the received pulse) and with initial null conditions.

2 Simulating the Dynamics of Dielectric Properties of Fatty Acids for External Optical Pulses

We are looking for a mathematical model for describing the dynamics of phenomena taking place inside the material, under the influence of external optical pulses. Due to the fact that the optical pulse possess a higher frequency, we can't

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 592–597, 2005.

use a linear equation of evolution, because the answer of such an equation at an external command of higher frequency is close to zero. It results that we must try nonlinear equations of evolution; we begin our study by looking for nonlinear dynamics able to generate pulses similar to test-functions. These functions (similar to a Dirac pulse) can be written under the form

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

where

$$t = t - t_{sym} \tag{2}$$

t_{sym} being defined as the middle of the working time interval. Such a function has nonzero values only for $t \in [-1,1]$. At the first step we are looking for a differential equation, which can have as solution this function φ. Such an equation can't generate the test function φ (the existence of such an equation of evolution, beginning to act at an initial moment of time would imply the necessity for a derivative of certain order n - noted φ^n to make a jump at this initial moment from the zero value to another value which differs to zero, and such an aspect would be in contradiction with the property of the test-functions to have continuous derivatives of any order on the whole real axis - in this case represented by the time axis). So it results that an ideal test-function can't be generated by a differential equation, but it is quite possible for such an equation to possess as solution a practical test function f (a function with nonzero values on the interval $t \in [-1,1]$. and a certain number of continuous derivatives on the whole time axis). So we will try to study evolutions depending only of the values $f, f^1, \ldots f^n$ (these values being equal to the values of $\varphi, \varphi^1, \ldots \varphi^n$ at a certain time moment very close to the initial moment $t = -1$). By trying equations as

$$f\,(1) = \frac{-2t}{t^2-1} f \tag{3}$$

(a first order differential equation), we obtain a pulse symmetric as related to $t = 0$. However, we try to use second order differential equations, which can admit as solution functions having the form

$$\varphi = exp\frac{a^2}{t^2-1}$$

with certain initial conditions; for a = 1 we must use the differential equation

$$f^{(2)}\,(t) = \frac{6t^4-2}{\left(t^2-1\right)^4} f\,(t) \tag{4}$$

and for $a = 0.1$ we must use the differential equation

$$f^{(2)}\,(t) = \frac{0.6t^4-0.36t^2-0.2}{\left(t^2-1\right)^4} f\,(t) \tag{5}$$

and with different initial conditions for $f, f^{(1)} \ldots$ (using Runge-Kutta functions in MATLAB) it resulted that the last equation can lead to functions similar to a rectangular unitary pulse (the amplitude is close to unity for more than 2/3 of the integration period).

At the second step we must change these equations of evolution, because all changes inside the material medium appear due to the external optical pulse. Due to this reason, we must consider initial null initial conditions for the system, and we must also add a free term in the differential equation - corresponding to the instantaneous value of the electrical field of the optical pulse (having a frequency of about $10^{14} Hz$). As working period we choose a value approximately equal to the period when the optical signal is received by the detector (about 0.2ms). So, the differential equation must be written as

$$f^{(2)}(t) = \frac{6t^4 - 2}{(t^2 - 1)^4} f(t) + u(t) \tag{6}$$

where u is represented by an alternating function with a frequency $10^{11} times$ greater than the working period (0.2 ms). By numerical simulations in MATLAB with Runge-Kutta functions we have obtained the results presented in figure 1 (for $u = sin\left(10^{11}\pi t\right)$) and in figure 2 (for $u = cos\left(10^{11}\pi t\right)$).

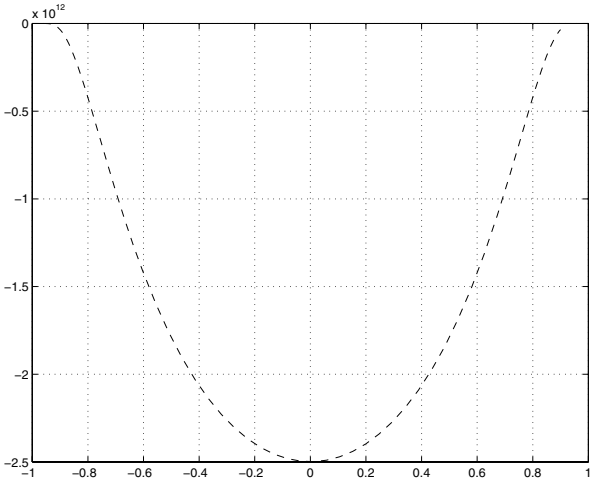


Fig. 1. f versus time for input $u = \sin(10^{11}\tau)$

The function f (generated by the material medium under the influence of the external optical pulse) can be integrated on this working time, the result z of this operation representing the physical quantity measured by the external observer (the dielectric constant ϵ of fatty acids under the influence of laser pulses, in our case). It can be noticed that, for the external observer, the behavior of the material medium is far of having a shape of an alternating signal, with a

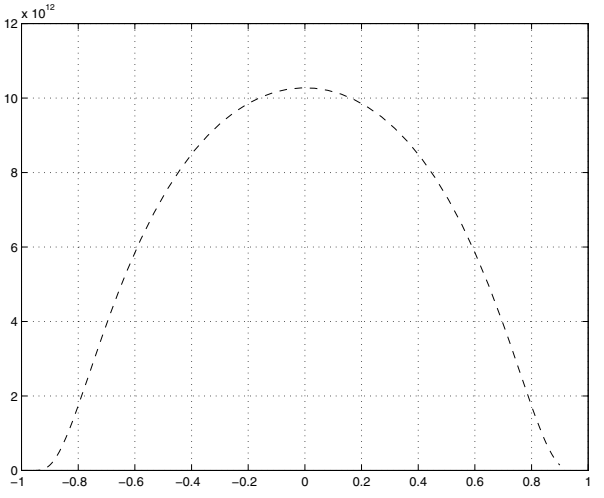


Fig. 2. f versus time for $u = \cos(10^{11}\tau)$

frequency similar to the frequency of the external optical pulse. A single oscillation on the whole working interval can be observed, similar to the dynamics of ϵ noticed during the experiments. This shows that the tackling way based on the use of systems described by differential equations able to generate pulses similar to test functions is correct.

3 Simulating the Dynamics of Dielectric Properties of Fatty Acids for Suddenly Applied Electric Fields

For simulating the dynamics of dielectric properties (the dielectric constant ϵ) of the fatty acids under the influence of an external electrical pulse (supposed to be constant over a certain time interval) we must consider that the function f (generated inside the material medium) is propagating toward the external observer under the form of a progressive wave. The external observer receives the influence of function f as a sum of all effects, which can be represented as an integral of function f multiplied by the progressive wave. Thus the dynamics of quantity z corresponding to the dielectric properties of the material can be written as

$$f^{(2)}(t) = \frac{6t^4 - 2}{(t^2 - 1)^4} f(t) + u(t) \tag{7}$$

$$z^{(1)}(t) = f(t) sin(\pi t - \Phi) \tag{8}$$

where Φ represents an initial phase of the progressive wave (caused by an internal response time inside the material). For

$$u(t) = 1, \;\; \Phi = \pi/12 \tag{9}$$

the function $z(t)$ is represented in figure 3, and for

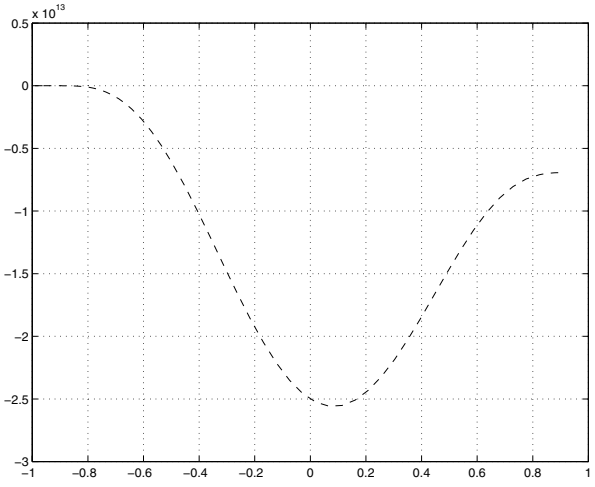


Fig. 3. z versus time for progressive wave, $\Phi = \pi/12$

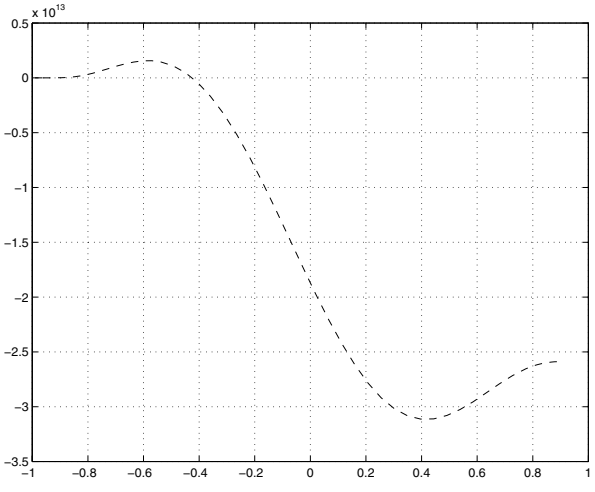


Fig. 4. z versus time for progressive wave, $\Phi = 5\pi/12$

$$u(t) = 1, \;\; \Phi = 5\pi/12 \tag{10}$$

the function $z(t)$ is represented in figure 4.

Studying figure 3, one can notice that function z is equal to zero at the initial time moment and then presents a minimum value at a time moment close to the initial time moment, before ending its evolution at a certain final value. This shape matches the dynamics of the dielectric constant ϵ under the influence of

an electric field suddenly applied to a material medium consisting of a mixture of fatty acids. So the use of dynamical models based on practical test-function (in conjunction with the hypothesis of progressive waves generated inside mixtures of fatty acids by suddenly-applied electric fields) can simulate the dynamics of electrical properties of such mixtures in a correct manner.

4 Conclusions

This study has presented simulations performed in Matlab for the change of dielectric properties of mixtures of fatty acids under the influence of external optical or electrical pulses. These simulations are based on the use of practical test-functions; the output of systems of differential equations simulating the behavior of the material medium was studied using Runge-Kutta equations. It was shown that, for an external observer, the behavior of the material medium for an alternating input of high-frequency (corresponding to an optical pulse) is far of having a shape of a high-frequency alternating signal. A single oscillation on the whole working interval can be observed, similar to the dynamics of ϵ noticed during the experiments.

For the case of suddenly-applied electric fields, it was shown that the use of a dynamical models based on practical test-function (in conjunction with the hypothesis of progressive waves generated inside mixtures of fatty acids by suddenly-applied electric fields) can simulate the dynamics of electrical properties of such mixtures in a correct manner. The shape of the output is equal to zero at the initial time moment and then presents a minimum value at a time moment close to the initial time moment, before ending its evolution at a certain final value, similar to the experimental observations. This shows that the tackling way based on the use of systems described by differential equations able to generate pulses similar to test functions is correct.

References

1. Toma, C. : An extension of the notion of observability at filtering and sampling devices, Proceedings of the International Symposium on Signals, Circuits and Systems Iasi SCS 2001, Romania 233–236
2. Carr, T.W., Bilings, L., Schwartz, I.B., Triandaf, I.: Bi-instability and the global role of unstable resonant orbits in a driven laser, Physica D **147** (2000) 59–82
3. Shilnikov, A., Nicolis, G., Nicolis, C. : Bifurcation and predictability analysis of a low-order atmospheric circulation model, Int. J. Bifur. Chaos Appl. Sci. Engrg. **5** (1995) 1701–1711

Statistical Aspects of Acausal Pulses in Physics and Wavelets Applications

Cristian Toma[1] and Rodica Sterian[2]

[1] Titu Maiorescu University, Department of Applied Informatics, Bucharest,Romania
[2] Hyperion University, Department of Electronics and Computers, Bucharest, Romania

Abstract. In the mathematical theory of distributions are widely used test-functions (which differ to zero only on a limited interval and have continuous derivatives of any order on the whole real axis). The use of such functions is also recommended in Fourier analysis of wavelets. However, less attention was given to connections between test-functions and equations used in mathematical physics (as wave equation). This paper shows that test-functions, considered at the macroscopic scale (that means not as δ -functions) can represent solutions for the wave-equation, under the form of acausal pulses (which appear under initial null conditions and without any source-term to exist). This implies the necessity for some supplementary requirements to be added to the wave-equation, so as the possibility of appearing such pulses to be rejected. It will be shown that such a possibility represents in fact a kind of bifurcation point, and a statistic interpretation (based on probability for state-variables to make certain jumps) is presented for justifying the fact that such pulses are not observed. Finally the advantage of using practical test function for wavelets processing is presented.

1 Introduction

As it is known, in Fourier analysis based on wavelets the user wants to obtain the mean value of the received signal multiplied by certain alternating functions over a limited time interval. Usually this operation is performed by a direct integration of the signal on this time interval. However, such structures are very sensitive at random variations of the integration period, due to stochastic phenomena appearing when an electric current is interrupted. For this reason, a multiplication of the received signal with a test-function - a function which differs to zero only on this time interval and with continuous derivatives of any order on the whole real axis - is recommended. Yet such test functions, similar to the Dirac functions, can't be generated by a differential equation. The existence of such an equation of evolution, beginning to act at an initial moment of time, would imply the necessity for a derivative of certain order to make a jump at this initial moment of time from the zero value to a nonzero value. But this aspect is in contradiction with the property of test-functions

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 598–603, 2005.

to have continuous derivatives of any order on the whole real axis, represented in this case by the time axis. So it results that an ideal test-function can't be generated by a differential equation (see also [1]); the analysis has to be restricted at possibilities of generating practical test-functions (functions similar to test-functions, but having a finite number of continuous derivatives on the whole real axis) useful for wavelets analysis. Due to the exact form of the derivatives of test-functions, we can't apply derivative free algorithms [2] or algorithms which can change in time [3]. Starting from the exact mathematical expressions of a certain test-function and of its derivatives, we must use specific differential equations for generating such practical test-functions.

For example, the bump-like function

$$\varphi(\tau) = \begin{cases} \exp\left(\frac{1}{\tau^2-1}\right) & \text{if} \quad \tau \in (-1,1) \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

is a test-function on $[-1,1]$. We are looking for an initial value problem for generating a practical test-function f on $[-1,1]$ by considering differential equations satisfied by the exact form of the amplitude and of the derivatives of the bump function φ. Such equations are

$$f^{(1)} = \frac{-2\tau}{\left(\tau^2-1\right)^2} f, \quad f(-0.99) = \varphi(-0.99) \tag{2}$$

$$f^{(2)} = \frac{6\tau^4-2}{\left(\tau^2-1\right)^4} f, \quad f(-0.99) = \varphi(-0.99),\ f^{(1)}(-0.99) = \varphi^{(1)}(-0.99) \tag{3}$$

Numerically integrations give solutions similar to φ, but having a very small amplitude.

2 Utility of Test-Functions in Mathematical Physics

Test-functions are known as having as limit the Dirac function when the interval on which they differ to zero decreases toward zero. However, less attention was given to the fact that such test-functions, considered at the macroscopic scale (that means not as Dirac-functions) can represent solutions for certain equations in mathematical physics (an example being the wave-equation). The main consequence of this consists in the possibility of certain pulses to appear as solutions of the wave-equation under initial null conditions for the function and for all its derivatives and without any free-term (a source-term) to exist. In order to prove the possibility of appearing acausal pulses as solutions of the wave-equation (not determined by the initial conditions or by some external forces) we begin by writing the wave-equation

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \phi}{\partial t^2} = 0 \tag{4}$$

for a free string defined on the length interval $(0,l)$ (an open set), where ϕ represents the amplitude of the string oscillations and v represents the velocity

of the waves inside the string medium. At the initial moment of time (the zero moment) the amplitude ϕ together with all its derivatives of first and second order are equal to zero. From the mathematical theory of the wave-equation we know that any solution of this equation must be a superposition of a direct wave and of a reverse wave. We shall restrict our analyze at direct waves and consider a supposed extension of the string on the whole Ox axis, ϕ being defined by the function

$$\phi(\tau) = \begin{cases} \exp\left(\frac{1}{(x-vt-1)^2-1}\right) \text{ for } x - vt < 1 \\ 0 \text{ for } x - vt \geq 1 \end{cases} \tag{5}$$

where $t \geq 0$. This function for the extended string satisfies the wave-equation (being a function of x-vt , a direct wave). It is a continuous function, having continuous partial derivatives of any order for $x \in (-\infty, \infty)$ and for $x \geq 0$. For $x \in (0, l)$ (the real string)the amplitude ϕ and all its derivatives are equal to zero at the zero moment of time, as required by the initial null conditions for the real string. We can notice that for $t = 0$ the amplitude ϕ and its partial derivatives differ to zero only on a finite space interval, this being a property of the functions defined on a compact set (test functions). But the argument of the exponential function is $x - vt$; this implies that the positive amplitude existing on the length interval $(-2, 0)$ at the zero moment of time will move along the Ox axis in the direction $x = +\infty$. So at some time moments $t_1 < t_2 < t_3 < t_4 < \ldots$ after the zero moment the amplitude ϕ will be present inside the string, moving from one edge to the other. It can be noticed that the pulse passes through the real string and at a certain time moment t_{fin} (when the pulse existing at the zero moment of time on the length interval $(-2, 0)$ has moved into the length interval $(l, l + 2)$) its action upon the real string ceases. We must point the fact that the limit points $x = 0$ and $x = l$ are not considered to belong to the string; but this is in accordance with the rigorous definition of derivatives (for this limit points can't be defined derivatives as related to any direction around them). The problem that a classical equation (such as the wave-equation) admits acausal solutions (for initial null conditions and without any external forces to exist) can be solved using deterministic methods, such as adding supplementary mathematical requirements to the wave-equation (the principle of least action, for example) or considering a causal chain:

a) external force (free-term) $\Longrightarrow$

$\Longrightarrow$ b) changes in the value of partial derivatives as related to space coordinates

$\Longrightarrow$ c) changes in the partial derivatives of the amplitude as related to time

$\Longrightarrow$ d) changes in the value of the function

so as the possibility of appearing acausal pulses (not yet observed) to be rejected. Such a causal chain can be represented in a mathematical form only as a differential equation able to generate functions similar to test functions, defined as practical test functions.

Another kind of method, based on statistical physics, is also available. Taking into account the fact that at the zero moment of time all derivatives of the amplitude of the real string are equal to zero on the whole length of the string

and after a very small time interval, at moment t' close to zero they may become different to zero in a small area inside it, we can consider the zero moment of time as a bifurcation point. At this moment of time there are several branches in the phase-space which satisfy the wave equation (the zero amplitude and the acausal pulse, for example). We consider the hypothesis that the string can choose a branch due to some stochastic jumps of the state-variables (the amplitude and some of its derivatives) around the zero moment of time, in a certain point. This implies that small changes in a small number of state-variables at the zero moment of time, imply a higher probability for that branch of evolution to appear. In the case we have presented, the acausal pulse (the test function) possess an infinite number of derivatives different to zero for any value of the argument for which the function differs to zero. This imply that a jump on this trajectory requires an infinite number of changes in the state-variables (the amplitude and its derivatives) at the edge $x = 0$ of the real string at a time moment t' very close to the zero moment. These changes have a very small module, but they establish in a very short time Δt the shape of the amplitude ϕ on a very small length Δx around the point $x = 0$ (the edge of the real string) along the positive part of the Ox axis (the real string). This nonzero amplitude appearing on length Δx can be considered as *part of an acausal pulse starting to move through the real string* .By noting these state-variables (the amplitude and its derivatives of different order at the point $x = 0$) with $a_0, a_1, \ldots a_k \ldots$ and by noting the state-variables of the acausal pulse at a moment of time t' very close to zero with $b_0, b_1, \ldots b_k \ldots$,we may write the probability of appearing a trajectory representing an acausal pulse as a consequence of such jumps under the form:

$$P_{ac} = P_0 \cap P_1 \cap P_2 \cap \ldots P_k \cap \ldots \tag{6}$$

In the previous equation P_0 is defined as

$$P_0 = P\left(a_0\left(t'\right) = b_0 \mid a_0\left(0\right) = 0\right) \tag{7}$$

and it represents the probability of the state-variable a_0 to become equal to b_0 at the time moment t' close to the zero moment, taking into account the fact that this state-variable was equal to zero at the zero moment of time. P_1 is defined as

$$P_1 = P\left(a_1\left(t'\right) = b_1 \mid a_1\left(0\right) = 0\right) \tag{8}$$

and it represents the probability of the state-variable a_1 to become equal to b_1 at the time moment t' close to the zero moment, taking into account the fact that this state-variable was equal to zero at the zero moment of time,.. P_k is defined as

$$P_k = P\left(a_k\left(t'\right) = b_k \mid a_k\left(0\right) = 0\right) \tag{9}$$

and it represents the probability of a state-variable a_k to become equal to b_k at the time moment t' close to the zero moment, taking into account the fact that this state-variable was equal to zero at the zero moment of time.

Considering possible independent jumps for each state-variable a_k and considering also that each factor P_k appearing in expression of P_{ac} is less than a certain value $m < 1$ (P_k corresponding to a probability) , we may write:

$$P_{ac} = P_0 \cap P_1 \cap P_2 \ldots \cap P_k \ldots \Rightarrow \tag{10}$$

$$P_{ac} = P_0 \cdot P_1 \cdot P_2 \ldots \cdot P_k \ldots \Rightarrow \tag{11}$$

(the probabilities P_k are considered to be independent, so P_{ac} is represented by the product of all P_k)

$$P_{ac} < P_0 \cdot P_1 \cdot P_2 \ldots \cdot P_n \Rightarrow \tag{12}$$

(all factors P_k, with $k > n$, are less than unity, so the right part of the previous equality increases if these factors are removed)

$$P_{ac} < m^n \Rightarrow \tag{13}$$

(because each factor is considered to be less than m, where $m < 1$)

$$P_{ac} \to 0 \; for \; n \to \infty \tag{14}$$

(the number of state variables a_k trends to infinite, and so n can be chosen as great as we want).

So $Pac \to 0$, the probability of appearing an acausal pulse being equal to zero. On the contrary, the probability for the string to keep its initial trajectory (the zero trajectory, which means that no changes in the amplitude appear) is very high, while this implies that the initial state-variables do not vary.

Another statistic method of solving this aspect consist in considering that at each moment the probability of appearing such an acausal pulse is equal to the amplitude of appearing an acausal pulse having the same amplitude, but with an opposite sign. So the resulting amplitude is be equal to zero and no motion appears.

3 Applications for Generating Wavelets

As shown in previous paragraph, acausal pulses similar to test-functions can't be generated by equations with partial derivatives (such as in [4]) such as the wave equation, due to the changes appearing at a certain moment of time, on a very small length Δx , for an infinite number of state-variables. The statistic method presented for justifying the fact that such acausal pulses are not observed implies also the fact that statistic computer methods for generating different functions using differential equations (by varying the initial conditions) are not adequate for test-functions. So for wavelets processing applications, we must use practical test-functions, generated by differential equations of evolution.

A first choice would be the use of a practical test-function for a primary multiplication of the received signal before multiplying this signal with an alternating function (the wavelet); yet this would imply two operations to be performed upon this received signal. This can be avoided if we use the associative property of multiplication. By noting the received signal with f, the wavelet with w and the practical test function with $\varphi(t)$, we can write the results z of both operations (the function which must be integrated) under the form

$$z(t) = w(t)\left[\varphi(t) f(t)\right] \tag{15}$$

or under the form

$$z(t) = [w(t)\varphi(t)](t) = w_\varphi(t) f(t). \tag{16}$$

The function $w_\varphi(t)$ can be obtained by multiplying the usual wavelet with a practical test function; it is used further for processing the received signal f (by multiplying and integrating the result on a limited time interval). Due to the fact that the values of the practical test function $\varphi(t)$ and of certain number of its derivatives are also equal to zero at the beginning and the end of the integrating period, the values of the function $w_\varphi(t)$ and of a certain number of its derivatives will be also equal to zero at these moments of time. Thus $w_\varphi(t)$ possess properties similar to test functions. Moreover, if the usual wavelet $w(t)$ is asymmetrical as related to the middle of the working period, than the function $w_\varphi(t)$ is also asymmetrical as related to this moment ($\varphi(t)$ being symmetrical as related to this moment) and the integral of $w_\varphi(t)$ on the whole real axis will be equal to zero. By adjusting the magnitude of the practical test function the integral of $[w_\varphi(t)]^2$ can be made equal to unity, and thus $w_\varphi(t)$ becomes also a wavelet. In this manner wavelets similar to test functions can be generated.

4 Conclusions

This paper has presented the possibility of some acausal pulses to appear as solutions of the wave-equation for a free string considered on the length interval $(0, l)$. Such pulses are in fact extended Dirac functions which can be imagined as coming from outside the string. It is shown that the possibility of appearing such pulses represents in fact a bifurcation in the phase-space of the string. This study tries to apply this concept to stochastic jumps on trajectories determined by test functions (having an infinite number of derivatives different to zero inside a limited open set and equal to zero outside it). Then the utility of using practical test-functions (functions similar to such extended Dirac-functions, which can be generated by a differential equation of evolution) in wavelets analysis is presented.

References

1. Toma, C. : Acausal pulses in physics-numerical simulations, Bulgarian Journal of Physics (to appear)
2. Morgado, J. M., Gomes, D.J. : A derivative - free tracking algorithm for implicit curves with singularities, Lecture Notes in Computer Science **3039** (2004) 221–229
3. Federl, P., Prudinkiewiez, P. : Solving differential equations in developmental models of multicellular structures using L-systems, Lecture Notes in Computer Science **3037** (2004) 65–82
4. Frankel, M., Roytburd, V. : Finite-dimensional model of thermal instability, Appl. Math. Lett. **8** (1995) 39–44

Wavelet Analysis of Solitary Wave Evolution

Carlo Cattani

DiFARMA, Università di Salerno
Via Ponte Don Melillo, I- 84084 Fisciano (SA)
ccattani@unisa.it

Abstract. The problem of solitary wave propagation in dispersive media is considered. The nonlinear equation (hyperbolic modification of Burger's one) in presence of a dispersive propagation law, give rise to nonlinear effects as the breaking down of the wave into localized chaotic oscillations. Finally, it is shown how to handle the representation of solitary profiles by using elastic wavelets.

1 Introduction

The hyperbolic modification of the Burgers equation [8] is

$$\tau\frac{\partial^2 u}{\partial t^2}+\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}-\sigma\frac{\partial^2 u}{\partial x^2}=0 \tag{1}$$

where τ is the relaxation time, σ viscosity , $u=u(x,t)$ velocity. The second order derivative has been considered in order to take into account memory effects, in particular when $\tau=0$ (absence of memory effects) we have the nonlinear Burger equation. This equation is the simplest modification of the one-dimensional Navier-Stokes system and includes in the parabolic-type equation non-local effects as it has been done by Cattaneo, Vernotte, Joseph [2, 7] in order to take into account memory effects.

The initial profile is taken in the form

$$u(x,0)=B\,F(x) \tag{2}$$

where $F(x)$ is a function with finite compact support [1, 3, 4, 5, 6, 9, 10, 11].

The solution of (1) in the form of simple (Riemann) waves is [3, 9, 10, 11]

$$u(x,t)=B\,F(z) \tag{3}$$

being $z=x-v\,t$. In particular, it is assumed, according to the dispersity hypothesis on the medium, that the phase velocity depends on z , i.e.

$$v=v(z) \quad , \quad z=x-v\,t \tag{4}$$

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 604 – 613, 2005.

so that the evolution implies a non linear dependence on the phase velocity giving rise to some nonlinear effects.

In the following, we will give the general solution of equation (1), under the hypothesis that the initial profile can be expressed, according to (3), as a given function which is deformed (in dependence of the phase velocity) during the evolution.

2 Method of Solution for the Propagation of a Solitary Wave

In order to simplify this model, and to derive the solution in the form of Riemann wave some hypotheses are done.

First assumption: It is assumed the weak dispersive law for materials, i.e. that during the evolution there results

$$\frac{\mathrm{d}\,v}{\mathrm{d}\,z} \cong 0 \tag{5}$$

in other words during the considered time interval the phase velocity is nearly constant while the kinetic energy is taken constant during the evolution. With this hypothesis we have $\frac{\partial u}{\partial t} = -B\ F' v,\ \frac{\partial^2 u}{\partial t^2} = B\ F'' v^2$ and $\frac{\partial u}{\partial x} = BF',\ \frac{\partial^2 u}{\partial x^2} = BF''$ with $(\cdot)' = \frac{d(\cdot)}{dz},\ (\cdot)'' = \frac{d^2(\cdot)}{dz^2}$.

Second assumption: Among all functions $F(z)$ we choose an eigenfunction, in the sense that

$$F' \equiv \frac{d\,F(z)}{dz} = f_1(z)\,F(z) \quad , \quad F'' \equiv \frac{d^2F(z)}{dz^2} = f_2(z)\,F(z) \tag{6}$$

with $F(z)$ given for all z (so that $f_1(z)$, $f_2(z)$ can be explicitly computed).

With this assumptions, equation (1) gives

$$\tau\ \ f_2\ \ v^2 - f_1\ \ v + B\,f_1\,F - \sigma\ f_2 = 0 \tag{7}$$

so that we have for the phase velocity ,

$$v(z) = \begin{cases} \dfrac{1}{2\,\tau\,f_2(z)}\ \left[f_1(z) \pm \sqrt{f_1^2(z) - 4\,B\,\tau\,f_1(z)\,f_2(z)\,F(z) + 4\tau\sigma f_2^2(z)} \right] & , \quad \tau \neq 0 \\ B\,F(z) - \sigma\,\dfrac{f_2(z)}{f_1(z)} & , \quad \tau = 0 \end{cases} \tag{8}$$

and for the $B\,F(z)$:

$$B\ F(z) = \sigma \frac{f_2(z)}{f_1(z)} + v(z) - \tau \frac{f_2(z)}{f_1(z)} \left[v(z)\right]^2 \tag{9}$$

There follows that the solution of system (1) is

$$u(x,t) = B\ F(z) = B\ F\left(x - v(z)\,t\right)$$

with $v(z)$ given by (8), that is

$$\begin{aligned} &u(x,t) = \\ &= B\ F\left\{x - \frac{1}{2\tau f_2(z)} \left[f_1(z) \pm \sqrt{f_1^2(z) - 4B\tau f_1(z) f_2(z) F(z) + 4\tau\sigma f_2^2(z)}\right] t\right\} \end{aligned} \tag{10}$$

when $\tau \neq 0$ and

$$u(x,t) = B\ F\left\{x - \left[B\ F(z) - \sigma \frac{f_2(z)}{f_1(z)}\right] t\right\} \tag{11}$$

for $\tau = 0$, or equivalently, according to (9),

$$u(x,t) = \sigma \frac{f_2\left(x - v(z)\,t\right)}{f_1\left(x - v(z)\,t\right)} + v(z) - \tau \frac{f_2\left(x - v(z)\,t\right)}{f_1\left(x - v(z)\,t\right)} \left[v(z)\right]^2 \tag{12}$$

The initial condition is

$$u(x,0) = B\ F(x) \tag{13}$$

If we give the initial functions as (13), and we look for a dispersive law (6) propagation, wave propagates according to (10)-(11), with phase velocity given by (8).

3 Example of the Evolution of Harmonic Wave

Equations (10)-(11) depend on the choice of the function $F(z)$ and then, according to (3), on the initial profile. In this section we consider a classical harmonic wave. In the next sections we will consider solitary profiles.

Let us take $F(x) = \cos x$. We have (see Fig. 1)

$$u(x,t) = B\cos\left\{x + \frac{1}{2\tau} \left[\tan z \pm \sqrt{\tan^2 z + 4B\tau \sin z + 4\tau\sigma}\right] t\right\} \tag{14}$$

$$u(x,t) = B\cos\left\{x - \left[B\cos\left(x - v_0\,t\right) - \sigma\,\mathrm{cotan}\left(x - v_0\,t\right)\right] t\right\}$$

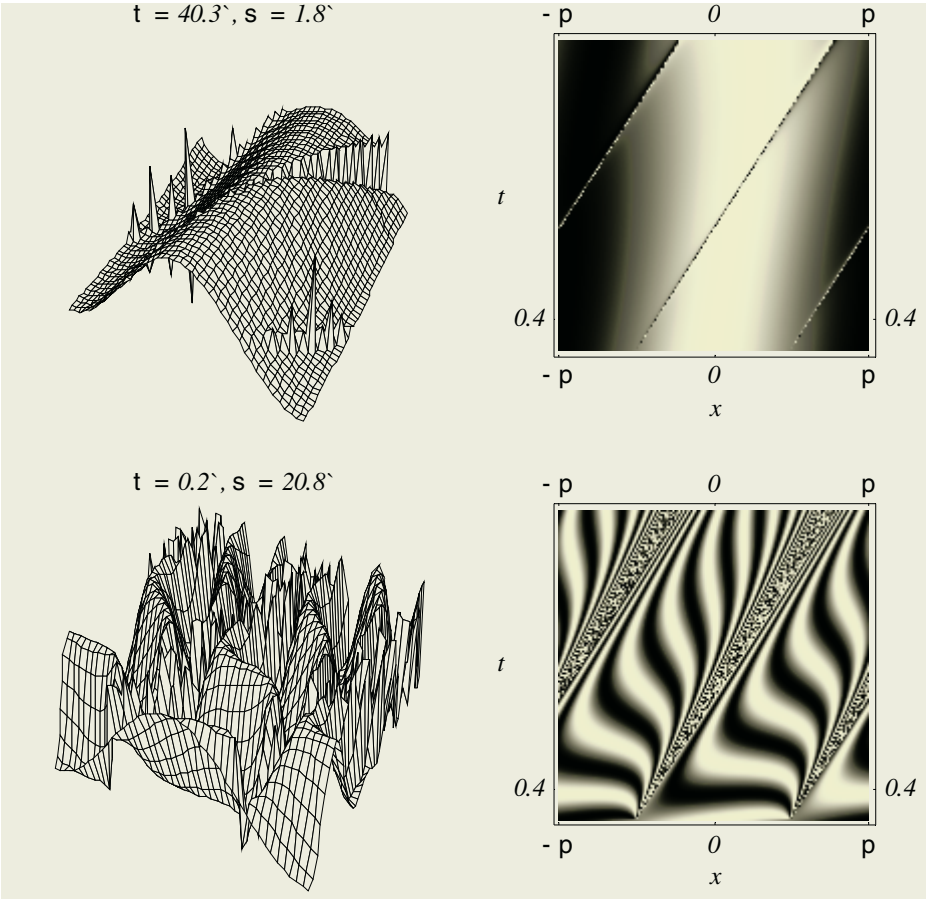

t = 40.3`, s = 1.8`
- p
0
p
t
0.4
0.4
- p
0
p
x
t = 0.2`, s = 20.8`
- p
0
p
t
0.4
0.4
- p
0
p
x

Fig. 1.

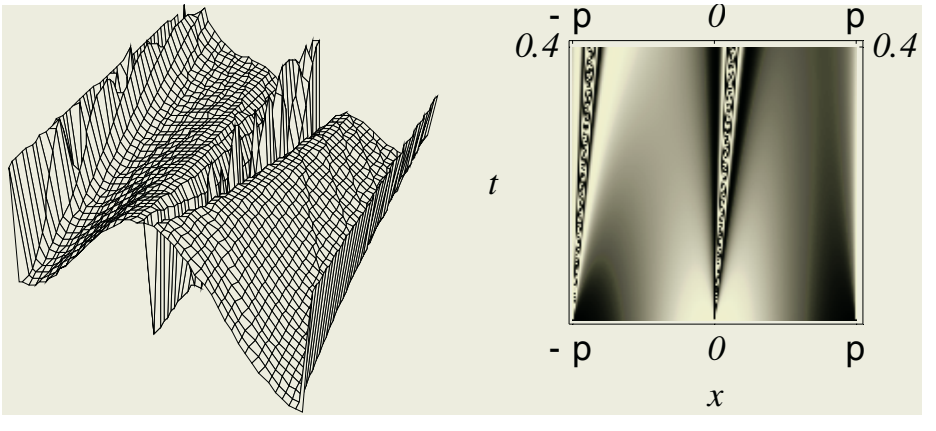

- p
0
p
0.4
0.4
t
- p
0
p
x

Fig. 2.

In absence of memory effect $\tau = 0$ it is:

$$u(x,t) = B\cos\left\{x - \left[B\cos(x - v_0\, t) - \sigma\,\mathrm{cotan}\,(x - v_0\, t)\right] t\right\}$$

In this case, with $\sigma = 2.7$ and $v_0 = 1$ the initial profile $u(x,0) = \cos x$ evolves as Fig. 2 .

In both cases there are some regions of chaotic behavior due to the viscosity $\sigma \neq 0$.

4 MH - Wavelets or Elastic Wavelets

As we have seen in the previous section the Riemann wave solutions are defined according to the initial function $F(x)$ that should be also an eigenfunction. If we want to take into account more general initial profiles we need to define a suitable representation of the initial profile into a suitable series of eigenfunctions. This can be done using wavelet families.

The MH-wavelet family is $\psi_{j,k}(x) = 2^{-j/2}\psi\left(2^{-j}x - 0{,}5k\right)$

$$\psi_{j,k}(x) = 2^{-j/2}\left(2/\sqrt{3}\right)\pi^{-1/4}\left[1-\left(2^{-j}x-\frac{k}{2}\right)^2\right]e^{-\left(2^{-j}x-\frac{k}{2}\right)^2\Big/2}. \tag{14}$$

If we define

$$A_{j,k}(x) = 2^{-j}x - \frac{k}{2} \quad , \quad \left(\frac{d}{dx}A_{j,k}(x) = 2^{-j}\right)$$

the first and second derivatives are

$$\frac{d}{dx}\psi_{j,k}(x) = \begin{cases} 2^{-j}\,\dfrac{A_{j,k}(x)\left[A_{j,k}^2(x)-3\right]}{\left[1-A_{j,k}^2(x)\right]}\,\psi_{j,k}(x) & , \quad x \neq 2^{j-1}(k\pm 2) \\ -2^{-j}\,\dfrac{A_{j,k}(x)}{\left[1-A_{j,k}^2(x)\right]}\,\psi_{j,k}(x) & , \quad x = 2^{j-1}(k\pm 2) \end{cases} \tag{15}$$

and

$$\frac{d^2}{dx^2}\psi_{j,k}(x) = 2^{-2j}\,\frac{\left[A_{j,k}^4(x)-6A_{j,k}^2(x)+3\right]}{\left[A_{j,k}^2(x)-1\right]}\,\psi_{j,k}(x) \tag{16}$$

which can be also written as

$$\frac{d^2}{dx^2}\psi_{j,k}(x) = 2^{-2j}\,\frac{\left[\left(A_{j,k}^2(x)-3\right)^2-6\right]}{\left[A_{j,k}^2(x)-1\right]}\,\psi_{j,k}(x)$$

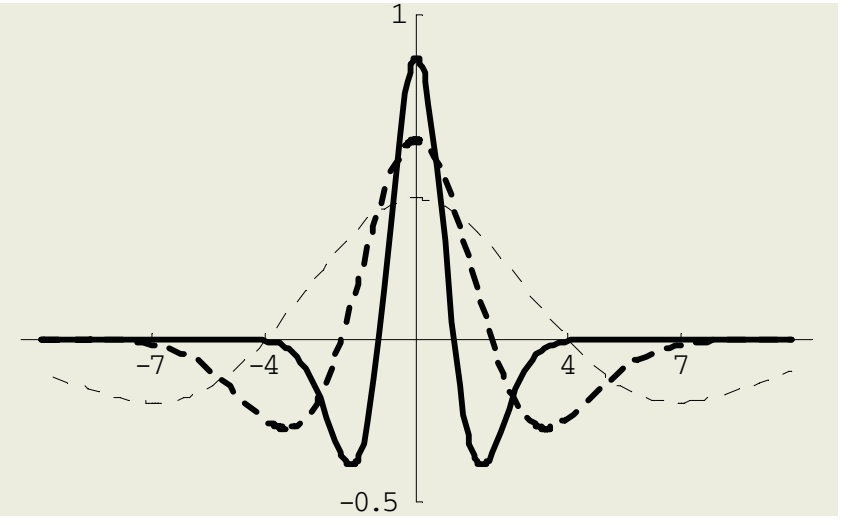


Fig. 3. The MH-wavelets $\psi_{0,0}(x)$ (plain), $\psi_{1,0}(x)$ (thick dashed), $\psi_{2,0}(x)$ (dashed)

Reconstruction Formula

An approximate (for choosing the finite numbers of scales from j_0 to j_{00}) representation of the function-signal using the wavelet (14) has the form

$$f(x) \approx \frac{1}{A}\sum_{j=j_0}^{j=j_{00}}\sum_{k\in Z} d_{j,k}\psi_{j,k}(x) \tag{17}$$

with $A \cong 6.819$ and

$$d_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx. \tag{18}$$

Of course the level of approximation depends on the coefficients $d_{j,k}$, which in turn depend on the function $f(x)$. For example the reconstruction of the function $H_0(x) = e^{-x^2/2}$ is given in the next section.

Reconstruction of the Function $H_0(x)$

The effective support (the interval, on which is accumulated the major weight of the function) of $H_0(x)$ is determined by the interval $(-5,5)$. It seems to be sufficient for the representation of any function to end the approximation up to certain scale levels from j_0 to j_{00} in order to fit as much as possible the convexity of the curve. While, in order to fit the "tails", it is enough to fix the upper bound of the translation parameter k_0. As a rough estimate of the approximation, we can take this function which, for a fixed family of wavelets, depends on $f(x), j_0, j_{00}, k_0$:

$$\varepsilon(f, j_0, j_{00}, k_0) = \left[\int_{-\infty}^{+\infty} \left(f(x) - \frac{1}{A} \sum_{j=j_0}^{j=j_{00}} \sum_{k=-k_0}^{k_0} d_{j,k} \psi_{j,k}(x) \right)^2 dx \right]^{1/2}$$

For $H_0(x)$, we have the following values of the error, therefore if we want to reconstruct $H_0(x)$, with an approximation error up to $\sim 3\ 10^{-3}$,

	$j_{00} = 2$	$j_{00} = 3$	$j_{00} = 4$	$j_{00} = 5$
$e(H_0, -2, j_{oo}, 10)$	0.189177	0.0961912	0.048710	0.0248426

	$j_{00} = 6$	$j_{00} = 7$	$j_{00} = 8$
$e(H_0, -2, j_{oo}, 10)$	0.0128928	0.0069158	0.00392709

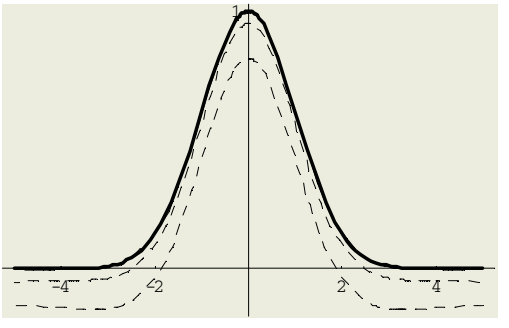


Fig. 4. Wavelet approximation of $H_0(x)$ at $j_{00} = 2$, $j_{00} = 4$ (dashed), and (plain) at the scale $j_{00} = 8$

it is enough to take $j_0 = -2, j_{00} = 8, k_0 = 10$, (see Fig. 4)

$$H_0(x) \approx (1/A) \sum_{j=-2}^{j=8} \sum_{k=-10}^{10} d_{j,k} \psi_{j,k}(x)$$

The total number of wavelet coefficients, $d_{j,k}$, $(j_0 = -2, j_{00} = 8, k_0 = 10)$, in the above approximate expansion is 231, but many coefficients have a very small values. Therefore by a suitable threshold we can select the coefficients (with corresponding wavelet) having a significant contribution. For example, if we neglect, in the above example, the coefficients less than 0.05, there remain only 45 significant coefficients and the approximation is still good (see Fig. 5, left), for a threshold 0.1 there remain 36 coefficients (see Fig. 5, center) , while for a threshold 0.2, we have only 24 coefficients, but the reconstruction is very coarse (see Fig. 5, right).

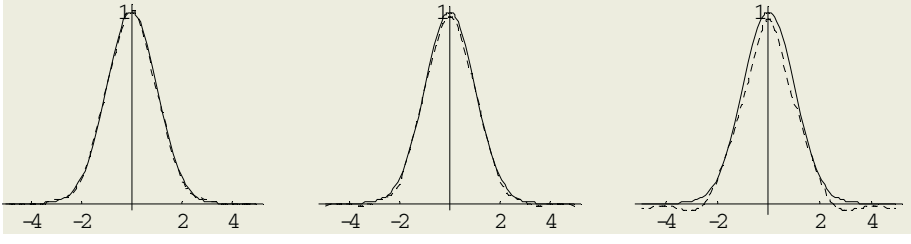


Fig. 5. Threshold reconstruction of $H_0(x)$ at $j_{00} = 2$, $j_{00} = 4$, with threshold 0.05, 0.1, 0.2

5 Solving the Problem on a Solitary Wave Using MH-Wavelets

The primary idea of using the system of MH-wavelets in solving the problem (1), (2) consists in representation of an initial profile of the wave through these wavelets

$$F(x) \approx F_4(x) = \frac{1}{A} \sum_{j=-4}^{4} \sum_{k=-20}^{20} d_{jk} 2^{-\frac{j}{2}} \psi\left(2^{-j}x - 0{,}5k\right) \tag{19}$$

and the assumption that the propagating in a weak dispersive medium initial profile will be changed slightly. So that the solution of problem (1), (2), expressed as a the evolution of a simple wave can be represented in the form like (7)-(15):

$$\begin{cases} u(x,t) = B\,F(z) \\ F(z) = (1/A)\sum_{j=j_o}^{j_{oo}}\sum_{k=-k_o}^{k_o} d_{j,k}\psi_{j,k}(z) \end{cases}$$

so that

$$\frac{d}{dz}F(z) = 2^{-2j}(1/A)\sum_{j=j_0}^{j_{00}}\sum_{k=-k_0}^{k_0} d_{j,k}\frac{A_{j,k}(z)\left[A_{j,k}^2(z)-3\right]}{\left[A_{j,k}^2(z)-1\right]}\psi_{j,k}(z)$$

and

$$\frac{d^2}{dz^2}F(z) = 2^{-2j}(1/A)\sum_{j=j_0}^{j_{00}}\sum_{k=-k_0}^{k_0} d_{j,k}\frac{\left[A_{j,k}^4(z)-6A_{j,k}^2(z)+3\right]}{\left[A_{j,k}^2(z)-1\right]}\psi_{j,k}(z)$$

Thus we have a hierarchical decomposition of waves (one for each level fixed j,k), which gives the solution according to (10).

References

1. Achenbach J .D., Wave propagation in elastic solids. Series "Applied Mathematic sand Mechanics". Vol.16 Amsterdam: North-Holland (1973)
2. Cattaneo C., Sur une forme de l'équation de la chaleur éliminant le paradoxe d'une propagation instantaée // C.R.Acad. Sc. Paris, (1958) 247, 431 – 433
3. Cattani C., Multiscale Analysis of Wave Propagation in Composite Materials // Mathematical Modelling and Analysis (2003). 8, № 4. 267-282.
4. Cattani C. The wavelet based technique in the dispersive wave propagation // Int. Appl. Mech. (2003) 39, № 4. 493 – 501
5. Dieulesaint E et Royer D. Ondes elastiques dans les solides. Application au traitement du signal. – Paris: Masson et Cie. (1974), 424
6. Hudson J.A. The excitation and propagation of elastic waves. - Cambridge: Cambridge University Press, (1980), 226
7. Joseph D., Preziosi L. //Rev. Mod. Phys. (1989) 61, № 2E. 47-73
8. Makarenko A.S. On blow-up solutions in turbulence// Physics Letters A (1997) 235, 391–397
9. Rushchitsky J.J., Cattani C., Terletska E.V. Analysis of solitary wave evolution in a material with the microstructure using wavelet analysis // Int.Appl.Mech. (2004) 40, № 3. 197 – 202

10. Rushchitsky J.J., Cattani C. Cubically nonlinear elastic waves versus quadratically ones: methods of analysis, propagation regularities, main effects // Int.Appl.Mech. (2003) 39, № 10. 1115 – 1145
11. Rushchitsky J. J., Tzurpal S. I. Waves in Materials with Microstructure, Institute of Mechanics of the Ukrainian Academy of Science (1997)

Numerical Analysis of Some Typical Finite Differences Simulations of the Waves Propagation Through Different Media

Dan Iordache[1], Stefan Pusca[1], and Ghiocel Toma[2]

[1] Politehnica University, Department of Physics, Bucharest, Romania
[2] Vimar Electro Co., Bucharest, Romania

Abstract. The Numerical Physics field has its specific phenomena , which intervene always but are confined usually to very restricted limits. However, the description possibilities (accuracy, computing time, etc) of computers are limited and sometimes - in strong connection with some specific features of the used algorithms and the computer errors - the numerical phenomena reach important amplitudes, the accomplished numerical simulations presenting significant distortions relative to the simulated (true) physical evolutions. That is why the goal of this work is to study the main features of some classical and newly found out numerical phenomena associated to the Finite Differences (FD) simulations of the wave propagation through media with sharp interfaces and attenuative character (considered as suddenly-emerging phenomena), and of other physical processes. The mechanisms of these numerical phenomena were studied in detail, the obtained findings allowing us to predict the distortions of the simulated physical processes. Of course, the good knowledge of the main features and of the mechanisms of the most important numerical phenomena allows us also to avoid the appearance of drastic distortions of the simulated evolutions, as well as the optimization of some numerical simulations.

1 Introduction

We live in a computerized world, our (21st century) civilization being a civilization of computers. The computers achieved a strong connection between the Technical Physics, Engineering and Numerical Approximation and Analysis, resulting Numerical Physics. This field has its specific phenomena, which intervene always but are confined usually to very restricted limits. However, the description possibilities (accuracy, computing time, etc) of computers are limited and sometimes - in strong connection with some specific features of the used algorithms and the computer errors - the numerical phenomena reach important amplitudes, the accomplished numerical simulations presenting significant distortions relative to the simulated (true) physical evolutions. Taking into account that the work of all present complex installations and devices is controlled by computers, the appearance (due to some numerical phenomena) of some important distortions of the simulated processes lead usually to major failures of the

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 614–620, 2005.

technical installations. Particularly, the events referring to the erroneous numerical simulation and design of the flight of the Patriot Missile which failed (with disastrous results) - during the Gulf War in 1991 - to stop a Scud Missile [1] and the self-destruction of the European space Agency's Ariane 5 rocket 37, at seconds after its launch, were both assigned to computer errors and their associated numerical phenomena. But why are computer simulations used? The frequent use of the numerical simulations in different scientific and technical problems has two main causes: a) the numerical simulations are considerably cheaper than the experimental determinations, b) they can be used even in inaccessible experimental conditions. The numerical simulations are accomplished by means of the computers, which are finite as action in time and in space. That is why the obtained numerical results are governed by specific laws, corresponding to a Virtual World. The computers errors generate the numerical phenomena, which correspond so to the Computers Virtual World. By the other side, we must think whether some FD methods doesn't correspond to the true world, the propagation of a wave being in this case represented by a succesion of some other phenomena specific to small space-time intervals.

The main numerical phenomena (instability, dispersion, non-convergence) specific to the Finite Difference (FD) simulations of the wave propagation in ideal media were observed even since 1920's years, when R. Courant, H. Friedrichs and W. Lewy [4] have established the conditions for the appearance of instabilities due to

$$C = \tau v_{\Phi} \ \epsilon > 1 \tag{1}$$

where C is the Courant's similitude criterion, τ, ϵ are the time and space FD steps, respectively, and for the appearance of dispersions due to

$$C = \tau v_{\Phi} / \epsilon < 1 \tag{2}$$

for the simulated pulse, during the propagation through an ideal medium (in such media, the phase propagation velocity v_{Φ} .is the same for all harmonic components of an arbitrary pulse). Later, prof. P.P. Delsanto (Politecnico di Torino) and his collaborators have pointed out [3] the appearance of some specific distortions in the case of the simulations with less than 1 values of the Courant's similitude criterion, the most important being the appearance of the opposite distortion, relative to the simulated pulse, which follows immediately this simulated pulse.

2 The Instability of the FD Simulations of the Ultrasonic Pulse Propagation

Taking into account that the knowledge of the causes which produce the instability of the numerical simulations is essential for a good understanding of this numerical phenomenon, we will analyze in following some studied cases, in relation with the instability cause.

a) Divergent oscillations and the monotonic divergence of the transfer coefficients

The study of the FD simulations of the interface phenomena (transmission, reflection) through/on a sharp interface between two homogeneous media [2] pointed out that between the components p'_m (for $t \geq N+2$) of the transmitted pulse and the components p_n $(n = 1, N)$ of the incoming pulse are related by the expression:

$$w_{1,m+2} = p'_m = \sum_{n=1}^{N} k_{m+1-n} p_n \tag{3}$$

where k_{m+1-n} are the transfer coefficients. If there is an upper limit M, so that:

$$k^2_{m+1-n} < M \text{ for any } m = 1, t-2 \text{ and } n = 1, N \tag{4}$$

the stability condition for the FD simulations becomes:

$$\sum_{i=I+1}^{\infty} |w_{i,t}|^2 = \sum_{m=1}^{t-2} |p'_m|^2 = \sum_{m=1}^{t-2} \left| \sum_{n=1}^{N} k_{m+1-n} P_n \right|^2 \leq N \sum_{m=1}^{t-2} \sum_{n=1}^{N} k^2_{m+1-n} p^2_n \tag{5}$$

and further

$$N \sum_{m=1}^{t-2} \sum_{n=1}^{N} k^2_{m+1-n} p^2_n < NM \sum_{m=1}^{t-2} \sum_{n=1} N p^2_n = NM(t-2) \sum_{-\infty}^{infty} |w_{i,0}|^2 \tag{6}$$

where $w_{i,t}$ represents the wave displacement in the FD node i at moment t, and I represents the value of i corresponding to the considered sharp interface. In these conditions, if there is the limit:

$$\xi = lim_{j \to \infty} \frac{k_{2j+1}}{k_{2j-1}} \tag{7}$$

and taking into account that for very large t we can write

$$k_t \approx \xi^{\frac{(t-1)}{2}} k_1 \tag{8}$$

it results that the FD schemes satisfy the stability condition only if $|\xi| \leq 1$.

Conversely, if $|\xi| \leq 1$, the FD scheme is unstable. A more detailed analysis shows that for $\xi > 1$ one obtains a divergent rise, while for $\xi < -1$ there appear some divergent oscillations of the simulated wave displacements. Generalizing these results, one finds that a such FD scheme is stable even it does not admit a limit , if there is a set of accumulation points $\xi_1, \xi_2, \ldots \xi_r$ which satisfy the condition $|\xi_r| \leq 1$ for any $i = 1, r$.

b) The "front" errors, corresponding to the simulation of the propagation of a finite sample of a monochromatic wave into an attenuative medium

Because the different monochromatic (Fourier) components of an ultrasonic pulse have different propagation velocities and present also different attenuation in the attenuative media, the simulation of the propagation of the monochromatic waves in such media is absolutely necessary. Taking into account that the computers cannot simulate an infinite wave (as it is the monochromatic one),

it is necessary to choose a characteristic wave sample (period or half-period). Since for the Courant's discretization $\epsilon/\tau = v_\Phi$, taking into account that for the Zener's media

$$\frac{\partial^2 \boldsymbol{w}}{\partial x^2} = \frac{\rho}{S}\frac{\partial^2 \boldsymbol{w}}{\partial t'^2} = \frac{\rho\cos\delta}{S}\exp -i\delta\frac{\partial^2 \boldsymbol{w}}{\partial t'^2} = \frac{\rho\cos^2\delta}{S}(1 - i\tan\delta)\frac{\partial^2 \boldsymbol{w}}{\partial t'^2} \tag{9}$$

the differential equation of the 1-D waves in these media:

$$\rho\frac{\partial^2 \boldsymbol{w}}{\partial x^2} = \boldsymbol{S}\frac{\partial^2 \boldsymbol{w}}{\partial t'^2} \tag{10}$$

reaches the discretized form

$$\boldsymbol{w}_{i+1} + \boldsymbol{w}_{i-1} - 2\boldsymbol{w} = c(1 - i\tan\delta)\left[\boldsymbol{w}_{t+1} + \boldsymbol{w}_{t-1} - 2\boldsymbol{w}\right] \tag{11}$$

where

$$c = \frac{\rho\cos^2\delta}{S}v_\Phi^2 = \frac{\cos\delta}{\cos^2\delta/2} \tag{12}$$

One finds so that if the complex displacements corresponding to the components of the chosen sample are given up to $\boldsymbol{w}_{i-1,t-1}$ (for the moment $t-1$, inclusively), and to $\boldsymbol{w}_{it} \equiv \boldsymbol{w}$ (for the moment t, inclusively), respectively, due to the absence of information concerning $\boldsymbol{w}_{i,t-1}, \boldsymbol{w}_{i+1,t}$, and $\boldsymbol{w}_{i+1,t-1}, \boldsymbol{w}_{i+2,t}$ respectively, the components $\boldsymbol{w}_{i+1,t-1}$ and $\boldsymbol{w}_{i+2,t}$ will be erroneous. So, for each new time step, two additional components of the simulated wave will be erroneous, the "front" errors spreading quickly and determining quickly the instability of the wave simulation. In order to avoid this unpleasant numerical phenomenon, it is necessary to substitute the two erroneous components from each time step by the corresponding correct values. This requirement can be accomplished if we simulate the space evolution, besides of the time evolution simulation, because in this case - for any space step - the wave is perfectly time periodical and the condition of time periodicity can be used. Of course, finally the space evolution can be converted into the time one, with errors considerably less than those corresponding to the "front" distortions.

c) The inverse wave generation for propagation in attenuative media

Using a FD lattice with the space ϵ and time τ steps, respectively:

$$x = I\epsilon, \;\; t' = t\tau \tag{13}$$

we obtain the discretized general solutions:

$$\boldsymbol{w}_{I,t} = A\exp(i\omega t\tau)\exp[\pm I(E + ik)\epsilon] \tag{14}$$

corresponding to the attenuated waves which propagate into the positive and negative directions of the Ox axis, respectively. Of course

$$\boldsymbol{w}_{I+1,t} = A\exp(i\omega t\tau)\exp[\pm(I+1)(E + ik)\epsilon] = \exp[\pm(E + ik)\epsilon]\boldsymbol{w}_{I,t} \tag{15}$$

therefore the product $A\exp[\pm I\epsilon(E + ik)]$ corresponds to the Fourier transform $\boldsymbol{w}_{I,t}^{FT}$ of the function $\boldsymbol{w}_{I,t}$, while the factors $\exp[\pm\epsilon(E + ik)]$ correspond to the

amplification factor $g(\xi)$ [5]. Taking into consideration the expression (10) and dividing the recurrence relation:

$$w_{i+1} = (p+q)w_{t+1} + (p-q)w_{t-1} - 2(1-p)w - w_{i-1} \tag{16}$$

where

$$p = \frac{\rho\epsilon 2}{S_0\tau^2} = \frac{\rho v_\Phi^2}{S_0} \tag{17}$$

and

$$q = \frac{R\epsilon^2}{2S_0\tau} = \frac{R\tau v_{|Phi}^2}{2S_0} \tag{18}$$

By $\boldsymbol{w} \equiv \boldsymbol{w}_{I,t}$, one obtains:

$$g + \frac{1}{g} = 2[\cosh E\epsilon \cos k\epsilon + i \sinh E\epsilon \sin k\epsilon] = 2\left[\left(p - i\frac{2q}{\omega\tau}\right)\cos\omega\tau + 1 - p + i\frac{2q}{\omega\tau}\right] \tag{19}$$

The equality of the last two members of the equation (14) leads to the dispersion-attenuation relations $k, E = f(\omega)$ corresponding to the FD scheme.

Both from equation (15), and from relation (19), one finds that the modulus of the first amplification factor $g_1 = \exp\epsilon(E + ik)$ corresponding to the Zener's differential (of the second order) equation is larger than 1:

$$|g_1| = \exp\epsilon E > 1 \tag{20}$$

It results that - according to the von Neumann's theorem - these FD schemes will be unstable. To understand the appearance of instabilities, we must underline that - though the initial conditions "launch" only the direct wave:

$$\boldsymbol{w}_{I,t}^{dir} = A \exp -EI\epsilon \exp i(\omega t\tau - kI\epsilon) \tag{21}$$

the errors introduced by the FD expressions of the DF partial derivatives produce the generation of the "inverse" wave:

$$\boldsymbol{w}_{I,t}^{inv} = A' \exp EI\epsilon \exp i(\omega t\tau + kI\epsilon) \tag{22}$$

Of course, the component of the simulated wave displacement corresponding to the "inverse" wave will increase continuously (for a "space" evolution along the positive direction of the axis Ox), determining finally the instability of the FD simulation.

Another reason of the instability corresponds to the mathematical solution:

$$\boldsymbol{w}_{I,t} = A \exp -i\omega t\tau \exp \pm(E + ik)I\epsilon \tag{23}$$

(equivalent to a wave amplification), admitted by the time and space symmetrical differential equations, as there are the equations (9), (10).

For this reason, the FD scheme corresponding to the complex stiffness is even more unstable than those corresponding to the wave asymmetric differential equations (as it is the FD scheme corresponding to the real wave-function).

To ensure an acceptable stability radius of the FD scheme with complex wave-function, it is necessary to use the analytical expression of the time derivative of the second order:

$$\frac{\partial^2 \boldsymbol{w}}{\partial t^2} = -\omega^2 \boldsymbol{w} \tag{24}$$

therefore the FD recurrence relation corresponding to equation (12) becomes:

$$\boldsymbol{w}_{i+1} = [2 - c\omega^2\tau^2(1 - i\tan\delta)]\boldsymbol{w} - \boldsymbol{w}_{i-1} \tag{25}$$

It is pointed out so the instability of the attenuated wave simulation, even for absolutely exact initial conditions, due to the generation of the amplified wave (mathematically possible, but without a physical meaning) by the stochastic local accumulation of some local "rounding" inaccuracies of the exact values corresponding to such waves, as well as to the extremely strong acceleration of the amplified wave generation when the complex wave-functions are used (stability and convergence radii of the magnitude order of $1dB$ or even smaller). The introduction of some: (i) corrective measures (the use of some analytical expressions of some partial derivatives, particularly), (ii) properly chosen effective parameters allows the weakening of these unpleasant numerical phenomena, ensuring stability and convergence radii of the magnitude order of $100dB$ [6], which represent sufficiently high values for accurate descriptions, by means of the finite difference method, of the cases of technical interest.

3 Conclusions

Taking into account that the most important ("living") part of the Numerical Physics refers to the Numerical phenomena met in frame of the usual simulations, this work focuses its study on the better knowledge the numerical phenomena and the optimizations of the numerical simulations.

Particularly, a general analysis of several numerical phenomena, which intervene in the FD descriptions of the wave propagation through different media was accomplished. These phenomena were identified in the frame of the classical classification of the main numerical phenomena of the FD schemes intended to the description of the wave propagation. For instability, there were three identified numerical phenomena. The pointed out numerical phenomena were analyzed in strong connection with their specific generating causes.

We believe that the study and a better knowledge of the FD numerical phenomena which intervene in the numerical calculations and simulations intended to the description of different physical processes (different physical applications of the search methods, numerical simulations of the wave propagation, diffusion and drift, etc) is useful:

a) in order to avoid major errors, leading to some catastrophic consequences (see the failures of the missiles, etc)

b) to avoid the frequent misleading pseudo-convergent results obtained in frame of certain numerical simulations

c) to accomplish some optimizations of the numerical simulations.

References

1. Skeel, R. : Round-off Errors and the Patriot Missile, SIAM News, **25**(4) (1992), 11–18
2. Delsanto, P., Iordache, D., Iordache, C., Ruffino, E.: Analysis of Stability and Convergence in FD Simulations of the 1-D Ultrasonic Wave Propagation, Mathl. Comp. Modelling, **25**(6) (1997), 19–29
3. Delsanto, P., Whitcombe, T., Chaskelis, H., Mignogna, R.: Connection Machine Simulation of Ultrasonic Wave Propagation in Materials. I: the One-dimensional Case, Wave Motion **16** (1992), 65–80
4. R. Courant, R., Friedrichs, K., Lewy, H.: Uber die partiellen differenzgleichungen der mathematischen physik, Math. Ann. **100** (32)(1928), 214–234
5. Strikwerda, J.: Finite Difference Schemes and Partial Difference Equations, Wadsworth-Brooks (1989)
6. Iordache, D., Scalerandi, M., Rugina, C., Iordache, V.,:Study of the Stability and Convergence of FD Simulations of Ultrasound Propagation through Nonhomogeneous Classical (Zener's) Attenuative Media, Romanian Reports on Physics, **50**(10) (1998), 703–716

B–Splines and Nonorthogonal Wavelets

Nikolay Strelkov*

Yaroslavl State University,
Sovetskaya, 14, Yaroslavl, 150000, Russia
strelkov@uniyar.ac.ru

Abstract. The necessary and sufficient conditions for the (nonorthogonal) wavelet multiresolution analysis with arbitrary (for example B-spline) scaling function are established.

The following results are obtained:

1) the general theorem which declares necessary and sufficient conditions for the possibility of multiresolution analysis in the case of arbitrary scaling function;
2) the reformulation of this theorem for the case of B-spline scaling function from W_2^m;
3) the complete description of the family of wavelet bases generated by B-spline scaling function;
4) the concrete construction of the unconditional wavelet bases (with minimal supports of wavelets) generated by B-spline scaling functions which belongs to W_2^m.

These wavelet bases are simple and convenient for applications. In spite of their nonorthogonality, these bases possess the following advantages: 1) compactness of set $\operatorname{supp}\psi$ and minimality of its measure; 2) simple explicit formulas for the change of level. These advantages compensate the nonorthogonality of described bases.

1 General Criterion

Let φ and ψ be some functions such that

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k \varphi(2x-k)\,, \tag{1}$$

$$\psi(x) = \sum_{k\in\mathbb{Z}} b_k \varphi(2x-k) \tag{2}$$

and let

$$V_0 = \operatorname{span}\left\{\varphi(\cdot-k)\right\}_{k\in\mathbb{Z}}\,,$$
$$V_1 = \operatorname{span}\left\{\varphi(2\cdot-k)\right\}_{k\in\mathbb{Z}}\,,$$

* This work is partly supported by Russian Basic Research Foundation (Grant 04-01-00590) and the program "Universities of Russia" (Grant 04.01.049).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 621–627, 2005.

$$W_0 = \operatorname{span}\{\psi(\cdot - k)\}_{k\in\mathbb{Z}}$$

be spaces generated by φ and ψ.

It can easily be checked that

$$\widehat{\varphi}(2t) = A(t)\widehat{\varphi}(t)\,, \qquad \widehat{\psi}(2t) = B(t)\widehat{\varphi}(t)\,, \tag{3}$$

where $\widehat{\varphi}$ and $\widehat{\psi}$ are Fourier images of φ and ψ,

$$A(t) = \frac{1}{2}\sum_{k\in\mathbb{Z}} a_k e^{-ikt}\,, \qquad B(t) = \frac{1}{2}\sum_{k\in\mathbb{Z}} b_k e^{-ikt}\,. \tag{4}$$

Note that if $f \in V_1$, $g \in V_0$, $h \in W_0$, and

$$f(x) = \sum_{k\in\mathbb{Z}} f_k\varphi(2x-k)\,, \quad g(x) = \sum_{k\in\mathbb{Z}} \alpha_k\varphi(x-k)\,, \quad h(x) = \sum_{k\in\mathbb{Z}} \beta_k\psi(x-k)\,,$$

then it can be shown that the equality $\widehat{f}(2z) = \widehat{g}(2z) + \widehat{h}(2z)$ may be rewrite in the form

$$\frac{1}{2}\widehat{\varphi}(z)\sum_{k\in\mathbb{Z}} f_k e^{-ikz} = \widehat{\varphi}(2z)\sum_{k\in\mathbb{Z}} \alpha_k e^{-2ikz} + \widehat{\psi}(2z)\sum_{k\in\mathbb{Z}} \beta_k e^{-2ikz}\,,$$

i.e. (see (3))

$$\widehat{\varphi}(z)F(z) = A(z)\widehat{\varphi}(z)G(2z) + B(z)\widehat{\varphi}(z)H(2z)\,. \tag{5}$$

Here F, G, H are 2π–periodical functions such that

$$F(z) = \frac{1}{2}\sum_{k\in\mathbb{Z}} f_k e^{-ikz}\,, \quad G(z) = \sum_{k\in\mathbb{Z}} \alpha_k e^{-ikz}\,, \quad H(z) = \sum_{k\in\mathbb{Z}} \beta_k e^{-ikz}\,,$$

A and B are 2π–periodical functions of the form (4). It follows from (5) that 2π–periodical functions F, G, and H satisfy the following equality:

$$F(z) = A(z)G(2z) + B(z)H(2z). \tag{6}$$

Substituting $z + \pi$ by z in (6), we obtain

$$F(z+\pi) = A(z+\pi)G(2z) + B(z+\pi)H(2z)\,.$$

Note that we can treat two last equalities as linear system

$$\begin{cases} A(z)G(2z) + B(z)H(2z) = F(z), \\ A(z+\pi)G(2z) + B(z+\pi)H(2z) = F(z+\pi) \end{cases}$$

with unknown $G(2z)$ and $H(2z)$. Thus if

$$A(z)B(z+\pi) \neq A(z+\pi)B(z),$$

then

$$G(2z) = \frac{F(z)B(z+\pi) - F(z+\pi)B(z)}{A(z)B(z+\pi) - A(z+\pi)B(z)}\,,$$
$$H(2z) = \frac{F(z+\pi)A(z) - F(z)A(z+\pi)}{A(z)B(z+\pi) - A(z+\pi)B(z)}\,. \tag{7}$$

Let $a * b$ denote the quasiconvolution of sequences a and b:

$$(a*b)_n = \sum_{s\in\mathbb{Z}} (-1)^s a_{n-s} b_s, \quad n \in \mathbb{Z}\,.$$

The following assertion describes the necessary and sufficient condition for the decomposition of V_1 into the direct sum of V_0 and W_0 (it means that for every $f \in V_1$ there exists one and only one representation $f = g + h$, where $g \in V_0$, $h \in W_0$). Moreover, this decomposition is constructive.

Theorem 1. *The decomposition $V_1 = V_0 \dot{+} W_0$ holds if and only if there exist $N \in \mathbb{Z}$ and $C \neq 0$ such that*

$$(a*b)_{2p+1} = C\delta_N^p \quad \textit{for all} \quad p \in \mathbb{Z}\,. \tag{8}$$

If (8) holds, then for every $f \in V_1$ there exist unique $g \in V_0$ and $h \in W_0$ such that $f = g + h$, and if

$$f(x) = \sum_{k\in\mathbb{Z}} f_k \varphi(2x-k)\,,$$

then

$$g(x) = \sum_{k\in\mathbb{Z}} \alpha_k(f)\varphi(x-k), \quad h(x) = \sum_{k\in\mathbb{Z}} \beta_k(f)\psi(x-k),$$

where

$$\alpha_k(f) = \frac{1}{C}(f*b)_{2k+2N+1}\,, \quad \beta_k(f) = -\frac{1}{C}(f*a)_{2k+2N+1}, \quad k \in \mathbb{Z}\,. \tag{9}$$

Remark 1. Note that this theorem is the consequence of equalities (7) above. Indeed, it is easy to prove that the denominator of (7) has the following form:

$$A(z)B(z+\pi) - A(z+\pi)B(z) = \frac{1}{2}\sum_{k\in\mathbb{Z}} (a*b)_{2k+1} e^{-i(2k+1)z}.$$

Similarly,

$$F(z)B(z+\pi) - F(z+\pi)B(z) = \frac{1}{2}\sum_{k\in\mathbb{Z}} (f*b)_{2k+1} e^{-i(2k+1)z},$$

$$F(z+\pi)A(z) - F(z)A(z+\pi) = \frac{1}{2}\sum_{k\in\mathbb{Z}} (a*f)_{2k+1} e^{-i(2k+1)z}.$$

Thus if (8) holds, then it follows from (7) that

$$G(2z) = \frac{1}{C} \sum_{k \in \mathbb{Z}} (f * b)_{2k+1} e^{-2i(k-N)z} = \frac{1}{C} \sum_{k \in \mathbb{Z}} (f * b)_{2k+2N+1} e^{-2ikz},$$

$$H(2z) = \frac{1}{C} \sum_{k \in \mathbb{Z}} (a * f)_{2k+1} e^{-2i(k-N)z} = -\frac{1}{C} \sum_{k \in \mathbb{Z}} (f * a)_{2k+2N+1} e^{-2ikz}.$$

This proves equalities (9).

Remark 2. If $\{\varphi(\cdot - k)\}_{k \in \mathbb{Z}}$ is the L_2–orthonormal system, then it follows from (1) that

$$\sum_{k \in \mathbb{Z}} a_k a_{k+2p} = 2\delta_p^0 \tag{10}$$

for all $p \in \mathbb{Z}$. Moreover, if

$$b_k = (-1)^k a_{1-k}, \quad k \in \mathbb{Z} \tag{11}$$

and ψ is of the form (2), then $\{\psi(\cdot - k)\}_{k \in \mathbb{Z}}$ is the L_2–orthonormal system and then it follows from (1) that

$$(\varphi(\cdot - p), \psi(\cdot - s))_{L_2(\mathbb{R})} = 0$$

for all $p, s \in \mathbb{Z}$ (see, for example, [1]). Note that

$$(a * b)_{2p+1} = \sum_{k \in \mathbb{Z}} a_{2p+1-k} a_{1-k} = 2\delta_p^0$$

for all $p \in \mathbb{Z}$ (see (10), (11)) and (8) holds with $C = 2, N = 0$.

Remark 3. It is obvious that the same condition (8) is necessary and sufficient for the decomposition $V_{j+1} = V_j \dot{+} W_j$ for every $j \in \mathbb{Z}$, where

$$V_j = \operatorname{span}\left\{\varphi(2^j \cdot - k)\right\}_{k \in \mathbb{Z}}, \quad W_j = \operatorname{span}\left\{\psi(2^j \cdot - k)\right\}_{k \in \mathbb{Z}}.$$

Note that this decomposition implies the representation

$$V_j = W_{j-1} \dot{+} W_{j-2} \dot{+} \ldots, \quad j \in \mathbb{Z}.$$

Theorem 1 enables us to construct (for the given function φ satisfying (1)) function ψ of the form (2) such that $V_{j+1} = V_j \dot{+} W_j$ for all $j \in \mathbb{Z}$.

2 *B*–Spline Scaling Function

Let us apply Theorem 1.1 to the case of Shoenberg B–spline scaling functions.

Let $\varphi \in W_2^m(\mathbb{R})$ be the m–order B–spline ($(m+1)$–multiple convolution of characteristic function of $[0, 1]$) with $\operatorname{supp} \varphi = [0, m+1]$. Then (see (1))

$$a_k = 2^{-m} \begin{pmatrix} m+1 \\ k \end{pmatrix};$$

therefore (8) and (9) take the form

$$\nabla^{m+1} b_{2p+1} = C\delta_N^p \quad \text{for all } p \in \mathbb{Z}, \tag{8'}$$

$$\alpha_k(f) = \frac{2^m}{C}(b * f)_{2k+2N+1}, \quad \beta_k(f) = \frac{1}{C}\nabla^{m+1} f_{2k+2N+1}, \tag{9'}$$

where $\nabla f_n = f_n - f_{n-1}$ is the backward finite difference. If (8') holds, then

$$\{\psi(2^s \cdot -k)\}_{s,k\in\mathbb{Z}}$$

is the unconditional basis of $W_2^m(\mathbb{R})$. The simplicity of (9') is the compensation for nonorthogonality of this basis.

Thus the condition (8') gives the full description of the class of wavelet W_2^m–bases generated by m–order B–spline.

The condition (8') may be treated as follows: there exist $N \in \mathbb{Z}$ and $u, v \in V_0$ such that $\psi \in V_1$ is of the form

$$\psi(x) = \begin{cases} u(x) & \text{if } x < N + 1/2, \\ v(x) & \text{if } x > N + 1/2 \end{cases}$$

and $\psi^{(m)}$ has nonzero jump at the point $x = N + 1/2$:

$$[\psi^{(m)}](N + 1/2) = \psi^{(m)}(N + 1/2 + 0) - \psi^{(m)}(N + 1/2 - 0) \neq 0$$

or (what is the same) there exists $N \in \mathbb{Z}$ such that the m–th derivative $\psi^{(m)}$ of $\psi \in V_1$ is continuous at all points $x = p + 1/2$, $p \in \mathbb{Z} \setminus \{N\}$ and discontinuous at the point $x = N + 1/2$.

3 Construction of Wavelet Basis

Let us construct function ψ with property (8') and support $[0, m]$ of minimal measure. It is convenient to describe this function in terms of its Fourier image.

Let $m \geq 0$ be fixed integer and let

$$\widehat{\varphi}(t) = \left\{ \frac{1 - e^{-it}}{it} \right\}^{m+1}, \quad \widehat{\psi}(t) = \left\{ \frac{1 - e^{-it/2}}{it} \right\}^{m+1} S_m(e^{-it/2})$$

be Fourier images of φ and ψ. Here $S_0(x) \equiv 1$ and if $m \geq 1$, then

$$S_m(x) = \sum_{n=0}^{m-1} s_n^m x^n$$

is the polynomial with coefficients

$$s_n^m = m \sum_{k=n}^{m-1} \frac{(-1)^{n+k+1} 2^{m-k}}{(m+k)(m+k+1)} \binom{m-1}{k} \binom{k}{n}, \quad n = 0, \ldots, m-1.$$

Then

1. $\varphi \in W_2^m(\mathbb{R})$ is m–order Shoenberg B–spline, $\operatorname{supp} \varphi = [0, m+1]$.
2. $\psi \in W_2^m(\mathbb{R})$ is piecewise polinomial function with support $[0, m]$ (if $m = 0$, then $\operatorname{supp} \psi = [0, 1/2]$).
3. If
$$V_0 = \operatorname{span} \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}} ,$$
$$V_1 = \operatorname{span} \{\varphi(2 \cdot - k)\}_{k \in \mathbb{Z}} ,$$
$$W_0 = \operatorname{span} \{\psi(\cdot - k)\}_{k \in \mathbb{Z}} ,$$
then $V_1 = V_0 \dot{+} W_0$, i.e. for every $f \in V_1$ there exists one and only one representation
$$f = g + h, \tag{12}$$
where $g \in V_0$, $h \in W_0$. More precisely, if
$$f(x) = \sum_{k \in \mathbb{Z}} f_k \varphi(2x - k),$$
then
$$g(x) = \sum_{k \in \mathbb{Z}} \alpha_k(f) \varphi(x - k), \quad h(x) = \sum_{k \in \mathbb{Z}} \beta_k(f) \psi(x - k),$$
where
$$\alpha_k(f) = \frac{2^m}{c} \sum_{n \in \mathbb{Z}} (-1)^{n+1} s_n^m f_{2k+1-n} , \tag{13}$$
$$\beta_k(f) = \frac{2^m}{c} \sum_{n \in \mathbb{Z}} (-1)^n \binom{m+1}{n} f_{2k+1-n} = \frac{2^m}{c} \nabla^{m+1} f_{2k+1} . \tag{14}$$
Here s_n^m are the coefficients of $S_m(x)$,
$$c = \sum_{k=0}^{m} \frac{(-1)^{k+1}}{2k-1} \binom{m}{k} ,$$
$\nabla f_p = f_p - f_{p-1}$ is the backward finite difference.
4. $\{\psi(2^n \cdot - k)\}_{n,k \in \mathbb{Z}}$ is the unconditional basis of $W_2^m(\mathbb{R})$.

Remark 4. These bases are nonorthogonal with the exception of the case $m = 1$ for which $\{\psi(2^n \cdot - k)\}_{n,k \in \mathbb{Z}}$ is the orthogonal basis of W_2^1 with the scalar product
$$(f, g)_{W_2^1(\mathbb{R})} = (f', g')_{L_2(\mathbb{R})}.$$
It is consequence of the following fact: derivatives
$$\{\psi'(2^n \cdot - k)\}_{n,k \in \mathbb{Z}}$$
generate the orthonormal (in $L_2(\mathbb{R})$) Haar basis.

Remark 5. Wavelet bases constructed in this section are rather simple and convenient for applications. In spite of their nonorthogonality, these bases possess the following advantages: 1) compactness of set $\operatorname{supp}\psi$ and minimality of its measure; 2) simple explicit formulas for the change of level (see (13), (14)), i.e., simple evaluation of g and h in (12) (in author's opinion, in the orthogonal case this process is often more complicated). These advantages compensate the nonorthogonality of described bases.

Reference

1. Daubechies, I.: Ten lectures on wavelets. SIAM, Philadelphia, Pennsylvania (1992)

Optimal Wavelets

Nikolay Strelkov and Vladimir Dol'nikov*

Yaroslavl State University,
Sovetskaya, 14, Yaroslavl, 150000, Russia
strelkov@uniyar.ac.ru

Abstract. The complete description of wavelet bases is given such that each of them is generated by the fixed function whose Fourier image is the characteristic function of some set. In particular, for the case of Sobolev spaces wavelet bases with the following property of universal optimality are constructed: subspaces generated by these functions are extremal for the projection-net widths (if $n = 1$, then also for Kolmogorov widths) of the unit ball in $W_2^m(\mathbb{R}^n)$ with $W_2^s(\mathbb{R}^n)$-metric for the whole scale of Sobolev classes simultaneously (i.e., for all $s, m \in \mathbb{R}$ such that $s < m$). Some results concerning completeness and basis property of exponential systems are established in passing.

1 Notation and Definitions

Let H^μ be Hilbert space with the scalar product

$$(u, v)_\mu = \int_{\mathbb{R}^n} \mu(x)\widehat{u}(x)\overline{\widehat{v}(x)}dx\,,$$

where $\widehat{u}$ is Fourier transform of u and μ is positive function defined on $\mathbb{R}^n$ for which there exist positive constants C and N such that

$$\mu(\xi + \eta) \le (1 + C|\xi|)^N \mu(\eta) \tag{1}$$

for all $\xi, \eta \in \mathbb{R}^n$. The symbol $\mathbf{M}$ denotes the set of all such *moderately increasing weight functions* (see [1], Definition 10.1.1).

Note that if $\mu(\xi) = \mu_s(\xi) = (1 + |\xi|^2)^s$, where $s \in \mathbb{R}$, then the spaces H^μ become Sobolev spaces $W_2^s(\mathbb{R}^n)$ (see, for example, [2], [3]). It is easy to show that for any $s \in \mathbb{R}$ the weight function μ_s belongs to the set $\mathbf{M}$ (for this function inequality (1) holds with the constants $C = 1\,, N = 2|s|$).

Let $f_1, \dots, f_n$ be a set of linearly independent elements of $\mathbb{R}^n$. We denote by $\Lambda = \Lambda(f_1, \dots, f_n)$ the lattice in $\mathbb{R}^n$ with basis $f_1, \dots, f_n$ [4], that is, the set of points of the form $\alpha = \alpha_1 f_1 + \dots + \alpha_n f_n$, where α_k take on arbitrary integer values. The notation Λ^* will be used for the lattice conjugate to Λ (that is, for the lattice whose basis forms a biorthonormal pair with the basis of Λ).

* This work is partly supported by Russian Basic Research Foundation (Grant 04-01-00590) and the program "Universities of Russia" (Grant 04.01.049).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 628–634, 2005.

Let

$$\{C_\gamma\}_{\gamma\in Q} \tag{2}$$

be a system of measurable sets of the space $\mathbb{R}^n$, where Q is a collection of indices of an arbitrary nature.

The system (2) is a *packing* (see, for example, [4]) if pairwise intersections of sets of this system have measure zero.

We say that the system (2) is a *tiling* of the measurable set $C \subseteq \mathbb{R}^n$ if, firstly, (2) is a packing and secondly, the measure of the symmetric difference of the union of the sets of system (2) and the set C equals zero.

In what follows, we shall often encounter the situation when the sets of the system (2) are translations of a fixed measurable set by vectors of a lattice, that is, Q is a lattice in $\mathbb{R}^n$, and $C_\gamma = \Omega + \gamma$, $\gamma \in Q$, where Ω is measurable set in $\mathbb{R}^n$. In this case we sometimes say that Ω generates a *Q–lattice packing* (respectively, *Q–lattice tiling*).

2 Preliminary Results

First we shall formulate two auxiliary assertions (which seem to be of interest in their own right) regarding the completeness and basis property in $L_2(\Omega)$ of the system of exponentials

$$\left\{e^{i(\alpha,x)}\right\}_{\alpha\in\Lambda} \tag{3}$$

(here Λ is a lattice in $\mathbb{R}^n$).

Lemma 1. *The system of functions (3) is complete in $L_2(\Omega)$ if and only if the set Ω generates a $2\pi\Lambda^*$–lattice packing.*

Lemma 2. *The system of functions (3) forms a basis of the space $L_2(\Omega)$ if and only if the set Ω generates a $2\pi\Lambda^*$–lattice tiling of the space $\mathbb{R}^n$. Moreover, the basis (3) is orthogonal.*

Remark 1. The assertions formulated above deal with the completeness and basis property of the system (3) in the spaces of L_2 type, which is quite sufficient for our purposes of constructing optimal wavelet bases. Nevertheless we note that these assertions remain valid for the spaces of L_p type as well (completeness for $1 \le p < \infty$ and the basis property for $1 < p < \infty$). Evidently for $p \ne 2$ it is not appropriate to talk about unconditional bases.

3 Wavelet Bases

In this section a method for constructing a family of wavelet bases is described. This approach allows us to remain within the framework of multiresolution analysis (see, for example, [5], [6]).

Let D and Ω be measurable bounded sets in $\mathbb{R}^n$, let $\{A_\gamma\}_{\gamma\in M}$ and $\{B_\omega\}_{\omega\in S}$ be two families of linear operators acting in $\mathbb{R}^n$, where M and S are some

countable sets of indices (in the examples below either $M = S = \mathbb{Z}^n$ or $M = S = \mathbb{Z}$), and let A_γ^* B_ω^* be operators conjugate to A_γ and B_ω, respectively.

Let ψ and ϕ be functions that depend on the sets D and Ω and whose Fourier transforms are of the form $\widehat{\psi}(t) = \chi_D(t), \widehat{\phi}(t) = \chi_\Omega(t)$, where χ_Ω and χ_D are the characteristic functions of the sets Ω and D.

Given that Λ and Σ are lattices in $\mathbb{R}^n$, we set

$$\psi_{\gamma,\alpha}(x)(2\pi)^{-n/2}(V(\Lambda)|A_\gamma|)^{1/2}\psi(A_\gamma x - \alpha)\,, \tag{4}$$

$$\phi_{\omega,\beta}(x) = (2\pi)^{-n/2}(V(\Sigma)|B_\omega|)^{1/2}\phi(B_\omega x - \beta) \tag{5}$$

(here $\gamma \in M$, $\omega \in S\,, \alpha \in \Lambda\,, \beta \in \Sigma$).

We choose an arbitrary $\mu \in \mathbf{M}$ and consider the space H^μ; for any $\gamma \in M$ and $\omega \in S$ we use the symbols W_γ and V_ω to denote the following closed linear spans: $W_\gamma = \operatorname{span}\{\psi_{\gamma,\alpha}\}_{\alpha\in\Lambda}$, $V_\omega = \operatorname{span}\{\phi_{\omega,\beta}\}_{\beta\in\Sigma}$.

Finally, we introduce the following functions, which depend on μ:

$$\psi^*_{\gamma,\alpha}(x)(2\pi)^{-n/2}(V(\Lambda)|A_\gamma|)^{1/2}\psi^*_\gamma(A_\gamma x - \alpha)\,,$$

$$\phi^*_{\omega,\beta}(x) = (2\pi)^{-n/2}(V(\Sigma)|B_\omega|)^{1/2}\phi^*_\omega(B_\omega x - \beta),$$

where the Fourier transforms of the functions ψ^*_γ and ϕ^*_ω are of the form

$$\widehat{\psi^*_\gamma}(t) = \mu^{-1}(A^*_\gamma t)\chi_D(t), \quad \widehat{\phi^*_\omega}(t) = \mu^{-1}(B^*_\omega t)\chi_\Omega(t)\,.$$

Theorem 1. *Let $\mu \in \mathbf{M}$. The systems of functions*

$$\{\psi_{\gamma,\alpha}\}_{\gamma\in M,\alpha\in\Lambda}$$

and

$$\left\{\psi^*_{\gamma,\alpha}\right\}_{\gamma\in M,\alpha\in\Lambda}$$

make up a biorthonormal pair of unconditional bases of the space H^μ if and only if the lattice Λ, the set D and the collection of linear operators $\{A_\gamma\}_{\gamma\in M}$ are such that each one of two systems of sets $\{D + 2\pi\alpha\}_{\alpha\in\Lambda^*}$ and $\left\{A^*_\gamma D\right\}_{\gamma\in M}$ are tilings of the space $\mathbb{R}^n$. Moreover, $H^\mu = \bigoplus_{\gamma\in M} W_\gamma$.

Theorem 2. *Let $\{D + 2\pi\alpha\}_{\alpha\in\Lambda^*}$ be a packing and for each $\omega \in S$ let M_ω be a subset of M. In order that for $\omega \in S$ the equality $V_\omega = \bigoplus_{\gamma\in M_\omega} W_\gamma$ be fulfilled, it is necessary and sufficient that the system of sets*

$$\{\Omega + 2\pi\beta\}_{\beta\in\Sigma^*} \tag{6}$$

be a packing and also that the system of sets $\left\{A^*_\gamma D\right\}_{\gamma\in M_\omega}$ be a tiling of the set $B^*_\omega\Omega$. Moreover he system of functions $\{\phi_{\omega,\beta}\}_{\beta\in\Sigma}$ is an unconditional basis of V_ω if and only if the system of sets (6) is a tiling of $\mathbb{R}^n$; here the systems $\{\phi_{\omega,\beta}\}_{\beta\in\Sigma}$ and $\left\{\phi^*_{\omega,\beta}\right\}_{\beta\in\Sigma}$ forms a biorthonormal (with respect to the scalar product in H^μ) pair of unconditional bases in V_ω.

Remark 2. In Theorems 1 and 2 purely geometric conditions are described which have to be fulfilled by the lattices Λ and Σ, the sets D and Ω, and the operators A_γ and B_ω in order that the systems of functions $\{\psi_{\gamma,\alpha}\}$, $\{\phi_{\omega,\beta}\}$ of the form (4) – (5) generated by these geometric objects possess specific properties in the spaces H^μ. We emphasize that these geometric conditions do not depend on the weight function $\mu \in \mathbf{M}$; in other words the same systems of functions of the form (4) – (5) possess the required properties simultaneously for the whole family of spaces H^μ where the weight function μ runs through the whole set $\mathbf{M}$ of moderately increasing functions.

4 Optimality of Wavelets

We now turn to the problem of optimality of wavelets constructed in the previous section.

Let $\mu, \nu \in \mathbf{M}$, $\omega \in S$ and for any $\varphi \in H^\mu$ let

$$H_\omega(\varphi) = \operatorname{span}\big\{\varphi(B_\omega \cdot - \alpha)\big\}_{\alpha\in\Sigma}$$

be a subspace of the space H^μ (in particular, $H_\omega(\phi) = V_\omega$). We define the following characteristics of approximation properties of these subspaces, which depend on the weight functions μ and ν:

$$\kappa_\omega(\varphi) = \sup_{\|u\|_\nu=1} \inf_{w\in H_\omega(\varphi)} \|u - w\|_\mu ,$$

$$\kappa_\omega = \inf_{\varphi\in H^\mu} \kappa_\omega(\varphi) . \tag{7}$$

Reformulating results of [7] one can show that $\kappa_\omega(\varphi)$ depends on the distributions of the values of the function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ of the form

$$f(x,\lambda) = \sum_{\beta\in 2\pi B_\omega^* \Sigma^*} \frac{\mu(x+\beta)|\widehat{\varphi}((B_\omega^*)^{-1}(x+\beta))|^2}{\mu(x+\beta)\nu^{-1}(x+\beta) - \lambda^2} ,$$

and κ_ω is determined by the behaviour of the ratio of the weight functions μ and ν.

More precisely, for any $c \in \mathbb{R}$ let Ω_c be the Lebesgue set of the function μ/ν of the form

$$\Omega_c = \left\{x \in \mathbb{R}^n : \sqrt{\mu(x)/\nu(x)} > c\right\} , \tag{8}$$

and $\overline{c} = \inf\{c \in \mathbb{R} : \operatorname{mes}(\Omega_c \cap (\Omega_c + 2\pi\alpha)) = 0 \quad \text{for all } \alpha \in B_\omega^*\Sigma^* \setminus \{0\}\}$ (that is $\overline{c}$ is the lower bound of the set $c \in \mathbb{R}$ such that the set Ω_c generates a $2\pi B_\omega^*\Sigma^*$–lattice packing). The following assertions are corollaries of the results of [7]:

1. $\kappa_\omega(\varphi) = \inf\{c > \overline{c} : f(\cdot, c) > 0 \quad \text{almost everywhere in } \Omega_c\}$ for all $\varphi \in H^\mu$;
2. $\kappa_\omega = \overline{c}$;

3. in order that a function $\varphi \in H^\mu$ be optimal (that is, the lower bound in (7) is attained for it; sometimes in this case we refer to the optimality of the subspace $H_\omega(\varphi)$ generated by φ) it is necessary and sufficient that each of the following two conditions holds almost everywhere in $\Omega_{\overline{c}}$:

$$\widehat{\varphi}((B_\omega^*)^{-1}\,\cdot\,) \neq 0\,; \quad f(\,\cdot\,,\overline{c}) \geq 0\,.$$

In particular, if the set $\Omega = \operatorname{supp}\widehat{\varphi}$ is such that (up to sets of measure zero) the relations

$$\Omega_{\overline{c}} \subseteq B_\omega^*\Omega \tag{9}$$

and

$$\{\Omega_{\overline{c}} + 2\pi B_\omega^*\beta\} \cap B_\omega^*\Omega = \emptyset \quad \text{for all } \beta \in \Sigma^* \setminus \{0\} \tag{10}$$

hold, then the function φ is optimal.

Theorem 3. *If conditions (9) – (10) hold, then the subspace V_ω constructed in section 3 is optimal (that is, at the function ϕ which generates this subspace the width κ_ω of the form (7) is realized).*

Remark 3. We note that in certain cases the same function φ turns out to be optimal for the hole collection of the spaces H^μ and H^ν defining widths of the form (7). This is the case, for example, when the weight functions μ and ν are such that $\mu(x)/\nu(x)$ depends only on $|x|$ and decreases monotonically as $|x|$ increases (a typical example is $H^\mu = W_2^s(\mathbb{R}^n),\ H^\nu = W_2^m(\mathbb{R}^n)$, where $s < m$). In this case the Lebesgue sets Ω_c of the form (8) are Euclidean spheres, the set $\Omega_{\overline{c}}$ is the ball of maximal radius generating a $2\pi B_\omega^*\Sigma^*$–lattice packing (it is essential that this ball does not depend on the weight functions) and, for example, any function φ whose Fourier transform does not vanish on the set $(B_\omega^*)^{-1}\Omega_{\overline{c}}$ and vanishes at the points of the set $\bigcup_{\beta\in\Sigma^*\setminus\{0\}}\{(B_\omega^*)^{-1}\Omega_{\overline{c}} + 2\pi\beta\}$ will be optimal.

5 Universally Optimal Wavelet Bases

We shall state an assertion (a direct corollary of Theorems 1, 2 and 3 for the case of Sobolev spaces) which provides a method for constructing universally optimal wavelet bases generating subspaces hat are extremal in the sense of projection–net widths of the unit ball of the space $W_2^m(\mathbb{R}^n)$ in the metric $W_2^s(\mathbb{R}^n)$ of the form

$$\inf_{\varphi\in W_2^s}\ \sup_{\|u\|_m=1}\ \inf_{w\in H_\omega(\varphi)} \|u-w\|_s \tag{11}$$

simultaneously for the whole scale of Sobolev classes (that is, for any pair $(W_2^m(\mathbb{R}^n), W_2^s(\mathbb{R}^n))$ for all $s < m$). In other words, the spaces V_ω do not have the saturation effect, namely, when approximating by means of elements of these spaces an automatic adjustment to the regularity of the approximated function takes place.

Theorem 4. *Let Λ and Σ be lattices in $\mathbb{R}^n$, let M and S be sets of indices of an arbitrary nature, and for any $\omega \in S$ let M_ω be a subset of M. Let the sets $D, \Omega \subset \mathbb{R}^n$ and collection of linear operators $\{A_\gamma\}_{\gamma\in M}$ and $\{B_\omega\}_{\omega\in S}$, acting in $\mathbb{R}^n$ be constructed such that the following conditions hold:*

1. *$\{D + 2\pi\alpha\}_{\alpha\in\Lambda^*}$, $\{\Omega + 2\pi\beta\}_{\beta\in\Sigma^*}$ $\{A^*_\gamma D\}_{\gamma\in M}$ are tilings of $\mathbb{R}^n$;*
2. *$\{A^*_\gamma D\}_{\gamma\in M_\omega}$ is a tiling of the set $B^*_\omega\Omega$ for any $\omega \in S$;*
3. *for any $\omega \in S$ the set $B^*_\omega\Omega$ possesses the following properties: a) it contains the open ball of maximal radius R_ω with the center at the origin, whose translations by the elements of the lattice $2\pi B^*_\omega\Sigma^*$ form a packing; b) it is disjoint from any other ball in this system of translations.*

If these geometric objects are constructed then the functions ψ and ϕ whose Fourier transforms are equal to the characteristic functions of the sets D and Ω, respectively, generate functions $\psi_{\gamma,\alpha}$ and $\phi_{\omega,\beta}$ of the form (4) and (5) such that for any $s \in \mathbb{R}$ the system of functions $\{\psi_{\gamma,\alpha}\}_{\gamma\in M,\alpha\in\Lambda}$ is an unconditional basis of the space $W_2^s(\mathbb{R}^n)$ (orthonormal for $s = 0$) and if V_ω and W_γ are subspaces of $W_2^s(\mathbb{R}^n)$ with bases $\{\phi_{\omega,\beta}\}_{\beta\in\Sigma}$ and $\{\psi_{\gamma,\alpha}\}_{\alpha\in\Lambda}$, then, firstly, for all $\omega \in S$ we have $V_\omega = \bigoplus_{\gamma\in M_\omega} W_\gamma$, and, secondly, for any $\omega \in S$ the subspace V_ω is optimal in the sense that the widths of the form (11) are realized on this subspace for all $s < m$.

Note that if $n = 1$, then the subspaces V_ω are extremal not only for projection–net but also for Kolmogorov widths of the same average dimension (see [7]).

Remark 4. In particular, condition (3) of Theorem 4 is satisfied if the set $B^*_\omega\Omega$ is the Voronoy polygon [4] of the lattice $2\pi B^*_\omega\Sigma^*$, that is,

$$B^*_\omega\Omega = \{x \in \mathbb{R}^n : |x| \le |x - \alpha| \text{ for all } \alpha \in 2\pi B^*_\omega\Sigma^*\} .$$

6 Examples

We provide two simple examples of constructions satisfying the conditions of Theorem 4.

Example 1. $$\Lambda = \Lambda^* = \Sigma = \Sigma^* = M = S = \mathbb{Z}^n ,$$

$$M_\omega = \{\gamma \in \mathbb{Z}^n : \quad \gamma_k < \omega_k \quad \forall k = 1, \dots, n\} \quad \forall\omega \in \mathbb{Z}^n,$$

$$\Omega = [-\pi, \pi]^n, \quad D = \{[-2\pi, -\pi] \cup [\pi, 2\pi]\}^n ,$$

$$A_\gamma x = A^*_\gamma x = B_\gamma x = B^*_\gamma x(2^{\gamma_1}x_1, \dots, 2^{\gamma_n}x_n) \in \mathbb{R}^n$$

$$\forall\gamma = (\gamma_1, \dots, \gamma_n) \in \mathbb{Z}^n, \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Then

$$B_\omega\Omega = \bigotimes_{k=1}^{n} [-2^{\omega_k}\pi, 2^{\omega_k}\pi] ,$$

$$A_\gamma D = \bigotimes_{k=1}^{n} \left\{ [-2^{\gamma_k+1}\pi, -2^{\gamma_k}\pi] \cup [2^{\gamma_k}\pi, 2^{\gamma_k+1}\pi] \right\} ,$$

$R_\omega = 2^a\pi$, where $a = \min_{1\le k\le n} \omega_k$, and conditions (1) — (3) of Theorem 4 are satisfied.

Example 2. $n = 2, \quad \Lambda = \Lambda^* = \Sigma = \Sigma^* = \mathbb{Z}^2, \quad M = S = \mathbb{Z}, \quad M_\omega = \{k \in \mathbb{Z} : k < \omega\} \quad \forall \omega \in \mathbb{Z}, \quad \Omega = [-\pi, \pi]^2$ is the Voronoy polygon of the lattice $2\pi\mathbb{Z}^2, \quad A_\gamma = B_\gamma = L^\gamma \quad \forall \gamma \in \mathbb{Z}, \quad D = L\Omega\backslash\Omega$, where L is an operator equal to the superposition of rotation by an angle $\pi/4$ in the positive direction and homothety with a coefficient $\sqrt{2}$; its matrix in the canonical basis has the form

$$L = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} .$$

The set D defined above is the so–called Brillouin zone of the second rank consisting of the points x of the plane $\mathbb{R}^2$ for which the zero node is the second furthest from the point x among all the nodes of the lattice $2\pi\mathbb{Z}^2$. In this case also all the required conditions are satisfied.

References

1. Hormander, L.: The analysis of linear partial differential operators. Vol.2. Springer–Verlag. Berlin (1983)
2. Volevich, L.P., Paneyakh, B.P.: Certain spaces of generalized functions and embedded theorems. Uspekhi Mat. Nauk **20** (1965) 3–74; English transl. in Russian Math. Surveys **20** (1965)
3. Lions, J.-L., Magenes, E.: Non-homogeneous boundary value problems and applications. Springer–Verlag. Berlin (1972)
4. Cassels, J.W.S.: An introduction to the geometry of numbers. Springer–Verlag. Berlin (1959)
5. Mallat, S.G.: Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. Trans. Amer. Math. Soc. **315** (1989) 69–87
6. Daubechies, I.: Ten lectures on wavelets. SIAM, Philadelphia, Pennsylvania (1992)
7. Strelkov, N.A.: Projection–lattice widths and lattice packings. Mat. Sb. **182** (1991) 1513–1533; English transl. in Math. USSR Sb. **74** (1993)

Dynamics of a Two-Level Medium Under the Action of Short Optical Pulses

Valerică Ninulescu and Andreea-Rodica Sterian

Department of Physics, "Politehnica" University of Bucharest,
Splaiul Independenţei 313, Bucharest, 060042, Romania
vninul@physics.pub.ro

Abstract. Optical Bloch equations are used to describe the dynamics of a two-level atomic medium subjected to ultrafast optical pulses. Starting with an uninverted system, the final atomic state is numerically studied. The generalization of the Bloch equations for a dense medium such that induced near dipole-dipole interactions are significant can result in a switching behavior of the atomic population inversion. Based on numerical investigations for typical pulse shapes, the mechanism of switching is explained in terms of dynamical systems theory.

1 Introduction

The interaction of a material medium with a light field is a basic subject in optics and its study can be achieved in a classical, semiclassical, or the full quantum theory [1, 2]. The quantum approach of the medium is usually an impossible task without approximations that make the problem a tractable one. The most simple model is the two-level system for the atomic constituents of the medium. This matter interaction with a classical light field can be described in many practical situations by the optical Bloch equations.

This work deals with the interaction of a short optical pulse with a collection of independent two-level atoms [2], or atoms coupled through dipole-dipole interactions [3, 4, 5]. We confine ourselves to the dissipation-free case and negligible propagational effects. Dynamical systems theory proves to be of much help in the study of the atomic response to the incoming light field. Numerical calculations make use of MATLAB.

For current experimental advances in the subject, see [5, 6, 7] and references therein.

2 Optical Bloch Equations

We consider a system of two-level atoms in interaction with an optical field. For the field, we assume a linear polarized electric field

$$E(t) = (1/2)[\mathcal{E}(t)\exp(-\mathrm{i}\omega t) + \mathrm{c.c}] \ , \tag{1}$$

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 635–642, 2005.

where $\mathcal{E}(t)$ and ω are the slowly-varying electric field intensity and the central angular frequency, respectively; for simplicity, the quantity $\mathcal{E}(t)$ takes on real values. In the dipole approximation and the rotating-wave approximation, the response of a two-level atom to the optical field is described by the optical Bloch equations [1, 2]

$$\dot{u} = -\Delta v \ , \tag{2}$$

$$\dot{v} = \Delta u + \Omega w \ , \tag{3}$$

$$\dot{w} = -\Omega v \ . \tag{4}$$

In the above, u and v are the slowly-varying in-phase and in-quadrature parts of the atomic dipole moment, respectively; w is the atomic population inversion. We use a frame rotating at the carrier frequency ω; in the absence of the external field, the atom is in the ground state and $(u, v, w) = (0, 0, -1)$. The parameter $\Delta = \omega_0 - \omega$ is the detuning of the atomic resonance frequency ω_0 from the field central frequency; the remaining parameter, Ω, is the instantaneous Rabi angular frequency [1, 2] which is proportional to $\mathcal{E}$. The vector (u, v, w), known as Bloch vector, completely characterizes the quantum state of the medium. For optical pulses whose duration is much less than the relaxation time of the induced dipoles, dissipation processes can be neglected and Eqs. (2)–(4) are appropriate to describe the medium.

In the derivation of Eqs. (2)–(4), it is assumed that atoms are driven by a local field which is equal to the incident one. We further consider a dense medium such that near dipole-dipole (NDD) interactions in the ground state lead to significant local field corrections. The generalization of the optical Bloch equations for this situation reads [3]

$$\dot{u} = -(\Delta + \varepsilon w)v \ , \tag{5}$$

$$\dot{v} = (\Delta + \varepsilon w)u + \Omega w \ , \tag{6}$$

$$\dot{w} = -\Omega v \ , \tag{7}$$

where ε ($\varepsilon/\omega_0 \ll 1$) is the strength of the NDD interactions with the dimension of angular frequency. To avoid propagational changes, a film of thickness much less than the field wavelength is considered. The additional terms introduce nonlinearities in the atomic variables which are sources of interesting phenomena such us the complete exciting of the atoms and optical switching [4] or optical bistability [5].

3 Atomic System Response to a Short Pulse

We start with the description of the atomic response through the optical Bloch Eqs. (2)–(4) in the on resonance case ($\Delta = 0$):

$$\dot{u} = 0, \quad \dot{v} = \Omega w, \quad \dot{w} = -\Omega v, \quad (u(-\infty), v(-\infty), w(-\infty)) = (0, 0, -1) \ . \tag{8}$$

The solution of this problem is [2]

$$u(t) = 0, \quad v(t) = -\sin\Theta(t), \quad w(t) = -\cos\Theta(t) \ , \tag{9}$$

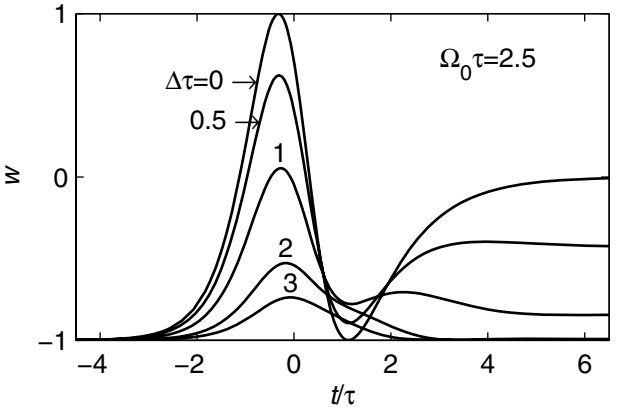


Fig. 1. Population inversion vs. time for the sech-pulse [Eq. (12)] of 2.5-pulse area and various detunings from ω_0 ($\Delta\tau \in \{0, 0.5, 1, 2, 3\}$)

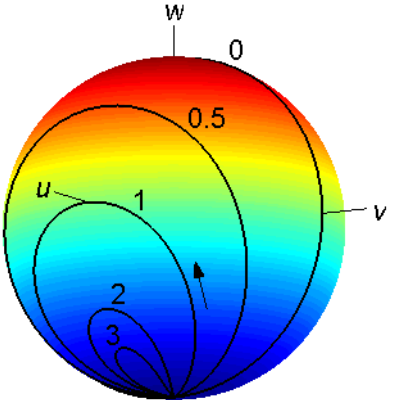


Fig. 2. Dynamics of the Bloch vector (u, v, w) on the unit sphere for the 2π-sech pulse for various detunings of the field central frequency from atomic resonance ($\Delta\tau \in \{0, 0.5, 1, 2, 3\}$). Starting from $(u, v, w) = (0, 0, -1)$ at $t \to -\infty$, the Bloch vector performes a loop in the direction indicated by the arrow; for $\Delta = 0$ the loop is the circle $v^2 + w^2 = 1$, $u = 0$

where $\Theta(t) = \int_{-\infty}^{t} \Omega(t')\mathrm{d}t'$. After the passage of the pulse, the atom is left in the state

$$u = 0, \quad v = -\sin\Theta_\mathrm{p}, \quad w = -\cos\Theta_\mathrm{p} \ , \tag{10}$$

where

$$\Theta_\mathrm{p} = \int_{-\infty}^{\infty} \Omega(t')\mathrm{d}t' \tag{11}$$

is referred to as the area of the pulse. According to Eq. (10), the final state depends on the pulse area and it is independent on the pulse shape. All values

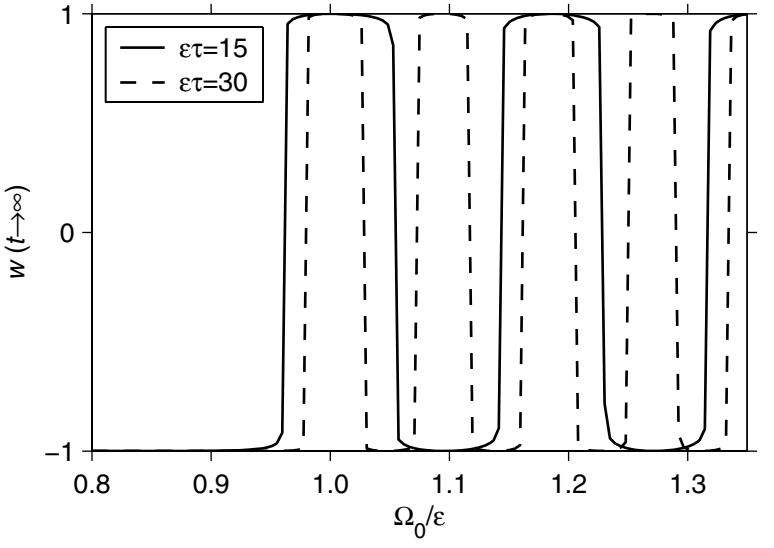


Fig. 3. Final state of population inversion vs. peak field strength of the sech-pulse [Eq. (12)] for two values of the parameter $\varepsilon\tau$

of the population inversion are accessible and this is smoothly reached. For example, a π-pulse excites the atom ($w = 1$) and a 2π-pulse leaves the atom in the ground state.

If the central frequency ω of the light pulse is detuned from the atomic frequency ω_0, for a given pulse the solution for the Bloch vector is generally Δ-dependent; this is shown in Fig. 1 for a sech-shaped pulse, i.e.,

$$\Omega(t) = \Omega_0 \mathrm{sech}(t/\tau) \ , \tag{12}$$

where Ω_0 is the peak value of the Rabi frequency and τ is a measure of the pulse width. More than that, the dynamics is dependent on the pulse shape.

To get an insight into the dynamics, let us notice that optical Bloch Eqs. (2)–(4) define a conservative dynamical system and the prime integral

$$u^2 + v^2 + w^2 = 1 \tag{13}$$

is easily derived for the initial condition $(u, v, w) = (0, 0, -1)$. Thus, in the uvw space the dynamics of the atomic system lays on the unit sphere. There exists an exceptional pulse of area 2π that leaves the atom in the ground state irrespective of the detuning, namely the sech-pulse (Fig. 2). This is the basis for the self-induced transparency phenomenon [1, 2].

We turn now to the investigation of the atomic dynamics in the presence of NDD interactions. To reveal the intrinsic features of these interactions influence on the atomic dynamics, we limit ourselves to the on resonance case ($\Delta = 0$); Eqs. (5)–(7) now become

$$\dot{u} = -\varepsilon v w \ , \tag{14}$$

$$\dot{v} = \varepsilon u w + \Omega w \ , \tag{15}$$

$$\dot{w} = -\Omega v \ . \tag{16}$$

In the absence of a closed-form analytical solution, a numerical approach is followed. We consider again a sech-pulse shape and a parameter $\varepsilon\tau \gg 1$. Figure 3 presents the final population inversion for a reduced Rabi frequency Ω_0/ε around unity. There exists a critical value of the parameter Ω_0/ε, $\Omega_{0,\mathrm{cr}}/\varepsilon < 1$, such that for $\Omega_0 < \Omega_{0,\mathrm{cr}}$ the final population inversion is $w = -1$, and for $\Omega_0 > \Omega_{0,\mathrm{cr}}$ the population inversion has an almost square wave shape, with the maximum at 1 and the minimum at -1. The greater the parameter $\varepsilon\tau$, the period of the oscillatory behavior decreases, the transitions are more abrupt, and the first cycle of the square wave becomes more centered with respect to $\Omega_0/\varepsilon = 1$. These characteristics preserve for other smooth pulse-shapes with one maximum, but not a square wave pulse for example.

4 Analysis of the Switching Effect

An analysis in terms of the theory of dynamical systems [8] follows. Equations (14)–(16) describe a conservative dynamical system. To start with, we determine the critical points; equating to zero the r.h.s. of these equations we find

$$u(t) = -\Omega(t)/\varepsilon, \quad v = 0, \quad w(t) = -\sqrt{1 - [\Omega(t)/\varepsilon]^2} \ , \tag{17}$$

$$u(t) = -\Omega(t)/\varepsilon, \quad v = 0, \quad w(t) = \sqrt{1 - [\Omega(t)/\varepsilon]^2} \ , \tag{18}$$

$$u = -1, \quad v = 0, \quad w = 0 \ , \tag{19}$$

$$u = 1, \quad v = 0, \quad w = 0 \ , \tag{20}$$

where the validity of eqs. (17,18) is limited to the case $\Omega(t)/\varepsilon \leq 1$.

The dynamical system starts from the point $(u, v, w) = (0, 0, -1)$, i.e., one end point of the quarter-circle (17). All our numerical simulations (Figs. 4–7) show that this curve is followed by the system. Physically, the real part u of the polarization field adiabatically follows the electric field and try to maintain 180° out of phase. Meanwhile, the imaginary part v of the polarization takes on small values or displays oscillations about the zero value. This happens as long as the ratio $\Omega(t)/\varepsilon$ is less than unity and not too close to this value; for $\Omega_0 < \Omega_{0,\mathrm{cr}}$, this is the case for $t \in (-\infty, \infty)$ (Fig. 4, $\Omega_0/\varepsilon = 0.95$).

For Ω_0/ε close to unity, at the peak of the pulse the dynamical system comes next to the fixed point (19) and the branch (18) of critical points. The switching to the curve (18) is performed if at the peak of the pulse $w > 0$ (Fig. 4, $\Omega_0/\varepsilon = 1$).

For $\Omega_0/\varepsilon > 1$, the curves given by eqs. (17, 18) dissapear when the instantaneous Rabi frequency $\Omega(t)$ satisfies $\Omega(t)/\varepsilon > 1$. This happens with the dynamical system close to the fixed point (19) that drives the system to an instantaneous limit cycle around it. After some time the pulse amplitude decreases enough so that critical points (17, 18) become real again; the system dynamics is influenced by that one which is nearest. If the number of cycles performed by the

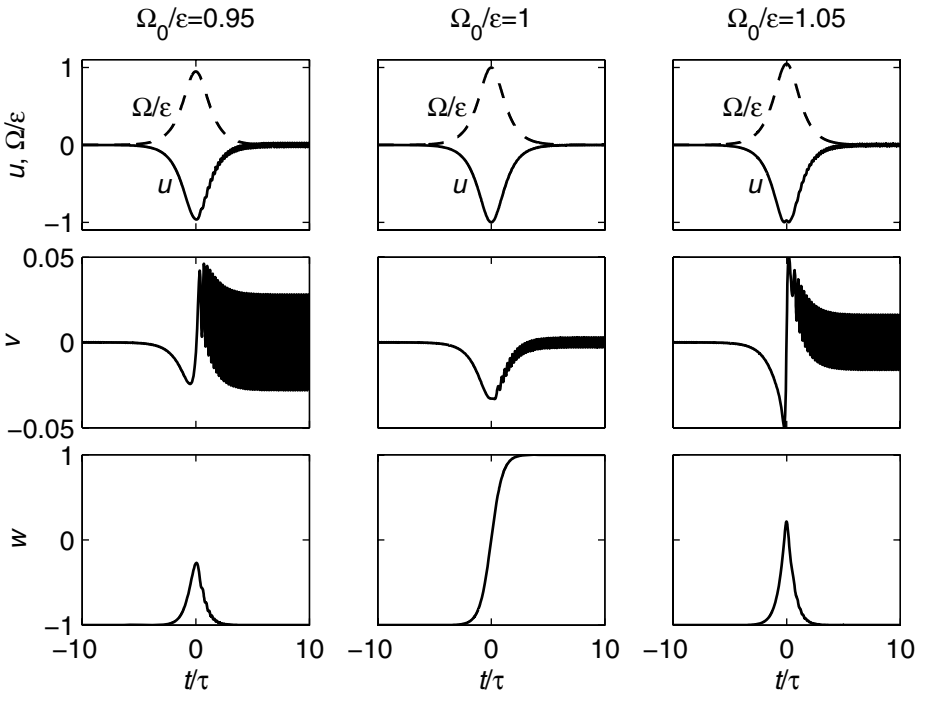
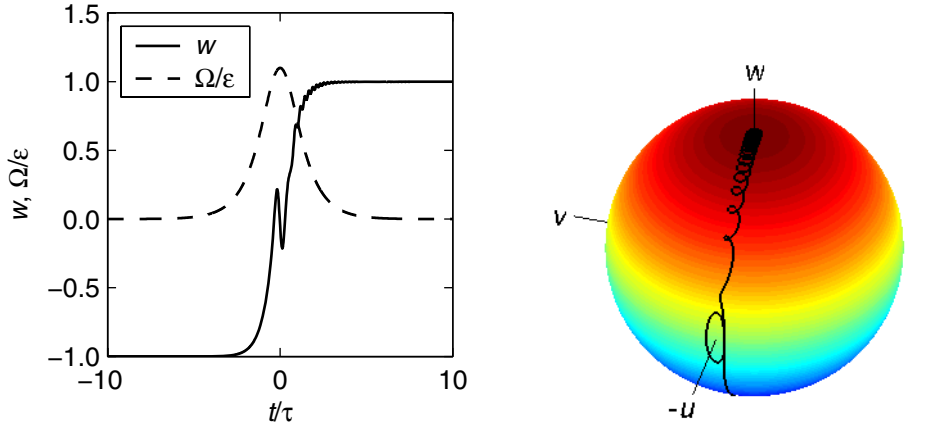
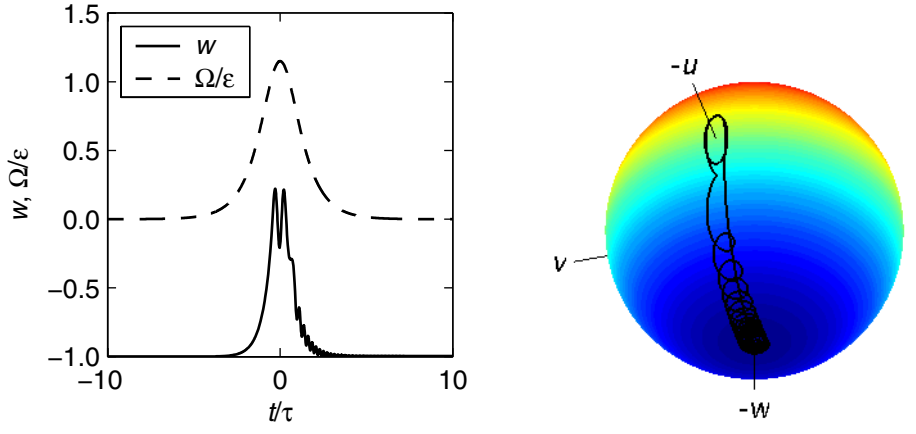


Fig. 4. Time evolution of the Bloch vector components at three values of the exciting sech-pulse [Eq. (12)] amplitude Ω_0/ε: before first switching (*left*), inside the first region of switching (*center*), and inside the first region of de-excitation (*right*)



(a) (b)

Fig. 5. (a) Time evolution of the population inversion (*continuous line*) induced by a sech-pulse (*dashed line*) with $\Omega_0/\varepsilon = 1.1$ and $\varepsilon\tau = 30$. (b) Dynamics of the Bloch vector (u, v, w): starting from $(u, v, w) = (0, 0, -1)$ at $t \to -\infty$, the Bloch vector approaches the point $(u, v, w) = (0, 0, 1)$ for $t \to \infty$



(a) (b)

Fig. 6. (a) Time evolution of the population inversion (*continuous line*) induced by a sech-pulse (*dashed line*) with $\Omega_0/\varepsilon = 1.15$ and $\varepsilon\tau = 30$. (b) Dynamics of the Bloch vector (u, v, w): starting from $(u, v, w) = (0, 0, -1)$ at $t \to -\infty$, the Bloch vector returns to this value for $t \to \infty$

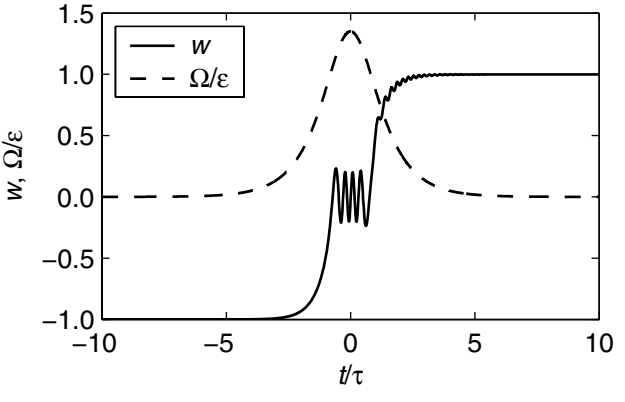


Fig. 7. Time evolution of the population inversion (*continuous line*) induced by a sech-pulse (*dashed line*) with $\Omega_0/\varepsilon = 1.35$ and $\varepsilon\tau = 30$

point (u, v, w) while $\Omega(t)/\varepsilon > 1$ is an integer, the system dynamics is influenced by critical points (17) and the atomic ground state is reached for $t \to \infty$; this is the case in Fig. 4 ($\Omega_0/\varepsilon = 1.05$) and Fig. 6. If on the contrary the number of cycles is a half-integer one (Figs. 5, 7), the system evolves to the excited state for $t \to \infty$. The phenomenon of switching in Fig. 3 for increasing the Rabi frequency corresponds to an increase of one half-cycle for the system dynamics in the uvw phase space.

5 Conclusion

The above results present an unifying point of view for the interaction of a two-level atomic system with a light pulse. The dynamics of the atomic system is characterized in terms of the Bloch vector motion on the unit sphere and it is explained based on the critical points of the dynamical system. Even if our results are explicitly derived for a sech-pulse, their validity also holds for other smooth pulse shapes with one maximum.

References

1. Meystre, P., Sargent M. III: Elements of Quantum Optics. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1999)
2. Mandel, L, Wolf, E.: Optical Coherence and Quantum Optics. Cambridge University Press, Cambridge New York Melbourne (1995)
3. Bowden, M., Dowling, J.P.: Near-Dipole-Dipole Effects in Dense Media: Generalized Maxwell-Bloch Equations. Phys. Rev. A **47** (1993) 1247–1251
4. Crenshaw, M.E., Scalora, M., Bowden, C. M.: Ultrafast Intrinsic Optical Switching in a Dense Medium of Two-Level Atoms. Phys. Rev. Lett. **68** (1992) 911–914
5. Hehlen, M.P., Güdel, H.U., Shu, Q., Rand, S.C.: Cooperative Optical Bistability in a Dimer System $Cs_3Y_2Br_9$:10 % Yb^{3+}. J. Chem. Phys. **104** (1996) 1232–1244
6. Xu, X.Y., Dai, D.C., Cai, Z.G., Luo, Q., Ninulescu, V., Zhou, J.Y.: Coherent Population Density Control by Ultra-high Repetition Rate Femtosecond Optical Pulse Pairs for Optical Switching. Opt. Commun. **164** (1999) 177–184
7. Redmond, S. M., Rand, S.C.: Intrinsic Chromatic Switching of Visible Luminescence in Yb^{3+}, Er^{3+}:$CsCdBr_3$. Opt. Lett. **28** (2003) 173–175
8. Broer, H.W., Dumortier, F., van Strien, S.J., Takens, F.: Structures in Dynamics. Finite Dimensional Deterministic Studies. North-Holland Elsevier Science Publishers B.V., Amsterdam (1991)

Nonlinear Phenomena in Erbium-Doped Lasers

Andreea Sterian and Valerică Ninulescu

Department of Physics, "Politehnica" University of Bucharest,
Splaiul Independenţei 313, Bucharest, 060042, Romania
vninul@physics.pub.ro

Abstract. The nonlinear dynamics of an erbium-doped fiber laser is explained based on a simple model of the ion pairs present in heavily doped fibers. The single-mode laser dynamics is reductible to four coupled nonlinear differential equations. Depending on the ion-pair concentration, the pumping level and the photon lifetime in the laser cavity, numerical calculations predict cw, self-pulsing and sinusoidal dynamics. The regions of these dynamics in the space of the laser parameters are determined.

1 Introduction

In recent years much attention has been paid to the study of nonlinear effects in optical fibers [1, 2, 3, 4, 5, 6]. Erbium-doped fiber lasers (EDFLs) present numerous applications in telecommunications and laser engineering. In the same time, it is also a promising system for theoretical nonlinear studies [7, 8]. Self-pulsing and chaotic operation of the EDFLs has been reported in various experimental conditions [3, 4], including the case of pumping near the laser threshold. We present a model for the single-mode laser taking into account the presence of the erbium ion pairs that act as a saturable absorber.

2 The Basic Model

The erbium-doped laser emitting around 1.55 μm ($^4I_{13/2} \rightarrow {}^4I_{15/2}$ transition) is a three-level system (Fig. 1). Er^{3+}-ions at a concentration N_0 are pumped with the rate Λ from level 1 ($^4I_{15/2}$) to level 3 (for example $^4I_{11/2}$, 980 nm above the ground state). The level 3 is fastly depopulated through a non-radiative transition on the upper laser level 2 ($^4I_{13/2}$) which is metastable with the lifetime $\tau_2 = 1/\gamma_2 = 10$ ms. The letter σ in Fig. 1 denotes the absorption cross section in the laser transition. In the rate equation approximation, the laser dynamics is described based on two coupled differential equations, one for the population inversion and the other for the laser intensity:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = 2\Lambda - \gamma_2(1+n) - 2in \ , \tag{1}$$

$$\frac{\mathrm{d}i}{\mathrm{d}t} = -i + Ain \ . \tag{2}$$

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 643–650, 2005.

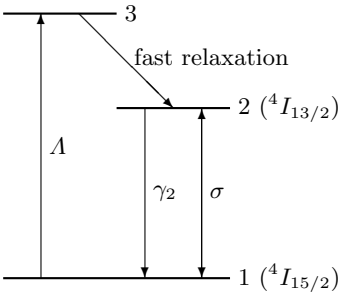


Fig. 1. Erbium-ion energy levels implied in the laser effect

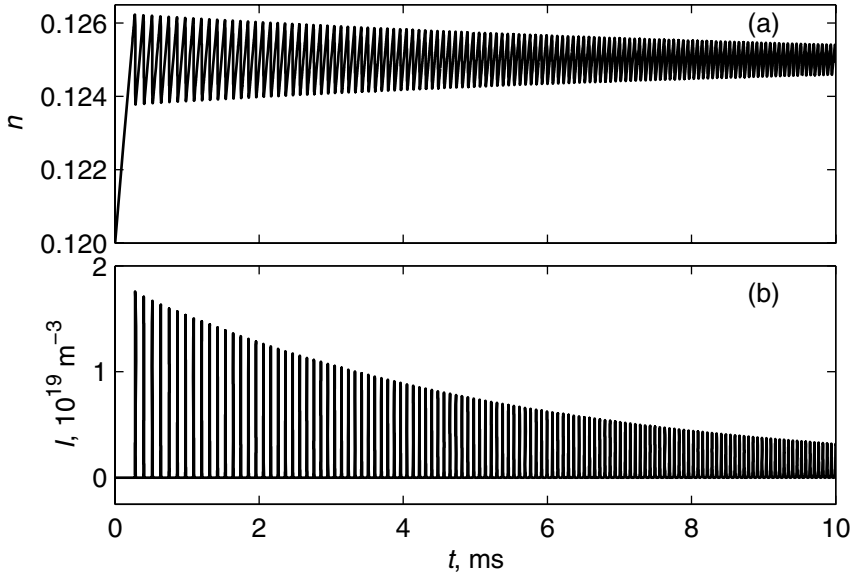


Fig. 2. Transient dynamics of the erbium laser. (a) Population inversion; (b) Laser intensity. Numerical values of the laser are given in text and photon lifetime is $\tau = 10^{-8}$ s. The plot starts at a point where the population inversion reaches a value close to the stationary one

Here, the time is expressed in units of the photon lifetime τ in the laser cavity, $n = n_2 - n_1$ is the difference of the occupation probability of level 1 and 2, respectively, and i stands for the dimensionless laser intensity; $i = \sigma I \tau$, where I denotes the photon density of the laser field. The parameter A in Eq. (2) is $A = \sigma N_0 \tau$; note that parameter γ_2 in Eq. (1) is dimensionless, i.e., the relaxation rate is expressed in $1/\tau$ units.

The determination of the steady-states furnishes the points

$$n = 2\Lambda/\gamma_2 - 1,\ i = 0 \quad \text{and} \quad n = 1/A,\ i = A(\Lambda - \Lambda_{\text{th}})\ , \tag{3}$$

where

$$\Lambda_{\rm th} = \frac{\gamma_2}{2}\left(1+\frac{1}{A}\right) . \tag{4}$$

Linear stability investigation gives laser action for $\Lambda > \Lambda_{\rm th}$ and the laser solution is stable for all pumping levels above threshold. Concerning the type of stability, we notice that the eigenvalues λ of the linearized system around the laser solution,

$$\lambda^2 + 2A\left(\Lambda - \frac{A}{A+1}\Lambda_{\rm th}\right)\lambda + 2A(\Lambda - \Lambda_{\rm th}) = 0 \ , \tag{5}$$

can be both negative (node) or complex-conjugate with negative real part (focus). We take $N_0 = 5\times 10^{24}\,{\rm m}^{-3}$, $\sigma = 1.6\times 10^{-16}\,{\rm m}^3{\rm s}^{-1}$ and $\tau = (2 \div 200)\,{\rm ns}$. In terms of the pumping strength $r = \Lambda/\Lambda_{\rm th}$, the changes in the stability type occur for $r_- = 1 + A\Lambda_{\rm th}/2(A+1)^2 \approx 1$ and $r_+ = 1 + 2/A\Lambda_{\rm th} \gg 1$. Practically, for all accessible pumping levels and all photon lifetimes, the eigenvalues are complex conjugate; consequently, the approaching to the stable laser solution is performed through relaxation oscillations (Fig. 2).

3 Ion-Pair Influence

At sufficiently large ion concentrations, their interaction gives rise to clusters with a distinct contribution to the laser effect. One model considers the amplifying medium as a mixture of isolated erbium ions and erbium ion pairs [3]. The strength of the interaction in an ion-pair is relatively small due to the screening effect of the $4d^{10}$ electrons on the $4f$ electrons. So, the ion energy levels are practically preserved and the energy in an ion-pair is the sum of the two ions energy. The ion-pair importance is related to the quasiresonance of the transition ${}^4I_{9/2} \rightarrow {}^4I_{13/2}$ with the laser transition. This makes it probable the up-conversion process in an ion-pair (Fig. 3). For an ion-pair the laser levels are 11 (ground level), 12 (intermediate level) and 22 (uppermost level), separated through the laser transition. Denoting by n_{11}, n_{12} and n_{22} the corresponding populations, these quantities satisfy the normalization condition $n_{11} + n_{12} + n_{22} = 1$.

Based on the above picture of the active medium, two supplementary equations are to be added to Eqs. (1,2). By use of the new variables $n_\pm = n_{22} \pm n_{11}$, the complete set of laser equations is

$$\frac{{\rm d}n}{{\rm d}t} = 2\Lambda - \gamma_2(1+n) - 2in \ , \tag{6}$$

$$\frac{{\rm d}n_+}{{\rm d}t} = \gamma_2(1-n_+) - \frac{\gamma_{22}}{2}(n_+ + n_-) + yi(2-3n_+) \ , \tag{7}$$

$$\frac{{\rm d}n_-}{{\rm d}t} = 2\Lambda - \gamma_2(1-n_+) - \frac{\gamma_{22}}{2}(n_+ + n_-) - yin_- \ , \tag{8}$$

$$\frac{{\rm d}i}{{\rm d}t} = -i + (1-2x)Ain + xAyin_- \ . \tag{9}$$

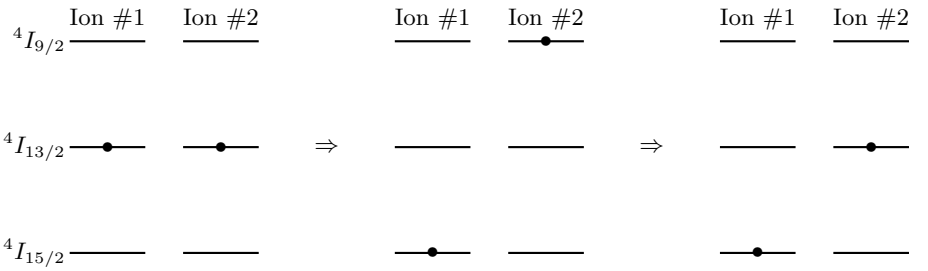


Fig. 3. Up-conversion (*center*) followed by relaxation (*right*) in an ion-pair

The quantity x is the fraction of ion pairs in the active medium, γ_{22} is the (reduced) relaxation rate of levels 22 and 12, and $y\sigma$ $(0 < y < 1)$ is the absorption cross section for the laser transition in an ion-pair.

4 Steady-States

The determination of the steady-states of system (6)–(9) reduces to the solutioning of a third-order polynomial equation for the laser intensity; this is

$$i^3 + c_1 i^2 + c_2 i + c_3 = 0 \ , \tag{10}$$

where

$$c_1 = Ax(\gamma_{22} - 2\gamma_2)/3 + (\gamma_2 + 2\gamma_{22})/3y + \gamma_2/2 - A(2\Lambda - \gamma_2)/2 \ , \tag{11}$$

$$c_2 = x[A\gamma_2(\gamma_2/6 + \gamma_{22}/6 - \Lambda) + A\gamma_{22}\Lambda/y - A\gamma_2(\gamma_2 + \gamma_{22})/3y] + \gamma_2(\gamma_2 + 2\gamma_{22})/6y + \gamma_2\gamma_{22}/3y^2 - A(\gamma_2 + 2\gamma_{22})(2\Lambda - \gamma_2)/6y \ , \tag{12}$$

$$c_3 = Ax\gamma_2(\gamma_2\gamma_{22} - 2\gamma_2\Lambda - \gamma_{22}\Lambda)/6y + \gamma_2^2\gamma_{22}/6y^2 - A(1 - 2x)\gamma_2\gamma_{22}(2\Lambda - \gamma_2)/6y^2 \ . \tag{13}$$

The laser threshold is obtained from the condition $c_3 = 0$ that yields

$$\Lambda_{\rm th} = \frac{1 - Ax(2 - y)/(A + 1)}{1 - (2 - y/2 - y\gamma_2/\gamma_{22})x}\Lambda_{\rm th}^0 \ , \tag{14}$$

where $\Lambda_{\rm th}^0$ is the pump threshold in the absence of the ion pairs and it is given by Eq. (4). Besides previously given laser parameters, we take the lifetime $\tau_{22} = 2\,\mu$s of the ion-pair level 22 and $y = 0.2$. The dependence of the threshold pumping level on the concentration of the ion-pair and the photon lifetime is presented in Fig. 4. The increase of the laser threshold due to the presence of the ion pairs is an important drawback of heavily doped fibers.

Above threshold, there exists only one steady-state laser intensity given by Eq. (10), the other two being unphysical (negative). The numerical calculation of the steady-state intensity for a significant range of the pumping parameter gives a laser intensity following a straight line dependency (Fig. 5) as in the absence of the ion pairs [see Eq. (3)].

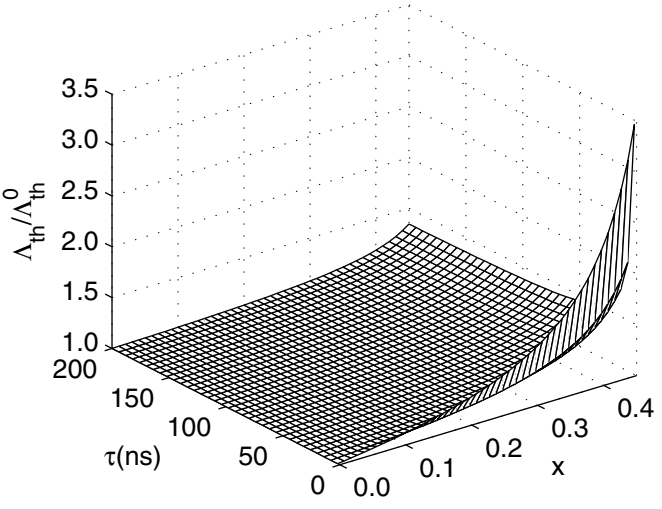


Fig. 4. Threshold pumping parameter vs. ion-pair percentage and photon lifetime. The laser parameters are given in text

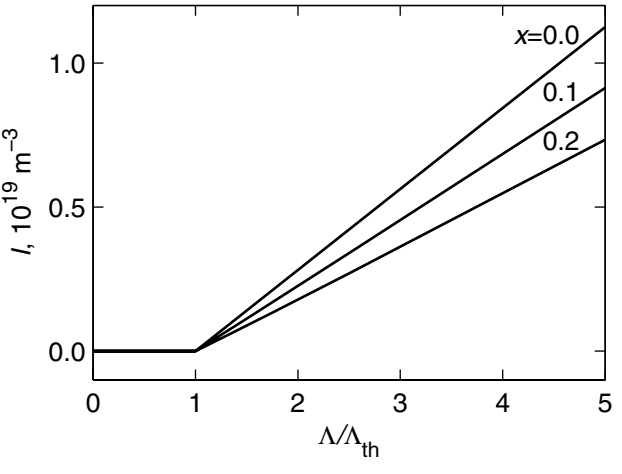


Fig. 5. Steady-state photon density of the laser field vs. the pumping strength $\Lambda/\Lambda_{\rm th}$. The same parameters as in Fig. 4 and $\tau = 10^{-8}$ s

5 Laser Dynamics

Linear stability investigation of the steady-states [7, 8] reveals the existence of a critical value of the ion-pair percentage under which the steady-state (cw) solution is stable whatever the pumping level. At larger concentrations of the ion pairs, the laser is in a steady-state or self-pulsing, depending on the value

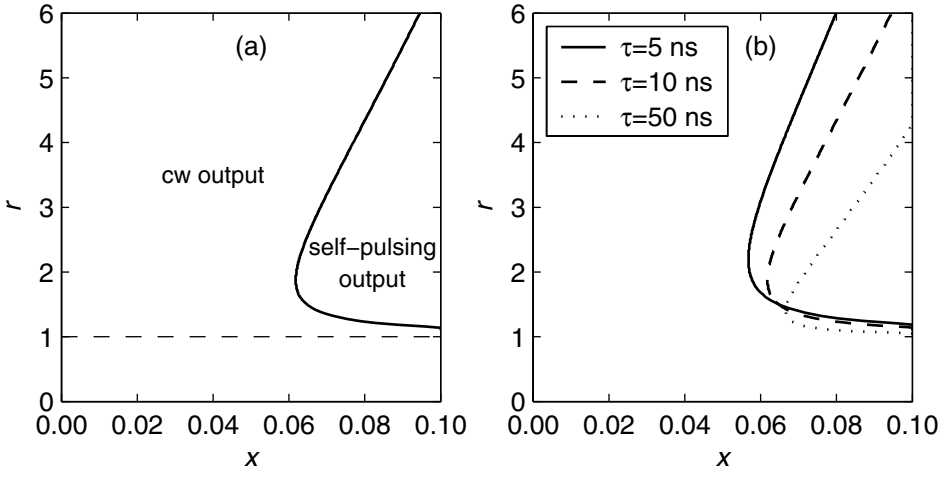


Fig. 6. (a) Calculated stability diagram for the EDFL. $\tau = 10\,\text{ns}$; (b) The influence of the photon lifetime on the margins of the stability domains

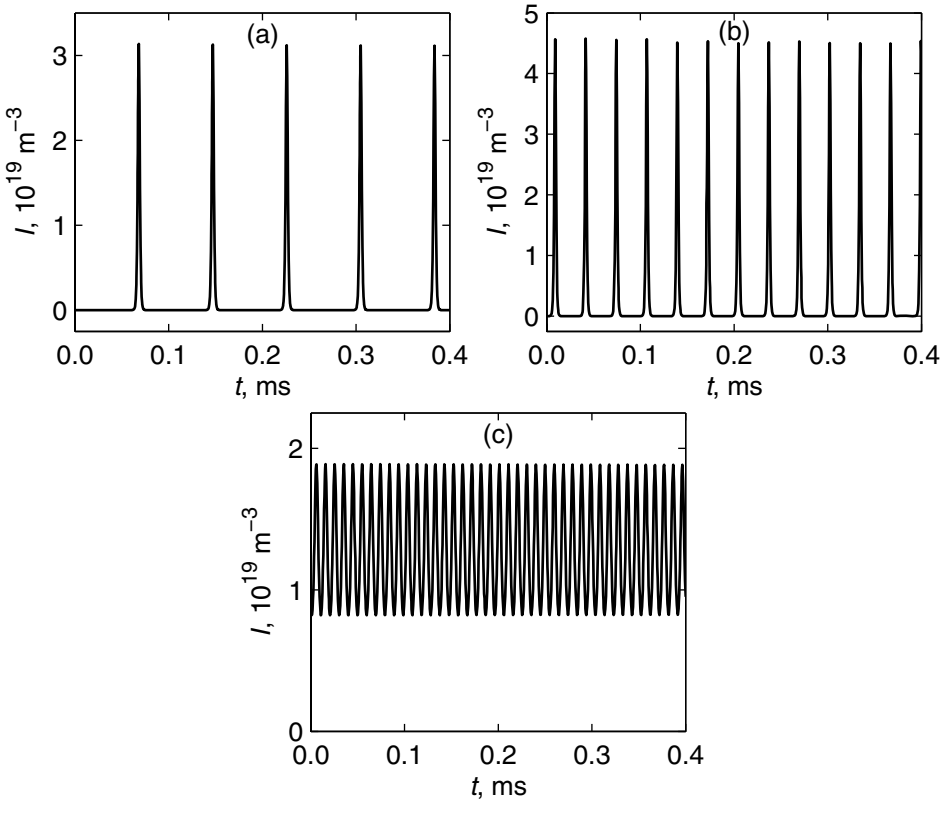


Fig. 7. Asymptotic temporal evolution of the photon density for a fraction $x = 0.1$ of ion pairs and the pumping level (a) $r = 1.5$, (b) $r = 2.5$, and (c) $r = 6.6$. $\tau = 10^{-8}\,\text{s}$

of the pumping parameter $r = \Lambda/\Lambda_{\text{th}}$ [Fig. 6(a)]. The transition from the cw dynamics to a self-pulsing one takes place when two complex conjugate eigenvalues of the linearized system cross the imaginary axis from left to right, i.e., a Hopf bifurcation occurs [7]; in such conditions the steady-state solution becomes unstable and the long term system evolution settles down on a stable limit cycle in the phase space. Figure 6(a) shows that the low branch of the Hopf bifurcation corresponds to pumping strengths close to the threshold value.

The quantitative changes of the laser intensity inside the self-pulsing domain are clarified in Fig. 7. At a fixed value of the ion-pair concentration, the increase in pumping gives rise to pulses of a higher repetition rate and close to the bifurcation point the intensity becomes sinusoidal. Besides, pulse amplitude reaches a maximum approximately in the middle of the self-pulsing domain.

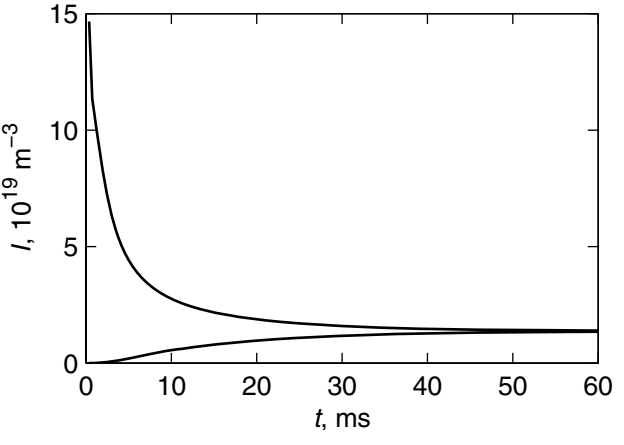


Fig. 8. Envelope of the transient photon density for a pumping parameter $r = 7$. The same laser parameters as in Fig. 7

Further increase of the pumping strength gives again a cw dynamics (Fig. 8).

The photon lifetime contribution to the stability diagram is emphasized in Fig. 6(b). This proves that cavities with low losses makes it possible to preserve the cw dynamics even at larger doping levels.

6 Conclusion

The erbium ion pairs in a fiber laser can explain the experimentally observed nonlinear dynamics of the system [2, 3]. In the single-mode laser description, the ion pairs are responsible for an increase of the laser threshold and a decrease of the laser power, and self-pulsations. Depending on the experimental conditions, the laser self-pulsations range from an oscillatory form to a well defined pulse-shape. The nonlinear dynamics can be limited by a choice of sufficiently low doped fibers at the expense of using longer fibers.

References

1. Agrawal, G. P.: Nonlinear Fiber Optics. 2nd edn. Academic Press, San Diego (1995)
2. LeFlohic, M., François, P.L., Allain, J.Y., Sanchez, F., Stephan, G.: Dynamics of the Transient Buildup of Emission in Nd^{3+} Doped Fiber Lasers. IEEE J. Quantum Electron. **27** (1991) 1910–1921
3. Sanchez, F., LeBoudec, P., François, P.L., Stephan, G.: Effects of Ion Pairs on the Dynamics of Erbium-Doped Fiber Lasers. Phys. Rev. A **48** (1993) 2220–2229
4. Sanchez, F., LeFlohic, M., Stephan, G.M., LeBoudec, P., François, P.L.: Quasi-Periodic Route to Chaos in Erbium-Doped Fiber Laser. IEEE J. Quantum Electron. **31** (1995) 481–488
5. Williams, Q.L., García-Ojalvo, J., Roy, R.: Fast Intracavity Polarization Dynamics of an Erbium-Doped Fiber Ring Laser: Inclusion of Stochastic Effects. Phys. Rev. A **55** (1997) 2376–2386
6. DeShazer, D.J., García-Ojalvo, J., Roy, R.: Bursting Dynamics of a Fiber Laser with an Injected Signal. Phys. Rev. E **67** (2003) 036602
7. Crawford, J.D.: Introduction to Bifurcation Theory. Rev. Mod. Phys. **63** (1991) 991–1037
8. Abarbanel, H.D.I., Brown, R., Sidorowich, J.J., Tsimring, L.S.: The Analysis of Observed Chaotic Data in Physical Systems. Rev. Mod. Phys. **65** (1993) 1331–1392

An e-Lottery Scheme Using Verifiable Random Function*

Sherman S.M. Chow, Lucas C.K. Hui, S.M. Yiu, and K.P. Chow

Department of Computer Science,
University of Hong Kong, Pokfulam, Hong Kong
{smchow, hui, smyiu, chow}@cs.hku.hk

Abstract. A number of e-lottery schemes have been proposed; however, none of them can satisfy all the identified requirements. In particular, some of them require a certain subset of players to remain online or the existence of a trusted third party (TTP) in order to generate the winning number(s) and some suffer from the forgery ticket attack. In this paper, we propose a new e-lottery scheme based on *Verifiable Random Function* that can satisfy all the identified requirements without the presence of TTP, yet the result of this generation is publicly verifiable.

Keywords: e-lottery, verifiable random function, applied cryptography.

1 Introduction

Lottery is a multi-billion dollar industry. Apart from gambling, lottery exists in other forms such as fund-raising coupon for charity. In a typical lottery, there is one dealer and a large number of players. Depending on the rules of the game, players bet on a single number or a combination of numbers chosen from a pre-defined domain. A random process (e.g. drawing of lots) is used to determine the combination of winning numbers. Due to the randomness of the process, it is not repeatable; therefore, the process is usually executed or monitored by a trusted auditing organization. With the popularity of the Internet, it is natural to ask whether we can have a secure e-lottery scheme, that is, having ticket purchase, winning result generation and prize claiming all done over the Internet. While there are many different interpretations of "security" issues in an e-lottery, we believe the following criteria are common to most practical e-lottery schemes.

1. *Random generation of the winning result.*
   Each possible combination of numbers from the domain is equally likely to be the winning result. No player can predict the result better than guessing.

* This research is supported in part by the Areas of Excellence Scheme established under the University Grants Committee of the Hong Kong Special Administrative Region (HKSAR), China (Project No. AoE/E-01/99), grants from the Research Grants Council of the HKSAR, China (Project No. HKU/7144/03E, HKU/7136/04E and HKU/7021/00E), and grants from the Innovation and Technology Commission of the HKSAR, China (Project No. ITS/170/01 and UIM/145).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 651–660, 2005.

2. *The winning result is publicly verifiable.*
   Every player can verify the winning result, i.e. the dealer cannot cheat in the generation process.
3. *The total revenue and the number of winning tickets are publicly verifiable.*
   The amount of winning prizes is usually related to dealer's revenue and hence the number of sold tickets should be publicly verifiable. In some lottery games, the prize for each winning player is not fixed but depends on the total number of winning tickets; hence, the number of winning tickets should also be publicly verifiable.
4. *Forgery of winning ticket is impossible.*
   Either players or the dealer cannot forge a ticket. Players cannot create a winning ticket which they did not purchase. The dealer cannot forge a ticket which is the same as a winning ticket after the winning result is generated.
5. *User is not required to be online for the generation of the winning result.*
   In some previous schemes, the generation of the winning result requires some players to remain online. This is not realistic, especially in a large scale game.
6. *Anonymity of player.*
   Player's anonymity is of paramount importance especially for the winning player(s). If a winning player's anonymity is compromised, he/she may face the threat of blackmail, beggars, etc.
7. *Confidentiality of the ticket's value.*
   In some lottery games, the players bet on a single number or a combination of numbers chosen from a pre-defined domain. In the case more than one winning players bet on the same value, the prize will be shared equally among them. In this setting, we need to keep the confidentiality of the ticket's value, or various attacks are possible. For examples, cutting down a certain ticket value's winning share by intentionally "duplicating" other's ticket value, or eavesdropping the channel between the dealer and the clients to learn the "popularity" of the values and betting on the most "profitable" number.
8. *Fairness on purchase of tickets and claiming of prize.*
   Purchase of tickets and claiming of prizes are based on the principle of fair exchange. That is, either both parties (the dealer and the player in our case) get the other party's item, or no party gets the other party's item at the end of a transaction.
9. *No early registration of player is required.*
   Some schemes (e.g. [6]) require early registration of players to avoid collusion of players. This is an unrealistic requirement since early registration is not required in traditional lottery.

Previous Work: There are many existing e-lottery schemes (e.g. [4, 6, 7, 8, 9, 10, 11, 14, 15]) but only three of them ([6, 11, 15]) address similar issues as our scheme without resort to a trusted third party (TTP)[1]. Some schemes assume a model different from ours: [7] considers the lottery game as the form of a *real-time* interactive game between players and dealer in the casino, while the

[1] TTP may be required for fair exchange.

security of [4] is guaranteed by using *multiple* dealers. Some schemes [8, 9, 14] simply assume the existence of a *trusted random source*, for example, from the random number generation process executed by a TTP. Moreover, [14] does not address the anonymity requirement, the forgery attack and the fairness on the purchase of tickets and claiming of prize; while [9] is shown to be insecure by [8].

The most recent e-lottery proposal is the "electronic national lotteries" [10]; however, a TTP (called "verifier" in their terms) is assumed to take into accounts of all players' winner tickets in the winning result generation (so that the dealer cannot predict the winning result until the last player's decision is made) and will not "insert" any forged tickets to affect the winning result generation. Due to the above mentioned reasons, we mainly compare our scheme with [6, 11, 15]. Table 1 shows a comparison between the existing schemes and our new scheme. None of these three existing e-lottery schemes can satisfy all the requirements that we have identified.

In [6], only the tickets sold in the "critical purchase phase" are used to determine the winning result. This makes forgery of winning ticket possible: After the winning result is generated, the dealer can "insert" the winning ticket to the tickets sold before the critical purchase phase, without altering the winning result. The dealer can do so since the sequence numbers and the timestamps associated with the tickets are not guaranteed to be sequential. For example, when multiple purchase requests are made simultaneously but one of the buyers refuses to pay for the ticket afterwards. Besides, the dealer may simply replace the ticket value of a certain ticket which is under his control and "sold" before the critical purchase phase. These malicious acts may not be noticed easily except there is a player who checks each ticket one by one with previously downloaded tickets list from time to time, which is impractical since the number of tickets sold is potentially in millions. As forgery of winning ticket is possible, the total number of winning tickets can be altered easily.

The first e-lottery scheme that uses delaying function (see the definition in Section 2) to prevent forgery ticket attack is [6]. In fact, our scheme also uses a similar idea to prevent the forgery ticket attack. However, their scheme applies the function only on a portion of the tickets sold while our scheme applies the function to all tickets sold so as to achieve a higher level of security.

Instead of applying delaying function in the winning result generation phase, [11] applies delaying function in the verification phase. Their scheme tries to prevent the forgery ticket attack by imposing a time limit for the prize claiming phase. Players must do the verification through a delaying function immediately afterwards, or they will suffer from losing the game due to the insufficient time for verification. This requirement is not realistic. Based on their scheme, the exact number of winning tickets cannot be accurately calculated unless all the winning tickets are *claimed* accordingly. In addition, this scheme has not addressed the anonymity and fair exchange issues.

The first piece of work integrating fair exchange with e-lottery is [15], this scheme satisfies most of the requirements we have identified. However, their winning result generation requires some players to be online, which make the

Table 1. A Comparison of e-Lottery Schemes

Scheme	Requirement for an e-Lottery Scheme									
	1	2	3	4	5	6	7	8	9	Trusted Third Party *Not* Required?
Verifiable Lotteries [6]	Y	Y	*N*	*N*	Y	*N*	Y[2]	*N*	*N*	Y
Digital Lottery Server [14]	Y	Y	*N*[3]	Y	Y	*N*	*N*	*N*	Y	*N*
Fair e-Lotteries [11]	Y	Y	*N*	Y	*N*	*N*	Y[2]	*N*	Y	Y
Lottery on the Internet [15]	Y	Y	Y	Y	*N*	Y	*N*[3]	Y	Y	Y[1]
Our scheme	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y[1]

scheme not realistic and not robust. For example, the winning result cannot be generated if none of the players participates in the generation process. Also, the winning result may be biased if all participating players collude, which is an important issue for small scale e-lottery game or when the e-lottery game is not yet popular in the beginning. Moreover, they have not considered the security issue of not encrypting the ticket's value.

Our Contribution: In this paper, we propose a new e-lottery scheme which satisfies all of the identified requirements by using the *verifiable random function*. Our scheme makes sure that the lottery dealer generates the winning number(s) based on all the tickets sold. The calculation of the winning number(s) does not depend on the participation of players but it is yet verifiable.

2 Preliminaries

We first discuss the cryptographic primitives used in our scheme.

Verifiable Random Function: *Verifiable Random Function* (hereafter referred as VRF) was introduced by [13]. Basically it is a pseudorandom function [5] providing a non-interactively verifiable proof for the output's correctness.

Based on the notation in [12], a function family $F_{(\cdot)}(\cdot) : \{0,1\}^k \mapsto \{0,1\}^{l(k)}$ is a verifiable random function if there exists polynomial-time algorithms ($\mathtt{G}(\cdot)$, $\mathtt{Eval}(\cdot,\cdot)$, $\mathtt{Prove}(\cdot,\cdot)$, $\mathtt{Verify}(\cdot,\cdot,\cdot,\cdot)$) such that

- $\mathtt{G}(1^k)$ is a probabilistic algorithm which output a secret key SK (the secret seed of the random function) and the corresponding public key PK.
- $\mathtt{Eval}(SK, x)$ is an algorithm that computes the VRF's output $y = F_{PK}(x)$.
- $\mathtt{Prove}(SK, x)$ is an algorithm that computes the proof π that $y = F_{PK}(x)$.
- $\mathtt{Verify}(PK, x, y, \pi)$ is an algorithm that verifies $y = F_{PK}(x)$.

And a VRF has to satisfy the following properties:

- *Uniqueness*: If $\mathtt{Verify}(PK, x, y_1, \pi_1) = \mathtt{Verify}(PK, x, y_2, \pi_2) = 1$, $y_1 = y_2$.
- *Computability*: $\mathtt{Eval}(SK, x) = F_{PK}(x)$ is efficiently computable.

[2] In their settings, players are not betting on any particular value.
[3] But we believe that the scheme may be modified to satisfy this requirement.

- *Provability*: $\texttt{Verify}(PK, x, \texttt{Eval}(SK, x), \texttt{Prove}(SK, x)) = 1$.
- *Pseudorandomness*: The probability that an adversary can tell any bit of $F_{PK}(x)$ for x he chose is negligible even if he has seen the values of many $F_{PK}(x')$ given $x' \neq x$.

Please refer to [3, 12, 13] for more details of VRF, e.g. its construction.

To see how VRF helps in the construction of e-lottery scheme, the following shows some key ideas. The random winning number generation is guaranteed by the pseudorandomness of VRF and the generation result is verifiable by the commitment made by dealer (public key). By keeping the private key secret, players are prevented from cheating even if they have access to the VRF's input.

Delaying Function: *Delaying function* was introduced by [6], which is a function that is moderately hard to compute and cannot be parallelized. The time to compute the function depends on the parameters chosen. It may take several hours or several days to complete, using the fastest existing implementation. The output is *publicly verifiable* by going through the same computation. Please refer to [6] for more details of delaying function, e.g. its construction.

Verifiable Encrypted Signature: *Verifiable Encrypted Signature (VES)* is an encrypted signature which can be proved that the decryption of it gives a certain party's signature on a public message, without revealing the signature. It is called *Certificate of Encrypted Message Being a Signature (CEMBS)* in [1]. VES can be constructed from aggregate signature scheme [2] or ElGamal public key encryption scheme. It is a useful primitive for enabling fair exchange [15].

3 A New e-Lottery Scheme

In this section, we describe our proposed e-lottery scheme. We show how to integrate VRF, hash function and delaying function to build a scheme that satisfies all the identified requirements. To ease our discussion, we assume that we only generate one winning number and each ticket bets on one number. It is easy to extend the scheme to cover a combination of numbers.

3.1 Our Proposed Scheme

Our scheme consists of 5 phases, namely, Setup, Ticket Purchase, Winning Result Generation, Prize Claiming and Player Verification. We assume that the numbers used in the lottery game is $\{1, 2, \ldots, u\}$ and the output domain for the VRF is $\{1, 2, \ldots, v\}$ with $v \geq u$. Let k be the security parameter, $H(\cdot)$ be a cryptographic hash function that maps a bit string of arbitrary length (i.e. $\{0,1\}^*$) into a bit string of fixed length (e.g. 160 bits) and let $H_0(\cdot)$ be another cryptographic hash function that maps $\{0,1\}^k$ to $\{0,1\}^{\lceil \log_2 u \rceil}$.

- Setup:
  1. Dealer generates a secret key SK and a public key PK by the algorithm $\texttt{G}$ with the security parameter k as the input.

2. Dealer publishes the following items.
   (a) Hash functions $H(\cdot)$ and $H_0(\cdot)$, delaying function $D(\cdot)$, and VRF's algorithms: $\mathtt{G}(\cdot)$, $\mathtt{Eval}(\cdot,\cdot)$, $\mathtt{Prove}(\cdot,\cdot)$ and $\mathtt{Verify}(\cdot,\cdot,\cdot,\cdot)$.
   (b) VRF public key PK and dealer's digital certificate.
   (c) The amount of time t in which dealer must release the winning ticket value generated. (This is the input parameter controlling the time complexity of the delaying function.)

– Ticket Purchase:
1. Player chooses his favorite number x and randomly picks a random bit string r from $\{0,1\}^k$. r is kept in secret.
2. Player obtains a sequence number s of the ticket from dealer.
3. Player computes $H_0(r)$ and $H(x||s||r)$, then sends $ticket_i = s||(x \oplus H_0(r))||H(x||s||r)$ to dealer.
4. Dealer publishes every single ticket $ticket_i$.
5. Dealer returns a signed ticket $signed\_ticket_i$ to player to acknowledge the recipient of player's purchase request. Any digital signature scheme which is existentially unforgeable against adaptive chosen message attack can do and the signature does not need to be encrypted.
   (Players do not need to enroll to the public key infrastructure as they need to verify the validity of the signature only.)
6. Dealer links the ticket to a one-way hash chain. This chain could be created by $chain_1 = H(ticket_1)$, $chain_i = H(chain_{i-1}||ticket_i)$ for $i > 1$.
7. Dealer publishes $chain_j$ where j is the number of tickets sold so far.

– Winning Result Generation:
1. Suppose the final value of the hash chain is h, computes $d = D(h)$ by the delaying function, and publishes it.
2. Dealer calculates $(w, \pi) = (\mathtt{Eval}(SK, d), \mathtt{Prove}(SK, d))$.
3. If $w > u$, VRF is applied on $w||d$ with padding bits if necessary (one possible source of bits is VRF public key) again.
4. Dealer publishes (w, π) (and intermediate tuples with number of the times VRF is applied, if VRF is applied more than once) within t units of time after the closing of the lottery session.

– Prize Claiming:
1. If $x = w$, player wins.
2. Player submits (s, r) to dealer (in a secure channel).
3. Dealer checks whether a ticket of $s||(w \oplus H_0(r))||H(w||s||r)$ is received.
4. If so, dealer checks whether the tuple (s, r) has already been published (i.e. the prize has been claimed by someone already).
5. If the prize is not yet claimed, dealer pays the player and publishes r.

– Player Verification:
1. Player checks whether his/her ticket(s) is/are included in the hash chain and checks whether the final output of the hash chain is correct, using the knowledge of $ticket_i$'s.
2. Player verifies the validity of d and whether $w = \mathtt{Eval}(SK, D(h))$ by checking whether $\mathtt{Verify}(PK, w, D(h), \pi) = 1$.
3. Intermediate (w, π) tuples may also be checked if necessary.

4. For each winning ticket published, players verify the validity of $s||(w \oplus H_0(r))||H(w||s||r)$.
5. If any mismatch occurs, player should report to the auditing organization by simply providing the parameters involved in the checking.

3.2 Design Philosophy

The rationales behind some important steps of our scheme are discussed here.

Probability of Winning: Suppose $v = u + b$ is the domain size of the VRF's output. The winning probability ensured by the Step 3 of the winning result generation is $\sum_{i=0}^{\infty}(\frac{b}{u+b})^i \frac{1}{u+b} = \frac{\frac{1}{u+b}}{1-\frac{b}{u+b}} = \frac{1}{u}$. We can control the domain size of VRF's output so the expected time of application of VRF can be kept small.

Mapping of the Winning Result to Winning Ticket: In most of the previous e-lottery schemes, the mapping of the winning result to the ticket value is not discussed in details. There may be a situation that the domain of the generated winning result is not the same as the domain of the ticket value. In particular, if the domain size of the winning result and the domain size of the ticket value are relatively prime, it is not straight-forward how one can ensure the pre-determined winning probability of each possible ticket value. In our scheme, the mapping of the winning result to the winning ticket value is trivial as it is a 1-1 mapping.

Use of VRF Key Pairs: To maximize the security level of the scheme, both dealer and players are prevented from gaining knowledge of some output values of VRF by the generation of a new pair of VRF keys for each lottery session.

Use of Sequence Number: The purpose of using sequence number is not the same as that in [6]. In our scheme, it only serves the purpose of making $H(x||s||r)$ distinct even if two players pick the same x and r. The sequence numbers are not required to be following each other one by one strictly. If the sequence number s is not used, the situation that two different winning players who used the same r may occur (although the probability is not high). In such case, when one of the players claimed the prize, dealer has already known the value of r and hence the dealer can falsely claim that the prize of the second winning player has been already paid. In our current design, even the values x and r associated with two different winning tickets are the same, the corresponding sequence numbers are different and hence the values of $H(x||s||r)$ are different.

Adding Fair Exchange Feature: By using CEMBS/VES, fair exchange between the player and dealer can be ensured. The fair exchange protocol used in [15] can be adopted in our scheme as well. For the purchase of tickets, if the player aborts the transaction, the dealer can send the cipher cash, sequence number s and $signed\_ticket_i$ to the TTP. Then TTP will send the e-cash to the dealer and the $signed\_ticket_i$ to the player. For the prize claiming, if the dealer aborts the transaction, the player can send the cipher cash, $ticket_i$ together with r to the TTP. Then TTP will send the e-cash to the player and r to the dealer.

Since there may be more than one player winning the same prize by betting on the same value x, we note that the e-cash used in the prize claiming should

not be of fixed value. It should be an agreement of the dealer to award the player a certain prize. By the publicly verifiability of the number of winning tickets, the amount of award shared by each player is guaranteed.

Use of Delaying Function: Dealer may cheat by trying to "insert" tickets to the hash chain (and "discard" these newly added tickets if the result is not favorable) until a favorable result is obtained. The use of delaying function prevents this kind of attack. One may argue that if the last ticket is sold much before deadline, then the dealer is able to compute delaying functions twice before deadline. However, it is not so probable in a large scale e-lottery. On the other hand, a malicious dealer can corrupt the late submissions, but this is prevented by the fair exchange protocol[4]. Moreover, we found that all existing e-lottery scheme without TTP employs delaying function ([15] also mentioned the use of delaying function is necessary to make their scheme secure). Schemes that do not employ the delaying function, for instance [10], assumed the integrity of the ticket submission pools is ensured by the TTP.

3.3 Analysis

Assuming the lottery is *pari-mutuel*, i.e. the amount of prize is solely based on the sold tickets, our scheme satisfies all the requirements we have identified.

1. *Random generation of winning number.*
   The random generation of winning number is guaranteed by the VRF's pseudorandomness. Each ticket submission from the players is random, thus nobody can predict the outcome before the time when the last ticket is sold. With the use of delaying function, the dealer cannot bias the generation of winning number by trying to insert tickets into the hash chain one by one. So, the dealer has no way to make sure that which number/player will win or loss even all the numbers purchased by the players are known.
2. *The winning result is publicly verifiable.*
   The winning number generation is verifiable by the provable property of VRF. Besides, the value of hash chain and the output of delaying function are also publicly verifiable.
3. *The total revenue and the number of winning tickets are publicly verifiable.*
   (a) The number of sold tickets is verifiable by the first step of player's verification since each player can check whether his/her ticket(s) is/are included in the hash chain.
   (b) For each ticket sold, the dealer will publish the corresponding number x, so each player can verify the total number of winning tickets easily after the winning result is announced.
4. *Forgery of winning ticket is impossible.*
   (a) The dealer cannot forge winning tickets after the winning number has been generated as all the tickets sold have been published and can be publicly verified by all players.

[4] For denial of service attack by dealer, there are other practices to resolve this issue and it is outside the scope of our discussion.

(b) Players cannot forge tickets that they have not actually purchased. The hash function $H(\cdot)$ makes sure that it is difficult to compute r based on the published ticket $s||(x \oplus H_0(r))||H(x||s||r)$.

5. *User is not required to be online at the time of winning number generation.*
6. *Anonymity of player.*
   Player's identity is not reflected in the process of buying ticket, verification and prize claiming. Players' account number can be encrypted as well.
7. *Confidentiality of the ticket's value.*
   The ticket's value is encrypted. Neither cutting down a certain ticket value's winning share nor eavesdropping for the "global statistics" is possible.
8. *Fairness on purchase of tickets and claiming of prize.*
   Fairness can be achieved by using CEMBS/VES as the protocol used in [15].
9. *No early registration of player is required.*

3.4 Other Issues

There are some other minor issues that have been addressed in the previous e-lottery schemes. We briefly discuss how our scheme addresses these issues.

1. *Total purchase of tickets.*
   It is natural that total purchase of all tickets guarantees a winning probability of 1. This threat can be resolved with external means, e.g. by setting the prize obtained by total purchase is less than the cost of total purchase. Because the total revenue is publicly verifiable, there is no use for the malicious dealer inserting all possible winning values to the hash chain and succeed to claim tie whenever some one claim to have a prize, since it will greatly increase the revenue and hence the value of prize to be given out by the dealer.
2. *Unclaimed tickets.*
   The dealer may not be aware of any unclaimed tickets as the tickets' value are encrypted. However, the player's verification is very efficient (in contrast to the time consuming verification in [11]) and hence it is not a real threat.
3. *Other threats.*
   Bogus server, *system database damage*, *reveal of network address of player*, etc. are not considered as there are other practices to resolve these threats.

4 Conclusion

A new e-lottery scheme, based on the verifiable random function, is proposed. Our scheme satisfies all the identified requirements and does not assume the existence of TTP for winning ticket generation.

To make the e-lottery a reality, the way the e-lottery game is conducted should be similar to a traditional one in which players' interaction with the dealer should be kept as simple and minimal as possible. Our scheme has a high degree of resemblance with a traditional one. Players are neither required to be online at the time of winning ticket generation nor to perform time-intensive operations for checking whether the tickets purchased are indeed winning tickets. Early user

registration is not required as our scheme is secure even in the presence of a large size of colluding players. And players are allowed to buy as many tickets as they want. We hope that the result of this work can help to increase the confidence level of customers in participating in e-lottery games.

It is still an open problem to design a publicly verifiable e-lottery scheme without using delaying function and the online participation of players during the generation of the winning result. Another research direction is to make a delaying function which is efficiently verifiable but the computation is still hard.

References

1. Feng Bao, Robert H. Deng, and Wenbo Mao. Efficient and Practical Fair Exchange Protocols with Off-line TTP. In *IEEE Symposium on Foundations of Security and Privacy*, pages 77–85, 1998.
2. Dan Boneh, Craig Gentry, Ben Lynn, and Hovav Shacham. Aggregate and Verifiably Encrypted Signatures. In *EUROCRYPT 03*, LNCS 2656, pages 416–432. Springer, 2003.
3. Yevgeniy Dodis and Aleksandr Yampolskiy. A Verifiable Random Function With Short Proofs and Keys. In *Public Key Cryptography 05*, LNCS 3386, pages 416–431.
4. Pierre-Alain Fouque, Guillaume Poupard, and Jacques Stern. Sharing Decryption in the Context of Voting or Lotteries. In *Financial Cryptography 00*, LNCS 1962, pages 90–104. Springer, 2001.
5. Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to Construct Random Functions (Extended Abstract). In *25th Annual Symposium on Foundations of Computer Science*, pages 464–479. IEEE, 1984.
6. David M. Goldschlag and Stuart G. Stubblebine. Publicly Verifiable Lotteries: Applications of Delaying Functions. In *Financial Cryptography 98*, LNCS 1465, pages 214–226. Springer, 1998.
7. C. Hall and B. Schneier. Remote Electronic Gambling. In *13th Annual Computer Security Applications Conference, ACM Press*, pages 227–230, December 1997.
8. Wooseok Ham and Kwangjo Kim. A Secure On-line Lottery Using Bank as a Notary. In *CISC 02*, pages 121–124.
9. K. Kobayashi, H. Morita, M. Hakuta, and T. Nakanowatari. An Electronic Soccer Lottery System that Uses Bit Commitment. *IEICE Transactions on Information and Systems*, E83-D(5):980–987, 2000.
10. Elisavet Konstantinou, Vasiliki Liagkou, Paul Spirakis, Yannis C. Stamatiou, and Moti Yung. Electronic National Lotteries. In *Financial Cryptography 04*, LNCS 3110. Springer, 2004.
11. Eyal Kushilevitz and Tal Rabin. Fair e-Lotteries and e-Casinos. In *CT-RSA 2001, The Cryptographer's Track at RSA Conference*, LNCS 2020, pages 100–109, 2001.
12. Anna Lysyanskaya. Unique Signatures and Verifiable Random Functions from the DH-DDH Separation. In *CRYPTO 02*, LNCS 2442, pages 597–612. Springer, 2002.
13. Silvio Micali, Michael O. Rabin, and Salil P. Vadhan. Verifiable Random Functions. In *IEEE Symposium on Foundations of Computer Science*, pages 120–130, 1999.
14. Kazue Sako. Implementation of a Digital Lottery Server on WWW. In *Secure Networking - CQRE (Secure) 99*, LNCS 1740, pages 101–108. Springer, 1999.
15. Jianying Zhou and Chunfu Tan. Playing Lottery on the Internet. In *Information and Communications Security - ICICS 01*, LNCS 2229, pages 189–201. Springer.

Related-Mode Attacks on Block Cipher Modes of Operation

Raphael C.−W. Phan[1] and Mohammad Umar Siddiqi[2]

[1] Information Security Research (iSECURES) Lab,
Swinburne University of Technology (Sarawak Campus), 93576 Kuching, Malaysia
rphan@swinburne.edu.my
[2] Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia
umar@mmu.edu.my

Abstract. In this paper, we present a generalization of the notion of the recently proposed related-cipher attacks. In particular, we show that when the cryptanalyst has access to an oracle under one mode, then almost all other related-cipher modes can be attacked with ease. Typically only one chosen plaintext/ciphertext query is required, while computational complexity is negligible.

1 Introduction

Block ciphers are often proposed with several variants, in terms of a different secret key size and corresponding number of rounds. Wu [13] presented the *related-cipher attack* model applicable to related ciphers in the sense that they are exactly identical to each other, differing only in the key size and most often also in the total number of rounds. Such related ciphers must have identical key schedules irrespective of their difference in the total number of rounds.

In this paper, we generalize the concept of the related-cipher attack model to apply to a larger class of related ciphers, in particular cipher encryptions with different block cipher modes of operation, but with the underlying block cipher being identical. We further show that when the cryptanalyst has access to an oracle for any one mode of operation, then almost all other related cipher modes can be easily attacked [1].

In Section 2, we briefly describe the standard block cipher modes of operation. In Section 3, we present our *related-mode attack* model and discuss how access to chosen plaintexts/ciphertexts in any one mode allows almost all other related-cipher modes to be attacked. We conclude in Section 4.

2 Standard Block Cipher Modes of Operation

When encrypting a plaintext, P that is longer than the block size, n of the underlying block cipher, this plaintext is divided into m number of n-bit blocks,

[1] Without loss of generality, we demonstrate this on the standard modes of operation.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 661–671, 2005.

P_i, and each one is encrypted at a time using a block cipher mode of operation that includes the electronic code book (ECB), cipher block chaining (CBC), cipher feedback (CFB) or output feedback (OFB) modes [6, 7].

The ECB mode is the simplest, where each plaintext block, P_i is independently encrypted to a corresponding ciphertext block, C_i via the underlying block cipher, E_K keyed by secret key, K:

$$C_i = E_K(P_i). \tag{1}$$

Figure 1 illustrates the ECB mode encryption on two consecutive plaintext blocks, P_{i-1} and P_i.

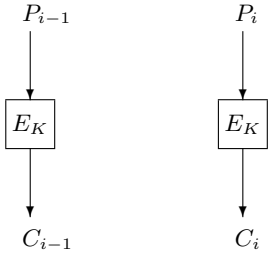


Fig. 1. ECB Mode Encryption

Meanwhile, the CBC mode uses the previous ciphertext block, C_{i-1} as the feedback component that is exclusive-ORed (XORed) to the current plaintext block, P_i, before the resulting XOR is encrypted to obtain the current ciphertext block, C_i. In particular:

$$C_i = E_K(P_i \oplus C_{i-1}) \tag{2}$$

where C_o = initialisation vector (IV). Figure 2 illustrates the CBC mode encryption on two consecutive plaintext blocks, P_{i-1} and P_i.

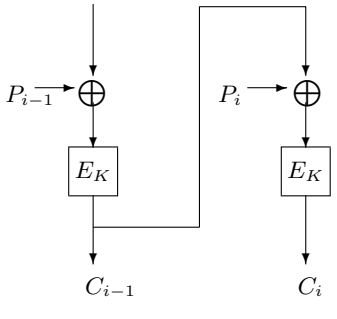


Fig. 2. CBC Mode Encryption

The CFB mode also uses the previous ciphertext block, C_{i-1} as feedback, which is first encrypted and then XORed to the current plaintext block, P_i to obtain the current ciphertext block, C_i:

$$C_i = P_i \oplus E_K(C_{i-1}) \tag{3}$$

where C_o = initialisation vector (IV). The CFB mode can also viewed as a stream cipher mode by treating $X_i = E_K(C_{i-1})$ as a keystream that is XORed to the plaintext, P_i to obtain the ciphertext, C_i. Figure 3 shows the CFB mode.

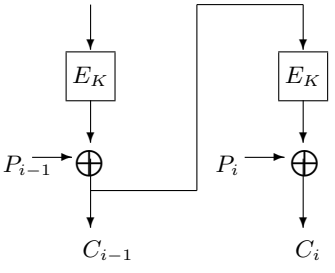


Fig. 3. CFB Mode Encryption

The OFB mode is similar to the CFB in that a keystream is also generated to be XORed to the current plaintext block, P_i to obtain the current ciphertext block, C_i. The difference is that the keystream is not a function of the previous ciphertext block, C_{i-1}, but is the previously encrypted feedback component, X_i:

$$X_i = E_K(X_{i-1})C_i = P_i \oplus E_K(X_i) \tag{4}$$

where X_o = initialisation vector (IV). Note that the keystream is independent of previous plaintext and ciphertext blocks. Figure 4 illustrates the OFB mode.

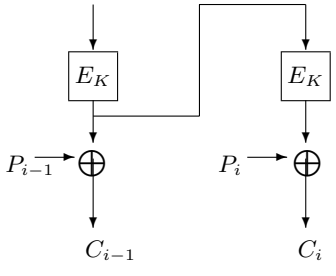


Fig. 4. OFB Mode Encryption

3 Generalization to Related Modes of Operation

We present generalizations of the concept of related-cipher attacks to related modes of operation. We first recall previous work [13] that considered similar ciphers which only differed in the number of rounds, but whose key schedules were the same. Such ciphers include Square [2], which has a standard number of rounds of 8, but the designers further commented that the user was free to choose a number of rounds of their choice. Wu [13] considered the interesting fact that given two such related ciphers, whose number of rounds was respectively r and $r + s$, then one could align their respective encryptions such that their first r rounds are the same. Then the remaining s rounds will allow the cryptanalyst to easily compute the key used in those rounds. We note that we can also view this as a type of slide attack [1] where the encryptions of the two related-ciphers are slid by their first r rounds, with the unslid rounds being the last s rounds:

$$P \rightarrow F \quad F \quad \ldots \quad F \quad F \rightarrow C$$
$$P \rightarrow \underbrace{F \quad F \quad \ldots \quad F \quad F}_{r} \quad \underbrace{F \quad \ldots \quad F}_{s} \rightarrow C'$$

Note that under the related-cipher attack model, the cryptanalyst is interacting not with two distinct related ciphers but in fact with the same cipher but whose number of rounds are variable. One may therefore question that if it is possible to vary the number of rounds, then what would stop the cryptanalyst from reducing the number of rounds down to merely one or two, and therefore render the cipher completely weak against trivial attacks?

We now consider a more common and likely scenario compared to the related-cipher attack setting, where we have the same cipher being used as the underlying component in different block cipher modes of operation. Standard modes [6] of operation are the electronic code book (ECB), cipher block chaining (CBC), cipher feedback (CFB) and output feedback (OFB). Consider the block cipher AES [5], for example, that is seeing widespread usage in various settings, such as in the Pretty Good Privacy (PGP) software [14] and the SSL protocol initially proposed by Netscape [8] for secure online transactions. Both may offer the user the flexibility of choosing a desired mode of operation to use. For example, the user could have chosen the default CFB mode for his PGP software, while the SSL server with which the user is communicating with could have selected the CBC mode. Also, we remark that it is quite common for a user to use the same password (and hence the secret key) with his various internet accounts in order to make it easier for him to remember, rather than having a different password for every account, and then having to write it down somewhere, which he considers even more unsafe and defeats the purpose of having a password altogether. To most, storing the password mentally in one's mind is safer than writing it down on paper. Therefore, since the user uses the same secret key, then when the underlying block cipher is the same, different modes of operation would give different but related-mode ciphers. This provides a fertile background for us to mount *related-mode attacks*: exploiting an oracle containing a certain mode to attack encryptions under other modes.

A similar attack model is the *padding oracle attack* model [11, 9] where the cryptanalyst has access to an oracle that upon receipt of a ciphertext encrypted under a secret key, K, will immediately decrypt it with the same secret key, K and replies if padding is valid or not. Further, their objective is similar to ours in that the cryptanalyst aims to directly recover the unknown plaintext blocks from intercepted ciphertext blocks.

Throughout this paper, we consider the case where the cryptanalyst has access to an oracle that is able to perform *either* encryption or decryption[2] for some fixed mode. This is similar to having access to known or chosen plaintext/ciphertext queries under that mode.

We show that this oracle allows the cryptanalyst to attack the other related-cipher modes, where the underlying block cipher is the same. P'_i and C'_i respectively denote the current plaintext and ciphertext block used in the interaction with the oracle being exploited, while P_i and C_i respectively denote the current plaintext and ciphertext blocks of the related-cipher mode being attacked.

For the mode being attacked, only the corresponding ciphertext blocks, C_i for $(i = 0, 1, \ldots, m)$ of some unknown plaintext blocks, P_i for $(i = 0, 1, \ldots, m)$ are known, where $C_0 = IV$ is the initialization vector[3]. It is the cryptanalyst's objective to directly recover these unknown plaintext blocks, P_i for $(i = 0, 1, \ldots, m)$, i.e. we assume a ciphertext-only scenario for the mode being attacked.

For the mode being exploited, access to its oracle allows the cryptanalyst to obtain known or chosen plaintext/ciphertext queries, − and as necessary known or chosen IV queries − though we assume for more concrete and interesting results that he can only access *either* a mode encryption or mode decryption oracle, and not both at the same time. Having said this, note that a standard mode of operation is expected to be secure against attacks where both encryption *and* decryption oracles are possible [4]. In fact, some attacks on block cipher modes, for instance the recent one [4] on EMD [10] specifically requires both encryption and decryption oracles to be available.

3.1 Exploiting an ECB Oracle

Consider that the cryptanalyst has access to either an encryption or decryption oracle under ECB mode. We will show how this oracle can be exploited to obtain the unknown plaintext blocks encrypted under other modes.

Attacking CBC. In our current case, the cryptanalyst has access to the ECB decryption oracle, and is exploiting it to attack another related cipher in CBC mode. In particular, given that he desires to know the unknown plaintext block, P_i corresponding to an intercepted ciphertext block, C_i of the CBC mode, he chooses $C'_i = C_i$ to feed to the ECB decryption oracle and hence obtains the corresponding P'_i. This P'_i is really gibberish since C'_i is not really a ciphertext that was encrypted under ECB. Instead, C'_i is really C_i encrypted under CBC.

[2] Encryption and decryption operations are generally considered symmetric and interchangeable in terms of complexity.

[3] Except for ECB where C_0 does not exist.

However, since $C'_i = C_i$, then $E_K^{-1}(C'_i) = E_K^{-1}(C_i) = P'_i$. Therefore, he can relate P'_i to a state within the CBC, as indicated in Figure 5, where the exploited oracle and the mode being attacked are on the left and right, respectively, and where the rectangular boxes delimit the parts inaccessible to the cryptanalyst.

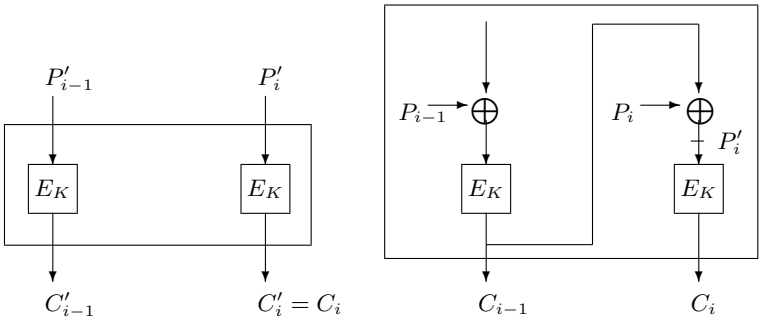


Fig. 5. Exploiting ECB to Attack CBC

This gives $P'_i = P_i \oplus C_{i-1}$, hence he can compute the CBC plaintext block as $P_i = P'_i \oplus C_{i-1}$. In summary, we require just one chosen ciphertext (CC) query encrypted under ECB to obtain the plaintext block corresponding to any ciphertext block encrypted under CBC.

Attacking CFB. Exploiting the ECB to attack the CFB mode is also straightforward. In this case, the cryptanalyst requires access to an ECB encryption oracle, feeding it with just one chosen plaintext (CP) query. In particular, he chooses $P'_i = C_{i-1}$, and so obtains the corresponding $C'_i = E_K(P'_i) = E_K(C_{i-1})$, which directly relates to a state in the CFB, in particular $C'_i = P_i \oplus C_i$, immediately allowing him to compute $P_i = C'_i \oplus C_i$. This is illustrated in Figure 6.

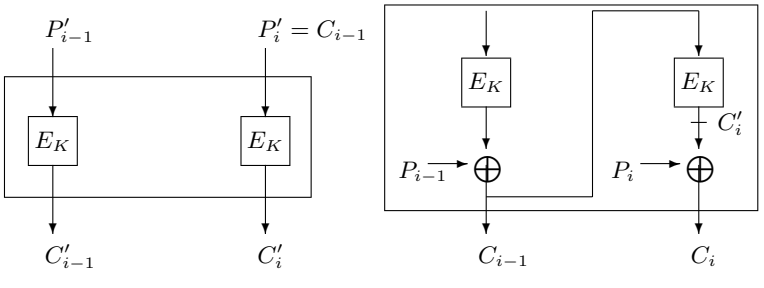


Fig. 6. Exploiting ECB to Attack CFB

Attacking OFB. Attacking the OFB given an ECB decryption oracle similarly requires a chosen plaintext query. In particular, given the known $C_0 = IV$, then

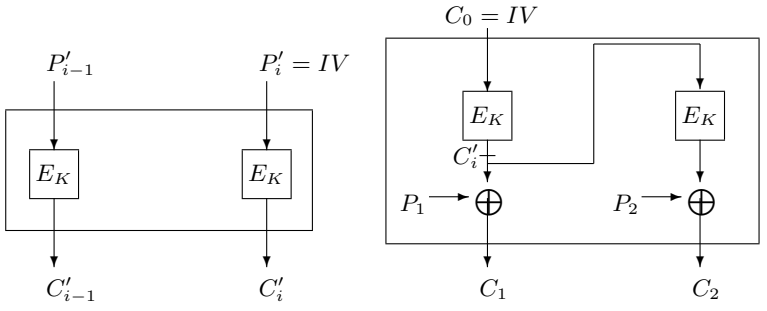


Fig. 7. Exploiting ECB to Attack OFB

merely choose $P'_i = IV$, and hence $C'_i = E_K(P'_i) = E_K(IV)$, which directly relates to an intermediate state in the OFB, in particular $C'_i = P_1 \oplus C_1$, allowing to compute $P_1 = C'_i \oplus C_1$. See Figure 7. Repeat as required to iteratively obtain the next plaintext blocks, P_i for $(i = 2, \ldots, m)$.

3.2 Exploiting a CBC Oracle

When the cryptanalyst has access to a CBC oracle, he can then similarly use this to attack the other related-cipher modes.

Attacking ECB. The cryptanalyst accesses a CBC decryption oracle and chooses $C'_i = C_i$, hence obtaining $P'_i = C'_{i-1} \oplus P_i$. He is then able to compute $P_i = P'_i \oplus C'_{i-1}$. See Figure 8.

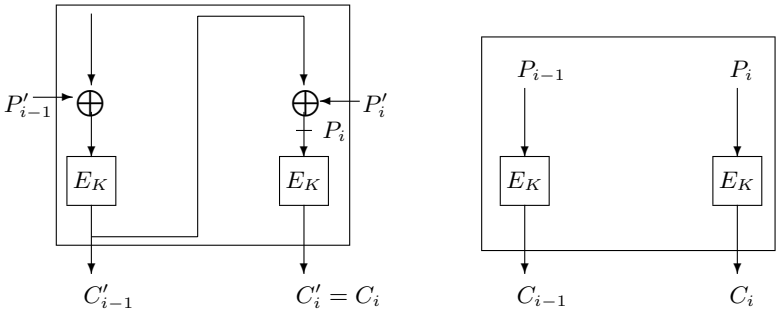


Fig. 8. Exploiting CBC to Attack ECB

Attacking CFB. Attacking this requires a CBC encryption oracle and choosing $P'_i = C'_{i-1} \oplus C_{i-1}$ such that $C'_i = E_K(P'_i \oplus C'_{i-1}) = E_K(C_{i-1})$ is directly related to an intermediate state in CFB, namely that $C'_i = P_i \oplus C_i$. Therefore, we can compute $P_i = C'_i \oplus C_i$, as shown in Figure 9.

Attacking OFB. Attacking the OFB mode is the hardest since both the input to and output from E_K is masked/XORed with consecutive unknown plaintext

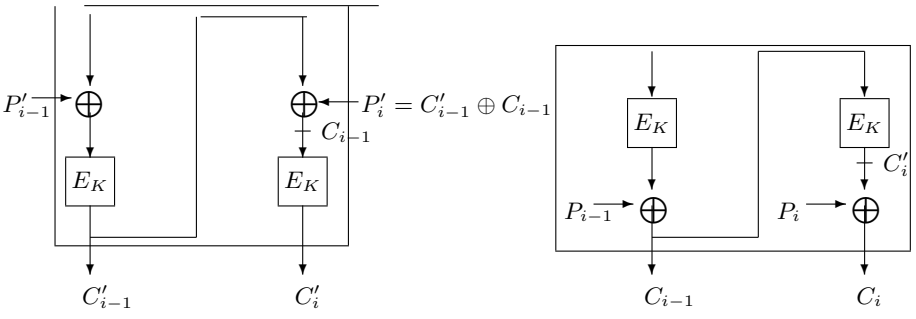


Fig. 9. Exploiting CBC to Attack CFB

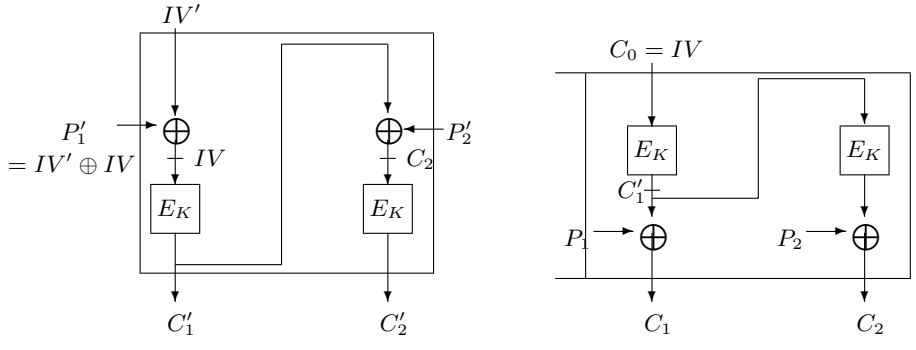


Fig. 10. Exploiting CBC to Attack OFB

blocks, P_{i-1} and P_i. Fortunately, under the known IV scenario [3], the cryptanalyst knows IV' and hence can choose $P'_1 = IV' \oplus IV$, and so have C'_1 be directly related to an intermediate state in OFB, namely $C'_1 = P_1 \oplus C_1$, hence $P_1 = C'_1 \oplus C_1$ can be computed. See Figure 10. This can be repeated to iteratively obtain the next plaintext blocks of the OFB.

3.3 Exploiting a CFB Oracle

In this case, the cryptanalyst has access to a CFB oracle, and uses this to attack other related-cipher modes.

The CFB and OFB modes are sometimes called stream-cipher modes since despite starting with an underlying block cipher, E_K, using it in these modes essentially results in a stream cipher. A stream-cipher mode uses the underlying E_K in both its mode encryption and decryption, in contrast to other non-stream-cipher modes such as the ECB and CBC that use E_K for mode encryption and correspondingly E_K^{-1} for mode decryption. Because of this, it appears that stream-cipher mode oracles can only be used to construct encryption oracles for other non-stream-cipher modes. This means that it will not be possible to exploit a stream-cipher mode oracle (such as CFB and OFB) to attack non-stream-cipher modes (such as ECB and CBC). Instead, we consider only how stream-cipher modes can be exploited to attack other stream-cipher modes.

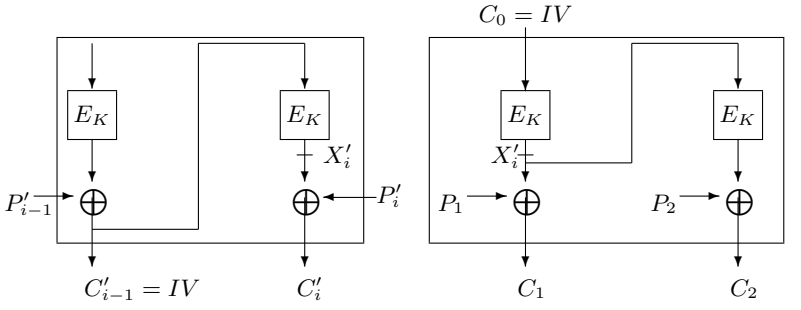


Fig. 11. Exploiting CFB to Attack OFB

Attacking OFB. The cryptanalyst accesses a CFB decryption oracle and chooses $C'_{i-1} = IV$, and hence $E_K(C'_{i-1}) = E_K(IV) = X'_i$ can be directly related to a similar intermediate state within OFB, namely $X'_i = P'_i \oplus C'_i = P_1 \oplus C_1$. He then computes $P_1 = X'_i \oplus C_1 = P'_i \oplus C'_i \oplus C_1$. See Figure 11. Repeating this attack will allow him to iteratively obtain the next plaintext blocks of the OFB.

3.4 Exploiting an OFB Oracle

Similarly, the OFB oracle can only be exploited to attack other stream-cipher modes. We will discuss here how to exploit this to attack the CFB mode.

Attacking CFB. This is so far the hardest attack to mount, and requires a chosen-IV (CIV) scenario [12]. In particular, the cryptanalyst chooses $IV' = IV$, and hence $E_K(IV') = E_K(IV) = X'_i$. This intermediate state relates between the two modes, OFB and CFB, namely $X'_i = P'_1 \oplus C'_1 = P_1 \oplus C_1$, and so he can compute $P_1 = X'_i \oplus C_1 = P'_1 \oplus C'_1 \oplus C_1$. This is shown in Figure 12. Note that in this case the plaintext and ciphertext blocks of the exploited oracle do not need to be chosen but are merely known.

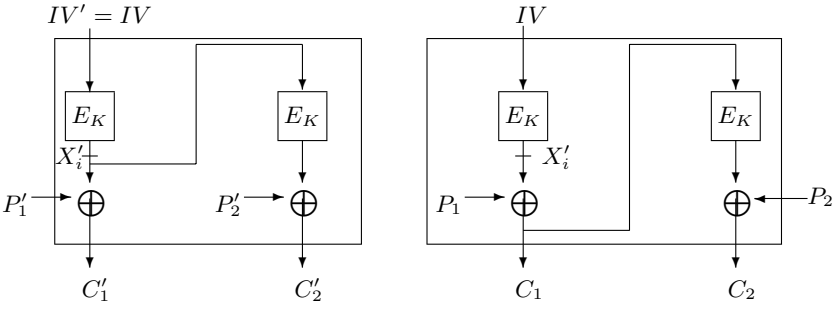


Fig. 12. Exploiting OFB to Attack CFB

4 Concluding Remarks

We have generalized the related-cipher attacks to block cipher modes of operation with the same underlying cipher. Our related-cipher-mode attacks show that when the cryptanalyst has access an oracle containing any one mode of operation, then almost all other related-cipher modes can be easily attacked. In Table 1, we list our attacks and the corresponding text complexities, while computational complexity is negligible.

Table 1. Related-mode Attacks on Modes of Operation

Oracle Exploited	Cipher Mode Attacked	Text Complexity
ECB	CBC	1 CC
	CFB	1 CP
	OFB	1 CP
CBC	ECB	1 CC
	CFB	1 CP
	OFB	1 CP
CFB	OFB	1 CC
OFB	CFB	1 KP, 1 CIV

Acknowledgement

We wish to thank God for His many blessings, and to the anonymous (past and present) referees for their comments.

References

1. A. Biryukov and D. Wagner, "Slide Attacks", FSE '99, LNCS, Vol. 1636, pp. 245–259, Springer-Verlag, 1999.
2. J. Daemen, L. R. Knudsen and V. Rijmen, "The Block Cipher SQUARE", FSE '97, LNCS, Vol. 1267, pp. 13-27, Springer-Verlag, 1997.
3. D. Hong, J. Sung, S. Hong, W. Lee, S. Lee, J. Lim and O. Yi, "Known-IV Attacks on Triple Modes of Operation of Block Ciphers", Asiacrypt '01, LNCS, Vol. 2248, pp. 208-221, Springer-Verlag, 2001.
4. A. Joux, "Cryptanalysis of the EMD Mode of Operation", Advances in Cryptology - Eurocrypt '03, LNCS, Vol. 2656, pp. 1-16, Springer-Verlag, 2003.
5. National Institute of Standards and Technology, "Specification for the Advanced Encryption Standard (AES)" *FIPS 197*. [Online] Available at: `http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf`
6. National Institute of Standards and Technology, "DES Modes of Operation" *FIPS 81*, 1980.
7. National Institute of Standards and Technology, "Recommendation for Block Cipher Modes of Operation − Methods and Techniques" *SP 800-38A*, 2001.
8. Netscape, "Secure Sockets Layer (SSL)", [Online] Available at http://wp.netscape.com/security/techbriefs/ssl.html.

9. K.G. Paterson and A. Yau, "Padding Oracle Attacks on the ISO CBC Mode Encryption Standard", Topics in Cryptology - CT-RSA '04, LNCS, Vol. 2964, pp. 305-323, Springer-Verlag, 2004.
10. P. Rogaway, "The EMD Mode of Operation (a Tweaked, Wide-blocksize, Strong PRP)", Cryptology ePrint archive, http://eprint.iacr.org/2002/148/, 2002.
11. S. Vaudenay, "Security Flaws Induced by CBC Padding − Applications to SSL, IPSEC, WTLS ...", Advances in Cryptology - Eurocrypt '02, LNCS, Vol. 2332, pp. 534-545, Springer-Verlag, 2002.
12. D. Wagner, "Cryptanalysis of Some Recently-proposed Multiple Modes of Operation", FSE '98, LNCS, Vol. 1372, pp. 254-269, Springer-Verlag, 1998.
13. H. Wu, "Related-Cipher Attacks", ICICS '02, LNCS, Vol. 2513, pp. 447-455, Springer-Verlag, 2002.
14. P. Zimmerman, "The Official PGP User's Guide", MIT Press, 1995.

A Digital Cash Protocol Based on Additive Zero Knowledge

Amitabh Saxena, Ben Soh, and Dimitri Zantidis

Dept. of Computer Science and Computer Engineering,
La Trobe University, Bundoora, VIC, Australia 3086
{asaxena, ben, dzantidis}@cs.latrobe.edu.au

Abstract. In this paper, we introduce the concept of *Additive* Non-Interactive Zero Knowledge (NIZK). We extend the notion of NIZK proofs to include the prover's identity as part of the theorem being proved. An additive proof allows a verifier to construct a new proof of knowledge using the information from an old proof. Intuitively, an additive proof is a proof of *knowledge of knowledge*. As an application of this concept, we propose a digital cash scheme with transferable coins.

1 Introduction

Zero knowledge proof systems were initially proposed to replace password-based authentication schemes. Essentially, a zero knowledge proof enables a prover to convince someone (the verifier) about the validity of a statement (the theorem) without revealing any extra information about the statement itself. That is, after the execution of the protocol, the verifier is convinced that the theorem is indeed true without getting any information whatsoever as to *why* it is true. As a classic example of this idea, supposing Alice the mathematician has settled Fermat's last theorem. Using a zero knowledge proof, she can convince a verifier that this is true without even revealing whether she has settled it in the affirmative or has actually found a counterexample.[1]

A formal discussion of zero knowledge proofs and their applications is beyond the scope of this paper and the reader is referred to [2, 3, 4, 5]. In the rest of the paper we will use the notion of zero knowledge informally (that is, without explicitly stating the correctness and soundness criteria if they are clear from the context).

2 Zero Knowledge Proofs

In the simplest form, a zero knowledge protocol is interactive. That is, it requires many interactions (in the form of challenge-response) between the prover and

[1] We note for the record that Fermat's last theorem has been recently settled in the affirmative [1].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 672–680, 2005.

verifier. After a sufficient number of successive correct responses, the verifier is convinced about the correctness of the theorem being proved. The number of challenges is chosen to be large enough so that the probability of a dishonest prover succeeding is very low.

Although suitable for most purposes, the interactive systems have an inherent weakness: the prover needs to be 'on-line' when proving a theorem. Subsequently, the concept of non-interactive zero knowledge proof systems was proposed where the interaction between the prover and verifier is eliminated [6, 7, 8].

2.1 Non-interactive Zero Knowledge

A Non Interactive Zero Knowledge (NIZK) proof system enables an honest prover to convince the validity of a statement to an arbitrary verifier by attaching a 'short' proof. The zero knowledge property is obtained by using a common random reference string [7, 9].

As an example, assume that Alice is a (possibly dishonest) prover and Bob is an arbitrary verifier. Alice knows a secret u and wants to convince Bob about this fact. To participate in the protocol, Alice and Bob must share a common random string a. Alice creates a statement $s =$ "*I know the value of* u" and a proof, p. We say (s, p, a) is a zero knowledge proof if Bob can verify that statement s is true without obtaining any 'useful' information about u.

Most NIZK proof systems require the prover and verifier to have interacted sometime in the past to share the random string. We introduce the notion of NIZK proofs that completely eliminate this interaction. Thus, using this idea Alice can create a 'public' NIZK proof of her statement without compromising her secret. This is done using an *identity* based proof.

Ordinarily any NIZK proof is automatically transferable. That is, Bob can reuse the triple $\{s, p, a\}$ using the same random string to falsely claim that he knows u. We shall see that an identity based proof is non-transferable.

2.2 Identity Based NIZK Proofs

Intuitively, an identity based proof can be considered as a proof of 'knowledge of knowledge'. The basic idea is the same; the verifuer is convinced of the validity of a statement in a non-interactive way. The statement in this case includes the identity of the prover. In other words, the statement s, above, is modified to read "*Alice knows the value of* u". To understand zero knowledge in this new scenario, observe that the proof only validates the correctness of s. It does not disclose the identity of the prover. Also, Bob does not have any significant advantage since the proof does not reveal any information that will enable him to claim knowledge of u. Note that the random string need not be pre-shared in this case since the proof is verifier independent.

2.3 Additive NIZK Proofs

We extend the notion of NIZK proofs to include multiple identities and propose the concept of *additive* zero knowledge. The word additive implies that a verifier can construct a new identity based NIZK proof of knowledge from an old proof

such that new proof is based on the same random string. If this process can go on many times, the participants in the protocol are said to be connected in a *chained* trust relationship.

To understand the concept of additive zero knowledge, include a third participant, Carol in the scenario. To begin with, Alice makes a NIZK proof p_A of the statement $s_A =$"*Alice knows the value of* u" and sends $\{s_A, p_A\}$ to Bob. After verifying that the proof is correct, Bob creates a new NIZK proof p_B from p_A of the statement $s_B =$ "*Bob knows that the statement* s_A *is true*". He keeps p_A secret (as evidence in case of a dispute) and sends $\{s_B, s_A, p_B\}$ to Carol. The proof p_B should convince Carol about the validity of s_B without revealing any information about the validity of s_A on its own. In other words, it should not be possible to compute p_A just from the information known to Carol. In a truly additive system, Carol should be able to further 'add' to the statement by creating the statement $s_C =$ "*Carol knows that the statement* s_B *is true*" and create a corresponding NIZK proof P_C such that the values $\{s_C, s_B, s_A, p_C\}$ also constitute a new NIZK proof. Without discussing this concept formally, we propose an application using a digital cash protocol that will demonstrate the additive property in an intuitive way.

2.4 Digital Cash Scenario

To explain the need for the additive property, we propose a transferable digital cash scheme. For an overview of digital cash, the reader is referred to [10, 11]. Our scenario is based on the following setting: a user can withdraw an *electronic coin* from the bank and pass it on to other users. Eventually, one of the participants will decide to deposit the coin into the bank. We don't limit the number of times a coin may get transferred before finally being deposited. Each coin has associated with it a secret value known only to the withdrawer who creates an identity based NIZK proof of this fact when transferring the coin. Subsequent participants create identity based proofs of the validity of the statement of the previous participant. To make this concept feasible, we need to introduce new cryptographic primitives which we describe next.

3 Associative One-Way Functions

The main building blocks of our protocol are a family of one-way functions defined in this section:

Fix the alphabet $\Sigma = \{0, 1\}$. The set of strings over Σ is denoted by Σ^*. For any string $x \in \Sigma^*$, let $|x|$ denote the length of x. Let $\sigma : \Sigma^* \times \Sigma^* \mapsto \Sigma^*$ be some total binary function. Define the following properties for σ:

1. Polynomial-time computable: We say σ is polynomial time computable if and only if there exists a polynomial p such that for all $a, b \in \Sigma^*$, computing time of $\sigma(a, b)$ is upper-bounded by $p(|a| + |b|)$.
2. Honest: If σ does not shrink its inputs by more than some fixed polynomial. Formally, we say σ is honest if and only if there exists a polynomial q such

that for each $c \in image(\sigma)$, there exists a pair $(a, b) \in \Sigma^*$ such that $c = \sigma(a, b)$ and $|a| + |b| \leq q(|c|)$.

3. Associative: If we have $\sigma(\sigma(a, b), c) = \sigma(a, \sigma(b, c))$ for all $a, b, c \in \Sigma^*$.
4. Non-invertible: If given any random $c \in image(\sigma)$, the probability of finding a pair (a, b) such that $c = \sigma(a, b)$ is negligible. Formally, σ is non-invertible if and only if there exists no polynomial r such that the running time of any algorithm for computing (a, b) is upper-bounded by $r(|c|)$.
5. Strongly non-invertible: If given any $c = \sigma(a, b)$ and a, the probability of finding $\hat{b}$ such that $c = \sigma(a, \hat{b})$ is negligible and, given any $c = \sigma(a, b)$ and b, the probability of finding $\hat{a}$ such that $c = \sigma(\hat{a}, b)$ is negligible. Formally, σ is strongly non-invertible if and only if there do not exist polynomials s_1 and s_2 such that the running times of any algorithms for computing such $\hat{a}$ or $\hat{b}$ are upper-bounded by $s_1(|c| + |b|)$ or $s_2(|c| + |a|)$ respectively.
6. Non-commutative: If for any pair (a, b) randomly selected from a chosen finite subset of Σ^*, we have $\sigma(a, b) \neq \sigma(b, a)$ with overwhelming probability.

If properties 1-4 hold, we say σ is an *Associative One Way Function* (AOWF). If property 5 also holds then σ is *strong* AOWF. We say that σ is a *strong non-commutative* AOWF if properties 1-6 are satisfied. The notion of honesty is needed to preclude functions that are non-invertible simply because some sets of images lack polynomially short preimages.

We propose an additive zero knowledge scheme for digital cash using strong non-commutative AOWFs. Strong AOWFs have also been used in multi-party protocols for digital signatures and key agreement [12, 13]. We note for the record that a practical construction of a strong AOWF for cryptographic applications is not known as yet[2]. See [14, 15] for some related work on complexity-theoretic strong one-way functions.

4 Our Digital Cash Protocol

Assume that P_0 is the first participant who withdraws a coin from the bank and $P_1, P_2, \ldots P_n$ are successive participants in the protocol. The coin is deposited to the bank by P_n. Denote the coin by M. Our protocol should ensure that a participant cannot use the same coin twice without being caught.

Let $f(\cdot, \cdot)$ be a strong non-commutative AOWF as defined in section 3 and let S be a finite set over which f is non-commutative and closed. Let m be the output of applying to M, a collision-resistant hash function mapping to S. It is assumed that finding random elements from S is easy. We will sometimes use the alternate infix notation $x \circ y$ to represent $f(x, y)$. Note that due to the associativity of f, the parenthesis in an expression like $(x \circ y) \circ z$ can be removed. We also use the idea of secure notice boards discussed in [16] to exchange public information in an open way. We assume that coin transfer is done over an insecure channel. This

[2] In the context of this paper, the definitions of 'non-invertible' and 'strongly non-invertible' are based on the average-case complexity model.

implies that a user can send the coin in an email. Our scheme demonstrates a public digital cash system. The coins can be transferred in public without compromising the security of the system.

4.1 Initial Setup

An initial setup is necessary during which a public key directory is created. We note that this is a one-time process. The notice board acts as the central authority. The protocol proceeds as follows:

1. For each P_i, the notice board generates a random $a_i \in S$
2. Each P_i generates a random private key $x_i \in S$ and computes $y_i = f(x_i, a_i)$. The public key is the pair (a_i, y_i).

The public keys are made available via the notice board. After this step, all such registered users will be able to participate in the protocol. That is, they can withdraw, transfer or deposit an electronic coin.

4.2 Withdrawal Protocol

For each coin M, P_0 generates a random secret *signing* key u and computes the corresponding public *verification* key $v = f(m, u)$. It asks the bank to certify the values (m, v). This is done using any standard signature scheme, the details of which we need not discuss here. Denote by B, the certificate from the bank. To avoid known cipher text attacks, a time stamp is included in the certificate. Users who created their public keys after this time are precluded from participating in this protocol. Additionally, the bank must ensure that u is indeed randomly generated.

4.3 Transfer Protocol

We assume that a signature is always attached to the coin. An arbitrary participant P_i will process the coin as follows: On receiving it from P_{i-1}, it first follows the verification procedure. Before passing on the coin to P_{i+1}, it follows the signing procedure. P_0, however, only follows the signing procedure. We assume that the certificate, B must always accompany the message. The following additional definitions will be useful:

Define $z_0 = f(u, x_0)$ and $z_i = f(f(z_{i-1}, a_{i-1}), x_i)$ if $i > 0$.
Define $w_0 = y_0$ and $w_i = f(w_{i-1}, y_i)$ if $i > 0$.
Define $r_i = f(f(m, y_{i+1}), x_i)$

Note that z_i can be computed directly from z_{i-1} while w_i can be computed easily from the public keys y_0, $y_1, \ldots y_i$. However, it is not possible to compute z_0 without knowledge of the secret signing key, u which is only known to P_0. Denote by $\tilde{M}_i$, the signature of P_i on M using a standard signature scheme.

Signing

The signature P_i on M is the set $\{B, z_i, \{r_0, r_1, \ldots r_i\}, \{\tilde{M}_0, \tilde{M}_1, \ldots \tilde{M}_i\}\}$. The string, "$P_0, P_1, \ldots P_i$" containing the list of names is also assumed to be

a part of the signature. Note that P_i receives z_{i-1} in the signature from P_{i-1} and computes z_i using the procedure described above. P_i keeps z_{i-1} secret as evidence that the coin was indeed received from P_{i-1}.

Verification

For clarity, we describe the verification procedure to be followed by P_{i+1}. It consists of four tests and the message is rejected if any of them fail:

1. Verify the certificate B and obtain verification key v.
2. Check that $f(f(m, z_i), a_i) = f(v, w_i)$.
3. Check that $f(f(m, y_{j+1}), y_j) = f(r_j, a_j)$ for all j from 0 to i.
4. Verify $\tilde{M}_j$ (the signature of P_j) for each j from 0 to i.

4.4 Deposit Protocol

The deposit protocol is followed when P_n wants to deposit the coin to the bank. Before passing the coin to the bank, P_n follows the signing procedure. Thus P_n sends $\{M, B, z_n, \tilde{M}_0, \tilde{M}_1, \ldots \tilde{M}_n\}$ to the bank with a deposit request. The bank follows the verification procedure described above. In other words the bank performs the following tests:

1. Verify the certificate B and obtain verification key v.
2. Check that $f(f(m, z_n), a_n) = f(v, w_n)$.
3. Check that $f(f(m, y_{j+1}), y_j) = f(r_j, a_j)$ for all j from 0 to n.
4. Verify $\tilde{M}_j$ (the signature of P_j) for each j from 0 to n.

If the procedure succeeds, the bank accepts the coin as valid. It also keeps a record of the ordered list of users who participated in the protocol. If it receives the same coin for deposit from a different user, it can easily find out the cheating user by comparing the two lists. The list will be different onwards from the point where some participant used the coin more than once.

5 Overview of the Protocol

The above protocol is an example of an additive zero knowledge proof system. To understand this, observe that z_0 cannot be computed without knowledge of u and x_0 which are both kept secret by P_0. Also note that it is not possible for another user $\hat{P}_0$ to create a signature $\hat{z_0}$ on M using v as the verification key without knowing u. Thus, z_0 is a *zero-knowledge* signature of P_0 on M with v being a common random reference string[3]. Moreover, since M and v cannot be

[3] To verify the signature, check that $f(f(m, z_0), a_0) = f(v, y_0)$. In this context, the terms 'Non-Interactive Zero Knowledge proof' and 'Zero Knowledge signature' are interchangeable since, alternatively, $\{z_0, m, v\}$ is a NIZK proof of knowledge of u with m being the random reference string.

un-linked due to the certificate B, it is ensured that a different reference string cannot be used for M.

The additive property is demonstrated by the fact that P_1 can add more information to z_0 by computing z_1 in a way that z_1 still retains zero-knowledge with respect to v. Note that computing any z_{i-1} just from z_i is considered infeasible due to the assumed properties of f.

Step 4 of the verification process is necessary to ensure that users cannot add arbitrarily extra information since otherwise, a user P_i could add a name P_j before itself simply by computing $z_i = f(f(f(z_{i-1}, a_{i-1}), y_j), x_i)$ corresponding to the modified list and there would be no way to tell if P_j's involvement was real or 'simulated'. Step 3 is necessary to ensure that the coin is useful only for the intended receiver.

Thus, each z_i is a NIZK proof of two facts: (a) P_0 knows u corresponding to (M, v) and is therefore the first name in the list and, (b) Information was added i times to z_0. The zero-knowledge property ensures that previously added information cannot be removed.

Assuming that all users are unique, a few points about this protocol are noteworthy:

1. Each P_i who passes the coin must include its name in the signature and in the right sequence for validation to succeed.
2. Users cannot remove names of other users from the list in the signature without knowledge of their private keys, nor can they change the order or add new names.
3. The protocol is completely non-interactive.

It is easily seen that the signing time is independent of the number of users. However, the signature length and the verification time increase linearly with the number of users in the list. This is not a problem unless the list becomes very large. To avoid this, we can restrict the number of times a coin may be passed before it becomes invalid.

6 Security Analysis

In this section, we outline a rough security analysis of our protocol. We consider an attack to be successful if any user can use a coin twice or modify the list to contain false information without detection. Assuming that P_i is the attacker, a combination of the following attacks are possible:

1. It uses the coin but does not include its name in the list.
2. It adds one or more names to the list.
3. It deletes one or more names from the list.
4. It changes the order of some names in the list.
5. It uses the coin more than once.

Additionally, the only useful secret information available to P_i is the signature from P_{i-1}. We consider each case separately.

1. The first possibility is ruled out since the message is always passed over a secure channel, which allows sender authentication.
2. Arbitrary names cannot be added to the list because P_i cannot compute signatures on behalf of other users. Thus, if a false user is added to the list, step 3 of the verification process will fail.
3. It appears that deleting names is not possible either. Firstly note that the verification key (used in step 2 of the verification process) necessarily has to be v. This implies that P_i would have to compute values $(\hat{w_i}, \hat{z_i})$ such that $f(f(m, \hat{z_i}), a_i) = f(v, \hat{w_i})$ and one of the following holds:
   (a) $\hat{w_i} = y_i$ or,
   (b) $\hat{w_i} = y_{r_0} \circ y_{r_1} \circ y_{r_2} \circ \ldots y_{r_j}$ where:
   $\{r_0, r_1, \ldots r_j\} \subset \{0, 1, \ldots i\}$ and $r_j = i$
   Due to the properties of f (specifically, non-invertibility and strong non-invertibility), it seems that neither of these cases are possible.
4. Similarly changing order is not possible due to the non-commutativity of f. To compute w_i, the public keys need to be used in the right order otherwise verification will fail with a high probability.
5. Finally if P_i spends the coin twice, it will be detected by the bank.

Note that even if an attacker manages to intercept the coin during transit, he or she does not have any additional advantage since step 3 of the verification process ensures that the coin is useful only for the intended recipient.

In the above analysis, we assumed that all participants in the protocol are unique. In the event that a coin comes back to a participant, he or she will have the ability to delete all subsequently added names in the list. However, this is not a problem since it make the participant automatically liable in case the coin is used more than once by one of those subsequent users.

7 Conclusion

In this paper, we demonstrated the concept of additive zero knowledge by proposing a digital cash protocol. The protocol uses the idea of transferable electronic coins. The scheme enables traceable cash and does not provide spender anonymity. An advantage of this scheme is that multiple transactions can be done with the same coin. Users are deterred from using a coin more than once due to the possibility of detection.

Additive zero knowledge proofs also have applications in mobile agent authentication. See [17, 18] for authentication protocols based on strong AOWFs using a similar concept.

We conclude this paper with an open question: is it possible to achieve a practical implementation of our digital cash scheme without using strong non-commutative AOWFs?

References

1. Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *AM*, 141(3):443–551, 1995.

2. Oded Goldreich. Zero-knowledge twenty years after its invention, 2002.
3. Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *Journal of Cryptology*, 7(1):1–32, 1994.
4. Oded Goldreich. *Foundations of Cryptography*, volume Basic Tools. Cambridge University Press, 2001.
5. Shafi Goldwasser. New directions in cryptography: Twenty some years later (or cryptography and complexity theory: A match made in heaven). In *In 38th Annual Symposium on Foundations of Computer Science*, pages 314–324. IEEE, 1997.
6. Manuel Blum, Paul Feldman, and Silvio Micali. Non-interactive zero-knowledge and its applications. In *STOC '88: Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 103–112. ACM Press, 1988.
7. Manuel Blum, Alfredo De Santis, Silvio Micali, and Giuseppe Persiano. Noninteractive zero-knowledge. *SIAM J. Comput.*, 20(6):1084–1118, 1991.
8. Mihir Bellare and Shafi Goldwasser. New paradigms for digital signatures and message authentication based on non-interactive zero knowledge proofs. In *CRYPTO '89: Proceedings on Advances in cryptology*, pages 194–211. Springer-Verlag New York, Inc., 1989.
9. Uriel Feige, Dror Lapidot, and Adi Shamir. Multiple noninteractive zero knowledge proofs under general assumptions. *SIAM J. Comput.*, 29(1):1–28, 2000.
10. Jan Camenisch, Ueli M. Maurer, and Markus Stadler. Digital payment systems with passive anonymity-revoking trustees. In *ESORICS*, pages 33–43, 1996.
11. Patiwat Panurach. Money in electronic commerce: digital cash, electronic fund transfer, and ecash. *Commun. ACM*, 39(6):45–50, 1996.
12. Muhammad Rabi and Alan T. Sherman. An observation on associative one-way functions in complexity theory. *Inf. Process. Lett.*, 64(5):239–244, 1997.
13. Amitabh Saxena and Ben Soh. A new paradigm for group cryptosystems using quick keys. In *Proceedings of the The 11th IEEE International Conference on Networks (ICON2003)*, pages 385–389, Sydney, Australia, 2003.
14. Lane A. Hemaspaandra, Jörg Rothe, and Amitabh Saxena. Enforcing and defying associativity, commutativity, totality, and strong noninvertibility for one-way functions in complexity theory. Technical Report UR CSD;854, University of Rochester, December 2004 2004.
15. Lane A. Hemaspaandra, Kari Pasanen, and Jörg Rothe. If p $\neq$ np then some strongly noninvertible functions are invertible. In *FCT '01: Proceedings of the 13th International Symposium on Fundamentals of Computation Theory*, pages 162–171. Springer-Verlag, 2001.
16. Amitabh Saxena and Ben Soh. Contributory approaches to centralized key agreement in dynamic peer groups. In *Proceedings of the The 11th IEEE International Conference on Networks (ICON2003)*, pages 397–402, Sydney, Australia, 2003.
17. Amitabh Saxena and Ben Soh. A novel method for authenticating mobile agents with one-way signature chaining. In *Proceedings of The 7th International Symposium on Autonomous Decentralized Systems (ISADS 05)*, China, 2005. to appear.
18. Amitabh Saxena and Ben Soh. Authenticating mobile agent platforms using signature chaining without trusted third parties. In *Proceedings of The 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE-05)*, Hong kong, 2005. to appear.

On the Security of Wireless Sensor Networks

Rodrigo Roman[1], Jianying Zhou[1], and Javier Lopez[2]

[1] Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
roman@lcc.uma.es
jyzhou@i2r.a-star.edu.sg
[2] E.T.S. Ingenieria Informatica, University of Malaga, 29071, Malaga, Spain
jlm@lcc.uma.es

Abstract. Wireless Sensor Networks are extremely vulnerable against any kind of internal or external attacks, due to several factors such as resource-constrained nodes and lack of tamper-resistant packages. As a result, security must be an important factor to have in mind when designing the infrastructure and protocols of sensor networks. In this paper we survey the "state-of-the-art" security issues in sensor networks and highlight the open areas of research.

1 Introduction

Wireless Sensor Networks [1], composed of hundreds or thousands of inexpensive, low-powered sensing devices with limited computational and communication resources, provide a useful interface to the real world with their data acquisition and processing capabilities. Sensor networks can be applied to a large number of areas, and its applications are continuously growing.

However, sensor networks are extremely vulnerable against any type of internal or external attacks, due to resource constraints, lack of tamper-resistant packaging, and the nature of its communication channels. In this scenario, any protocol, architecture or application which is not developed with security in mind is hardly useful.

As a result, it is essential to incorporate security, or at least to discuss whether it should be applied or not and why, inside the design of every aspect of a sensor network. In this paper, we present a survey of the "state-of-the-art" security issues and algorithms that a designer must have in mind while working with sensor networks.

The rest of the paper is organized as follows. In Section 2, we introduce the sensor network infrastructure and elements. In Section 3, we investigate the "state-of-the-art" security issues in sensor networks such as security primitives, key infrastructure, routing, data aggregation, auditory etc. In Section 4, we conclude our paper, highlighting the research challenges in those areas.

2 Sensor Network Infrastructure

The infrastructure of a sensor network can be divided into two parts, *data acquisition network* and *data dissemination network*.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 681–690, 2005.

- The data acquisition network contains the sensor network "per se": a collection of sensor nodes with the task of measuring the physical data of its surroundings, and one or more base stations in charge of collecting data from the nodes and forwarding control information from the users.
- The data dissemination network is a combination of wired and wireless networks that provides an interface of the data acquisition network to any user, and its security is out of scope of this paper.

Sensor nodes are densely deployed either very close or inside the object to be observed, and all measurements must be routed to the base station where users can have access to them. All the nodes are highly constrained devices. Their memory, computational power and battery life are very limited. On the other hand, base stations are not as constrained as the sensor nodes, and in most cases have no battery shortage problem.

Due to the extreme constraints of the network infrastructure, a sensor network is highly vulnerable against any external or internal attack, thus the infrastructure and protocols of the network must be prepared to manage these kinds of situations. Protecting the information flow not only requires a set of power-efficient encryption schemes, but also an effective key infrastructure in terms of key storage policies, key distribution procedures and key maintenance protocols. Collecting the information from a static or dynamic set of nodes and routing it through the error-prone, unreliable network is a difficult task as well. Moreover, the network should be able to monitor over any failures or security breaches in any of its members while self-configuring and self-healing itself.

3 Security Issues in Sensor Networks

3.1 Security Primitives

All sensor nodes inside a sensor network use power-efficient radio transceivers for their communications. Most of the existing sensor nodes operate in unlicensed frequency bands, but some nodes follow the IEEE 802.15.4 standard for Personal Area Networks [3]. In any case, a malicious adversary can easily get access to the information flow of a sensor network, because sensors are usually scattered in an uncontrolled or public environment and the wireless communication channel is inherently insecure. Consequently, any device can eavesdrop or inject packets inside the sensor network.

It is indispensable to provide basic security primitives to the sensor nodes in order to give a minimal protection to the information flow and a foundation to create secure protocols. Those security primitives are *symmetric key encryption* schemes (SKE), *message authentication codes* (MAC), and *public key cryptography* (PKC). Since sensor nodes are highly constrained in terms of resources, implementing the security primitives in an efficient way (using less energy, computational time and memory space) without sacrificing the strength of their security properties is one of the major challenges in this area.

Existing sensor nodes are able to incorporate software-based SKE with minor overhead in terms of CPU, energy and memory footprint. A proof-of-concept implementation is the TinySec [4] project. Hardware-based SKE is also provided in nodes with radio chips conforming to the 802.15.4 standard [3], although not all the security suites used by the standard are actually secure [5].

Regarding the MAC, it is usually computed using a cipher block chaining construction, called CBC-MAC. It is efficient and fast, and reduces the memory footprint needed for the MAC calculations due to the shared primitives between itself and the SKE. It is used by both hardware [3] and software [4] configurations.

Finally, PKC in software has been usually rejected as "not possible" in a sensor network environment, but there were almost no experiments that backed up the claim. Some studies [6] claimed that *elliptic curve cryptography* (ECC) seemed to be a good candidate for implementing PKC over sensor networks due to its small key size, faster computation, as well as memory and energy savings compared with other algorithms such as RSA. That claim was empirically demonstrated by a recent work in the area that developed a usable PKC in TinyOS [7].

3.2 Global Key Infrastructure

The communication channels between any pair of devices inside the sensor network must be protected to avoid attacks from external parties. This protection is provided by the security primitives introduced in the previous section, but all those primitives need to store a set of secret keys inside every node. Thus it is necessary to have a global key infrastructure.

There are three basic factors in the design of a key infrastructure for sensor networks: *key storage*, *key distribution*, and *key maintenance*.

- Key storage policies indicate the number of keys that a sensor node needs to store in order to open secure communication channels with other peers. It will influence over the network resilience, which defines the percentage of the network that can be controlled by an adversary after he steals the keys from a subset of the nodes, and over the amount of memory available to the node.
- The key distribution protocols define how the keys are issued to the sensor nodes. A node can receive its keys before the initial deployment of the network or create its keys after the deployment using preloaded information.
- The key maintenance procedures specify how a node can be included into or erased from the network, receiving or nullifying a set of keys in the process.

In terms of key storage, there are two extreme design cases: *global keying* (a single key is created for the entire network) and *pairwise keying* (a node must store a key for every other node inside the network). Neither of these cases is feasible in a real scenario: Global keying has no network resilience while pairwise keying is not a scalable solution due to the memory constraints of the nodes. Therefore, security researchers have been trying to develop more optimal solutions, such as pairwise keying only with every direct neighbor.

"Key pool" paradigm, introduced in [8], seeks to obtain a balance in the key storage policy while predistributing the secret keys before the deployment. In this paradigm, every sensor retrieves a certain number of keys from a common "key pool", and only the nodes that share a key (or a certain set of keys [9]) from their own pools can interchange messages securely. The number of keys of a pool and the number of keys that every node retrieves from the pool are factors that influence over the network resilience, memory usage and network connectivity.

Evolutions of the original "key pool" scheme aim to optimize the construction of the key pool and/or the distribution of keys from the pool, assuring a local coverage of 100% in most cases while decreasing the size of the local pools. Those optimizations are based on the Bloom scheme [10, 11, 12] or on combinatorial designs [13]. Other schemes improve the distribution of the keys using "a priori" information about the physical deployment of nodes [14], where nodes that are in the same physical area receive keys from the same subset of the key pool.

Other protocols can offer a balanced key storage policy while creating the keys after the network deployment. There is one solution that relies on negotiating the pairwise keys of a neighborhood with the base station [15], although this would be inconvenient for highly populated networks. A simpler model allows nodes to negotiate with their neighbors the pairwise keys, in "clear", just after the network deployment [16], because the treat level at this point is, in most scenarios, very low.

The advent of public key cryptography over sensor nodes [7] opens a new, uncharted area in the field of key infrastructures for sensor networks. PKC could help on the secure creation of pairwise keys after deployment and on the key maintenance procedures, which is a field in key infrastructure not fully addressed in the previous schemes.

3.3 Local Key Infrastructure - Secure Groups

There are some situations during the lifetime of a sensor network where a subset of the sensor nodes must group themselves in order to cooperate and fulfill a certain task. These groups must have a local key infrastructure, allowing them to open secure channels between members of the group and to broadcast messages inside the group. Securing a group inside an already protected sensor network is not redundant, because there are some cases where the group needs that protection.

Authentication is an important issue in protecting secure groups. A message addressed to all or a subset of the group must be properly authenticated, or any message from inside or outside the sensor network can be mistaken, deliberately or not, as addressed to the group. Confidentiality is also important, because in some cases, such as measuring critical factors in nuclear power plants, the group would want to hide the measurements from the other parts of the network. Finally, the integrity of the messages is critical as well, because without it the measurements and control messages would be prone to be attacked.

As in the global key infrastructure, there are three basic factors to be solved when designing the key infrastructure of a secure group: key storage policies,

key distribution protocols, and key maintenance procedures. However, securing a sensor group is completely different from securing the entire network. First, groups are normally created dynamically, when the base station commands to do so or where a particular set of readings (e.g., a truck approaching) force the network to organize itself. In these cases the keys of the group must be negotiated and distributed to all the members. Second, the nodes belonging to a group must be able to store all the necessary keys for the secure communications, having in mind that there may be no memory space at all in some extreme cases. Third, nodes will be added and deleted from the group in a more frequent basis, as for example when a truck is moving over the sensor field. These maintenance operations must be safe for the group, in the sense that an external node should not enter into the group when it is not invited and an internal node should not leave the group when it is not the time. Finally, the group should satisfy two more requirements: "forward security", where nodes left the group should not be able to access the current information flow of the group, and "secure tunnel", where a measurement made by the group and directed to the base station should not be accessed by the routing nodes in some essential cases, such as the nuclear power plant scenario.

The topic of secure grouping has not been intensely researched over the past years, and only few resource-demanding solutions exist [17]. An exception has been the protection of static groups, created before the initial deployment of the network, where more powerful nodes called "cluster heads" are in charge of managing and protecting the group [18]. Still, new optimal schemes that allow the sensor network to create and maintain secure groups by itself using as less resources as possible are needed.

3.4 Routing

In wireless sensor networks it is not possible to transmit messages directly (i.e., in one hop) from one node in the network to another. Therefore, it is necessary to provide a routing infrastructure. Designing routing algorithms is a challenging area [19]. All the nodes inside the sensor network should be reachable (*connectivity*) while covering the maximum possible area of environment using their sensors (*coverage*), even when the nodes inside the network start to fail due to energy shortage or other problems (*fault tolerance*). The algorithm should also work with any network size and node density (*scalability*) and provide a certain quality of service. At the same time, designers must try to lower the memory usage and energy consumption of the algorithms.

Security is another factor that cannot be ignored in the design of routing algorithms. Any potential adversary has a wide range of attacks at his disposition [20, 21] to manipulate the routing subsystem and take control over the routes, resulting in eavesdropped, altered, spoofed or discarded packets. The key infrastructure may help in the defense against routing attacks by authenticating nodes and protecting the confidentiality and integrity of the packets, but it is not enough to protect the whole routing infrastructure. Therefore, it is essential to make the routing algorithm robust against such attacks.

In the literature, there has been some work that protects previously existent routing protocols such as directed diffusion [22]. Another branch of research focused on discovering new protection techniques. For example, in [23] the route discovery algorithm creates redundant paths from the nodes to the base station, taking advantage of the network density. Also, in [24] nodes equipped with physical location services can locate and map the routing holes in the network. But most of the existent routing protocols do not take security into account in any step of their design. As a conclusion, the main challenge in this area is to discover new protection techniques and apply them into new algorithms, without sacrificing primary design factors such as connectivity, coverage, scalability, etc.

3.5 Data Aggregation

Inside a sensor network, the nodes generate an immense amount of raw data product of their measurements. In most cases these data are needed at the base station, thus there is a great cost, in terms of energy consumption and bandwidth usage, on transporting all the data from the nodes to the base station. However, since nodes are physically near each other, the data likely have some type of redundancy. The role of aggregation is to exploit this redundancy by collecting the data from a certain region and summarizing it into one report, hence decreasing the number of packets sent to the base station.

Aggregated data can be easily attacked by a malicious adversary, even if the communications are protected against any data injection attack or data integrity attack. If an aggregator node is being controlled by an adversary, it can easily ignore the data received from its neighbors and create a false report. Trusted aggregators can still receive false data from faulty nodes or from nodes being controlled by an adversary.

By using strong aggregation functions that are resilient against internal attacks, it is possible to defend the network against false data coming from malicious or faulty nodes. As an example, the author in [25] developed a theoretical framework for analyzing the resilience of a number of natural aggregation functions borrowing ideas from the field of robust statistics, although the aggregation is supposed to be carried out in the base station.

There are also solutions that discover whether the reports sent by a malicious aggregator are forged or not. In one approach [26] the aggregator must create a proof of its neighbors' data (e.g. using a Merkle hash tree), which will be used in a negotiation with the base station to demonstrate the authenticity of the data used to construct the report. Other approaches [27] take advantage of the density of sensor networks by using the nodes in the neighborhood of the aggregator as witnesses. Finally, it is also possible to filter the packets containing the report and the proofs in their way to the base station, hence decreasing the amount of traffic created by false aggregations (e.g. by using a Bloom filter [28]).

The field of secure aggregation has still room for more improvements. Interactive protocols between aggregators and the base station require more traffic for the negotiation, introduce a delay in the aggregation service, and are not scalable without an aggregation testing hierarchy. Proof-based systems usually require a negotiation between the aggregator node and its witnesses and increase the size

of the reports sent to the base station. New solutions should try to minimize the amount of negotiations carried out by these algorithms, and to introduce new ways to early detect and eliminate false reports.

3.6 Auditory

In a sensor network the user will have, in most cases, only access to the base station or to the data dissemination subsystem connected to the base station. As a result, any change of the internal state of a node in the network, such as low energy level or hardware failure, will go unnoticed unless the node itself reports to the base station. Therefore, it could be interesting to provide an auditory subsystem inside the network to query about its internal status and to receive information about internal events.

A possible application of that auditory subsystem could be an *intrusion detection system* (IDS). IDS can monitor the activities of a network, gathering and analyzing audit data, in order to detect intrusions and alert users when an attack is taking place. IDS is in fact a "second line of defense" - if a malicious adversary takes control of certain parts of a system, IDS is able to detect it and activate countermeasures.

An IDS architecture for sensor networks could take advantage from concepts and techniques of IDS schemes employed in ad hoc networks [29]. However, those IDS techniques cannot be applied directly to sensor networks due to their unique features. Every sensor node cannot have a full-powered IDS agent, because of its high constraints in terms of battery life and processing power. Besides, since the density of sensor networks is usually high, it is also redundant and a waste of resources to force every node to analyze all the packets from its neighborhood. Therefore, the most basic problem that an IDS must face is how to distribute the detection tasks over the nodes.

There are other challenging problems to be solved in the field of IDS over sensor networks as well. An IDS architecture must be simple and highly specialized, able to analyze the specific protocols used over the network and react against specific sensor network threats. The set of rules used by the IDS algorithms for detecting rogue nodes must be easy to parse, their results must consume little memory space, and there must be some policy to manage those results when the memory is full. The alerts generated by the IDS infrastructure should reach the base station as soon as possible, no matter where they were generated. Finally, the IDS agents located inside the network should be able to interchange information in order to achieve a better detection performance.

As a side note, there are partial solutions in the literature that are able to check the integrity of the nodes belonging to the network, such as health monitoring [30], sensor readings analysis [31], and code attestation techniques [32]. These solutions could be integrated into an IDS system for improving its effectiveness.

3.7 Other Issues

A sensor network needs a secure infrastructure to protect itself from external or internal attacks targeting the confidentiality, integrity and authentication

properties of its communication channels. However, this is not enough for certain scenarios. There are extra properties that some networks must comply with (e.g., privacy) and there are some applications whose security requirements over a constrained scenario are still unknown (e.g., agents).

Privacy, in certain situations such as a battlefield, is an essential property. There are three types of privacy threats [33]. If an adversary can determine the meaning of a communication exchange because of the existence of a message and the context of the situation, there is a *content privacy threat*. If an adversary is able to deduce the identities of the nodes involved in a communication, there is an *identity privacy threat*. And if the adversary is able to infer the physical location of a communication entity or to approximate the relative distance to that entity, there is a *location privacy threat*.

There are some preliminary studies on originator location privacy and content privacy [33] that explore the privacy of some existing routing protocols. But in general, privacy over sensor networks is an unexplored area of research. It is important to discover and explore the scenarios where a privacy threat exists, and investigate and develop new solutions to solve these problems.

Since the sensor networks are still in their infancy, there are some applications whose security is not yet fully investigated. An example is the area of mobile agents [34], which provide an interesting tool for collaborative processing. However, any adversary could be able to inject a malicious agent inside a node or to modify the results being collected by the agent. Therefore, researchers should investigate how to provide secure code and secure results inside a highly constrained environment.

4 Conclusion

Security in wireless sensor networks is a field of research that is growing rapidly and achieving tangible results applicable to real-life scenarios. Nevertheless, there is still room for more improvements in this area. Fields such as public key cryptography and intrusion detection systems over sensor networks are fairly new. It is necessary to develop secure routing algorithms while complying with essential design properties, such as connectivity, coverage and fault tolerance. Also, secure data aggregation algorithms should be more optimal, and the privacy of the information flow should be taken into account.

Other open areas of research [35] include tolerating the lack of physical security, optimizing the security infrastructures in terms of resources (energy and computation), detecting and reacting over denial of service attacks, and raising the question of the social privacy problems that sensor networks might create. Finally there are some areas, such as the management and protection of mobile nodes and base stations, and the secure administration of multiple base stations with delegation of privileges, that are yet developed.

References

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci. *Wireless sensor networks: a survey.* Computer Networks, 38(4), March 2002.

2. Crossbow Technology, Inc. *MICA2 and MICAz, Wireless Measurement Systems.* http://www.xbow.com.
3. IEEE Standard, 802.15.4-2003. *Wireless medium access control and physical layer specifications for low-rate wireless personal area networks.* May 2003, ISBN 0-7381-3677-5.
4. C. Karlof, N. Sastry, D. Wagner. *TinySec: a link layer security architecture for wireless sensor networks.* In Proceedings of 2nd International Conference on Embedded Networked Sensor Systems (SensSys'04), November 2004.
5. N. Sastry, D. Wagner. *Security considerations for IEEE 802.15.4 networks.* In Proceedings of 2004 ACM Workshop on Wireless security (Wise'04), October 2004.
6. N. Gura, A. Patel, A. Wander, H. Eberle, S. C. Shantz. *Comparing elliptic curve cryptography and RSA on 8-bit CPUs.* In Proceedings of 2004 Workshop on Cryptographic Hardware and Embedded Systems (CHES'04), August 2004.
7. D. J. Malan, M. Welsh, M. D. Smith. *A public-key infrastructure for key distribution in TinyOS based on elliptic curve cryptography.* In Proceedings of 1st IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (Secon'04), October 2004.
8. L. Eschenauer, V. D. Gligor. *A key-management scheme for distributed sensor networks.* In Proceedings of 9th ACM Conference on Computer and Communications Security (CCS'02), November 2002.
9. H. Chan, A. Perrig, D. Song. *Random key predistribution schemes for sensor networks.* In Proceedings of 2003 IEEE Symposium on Security and Privacy (S&P'03), May 2003.
10. W. Du, J. Deng, Y. S. Han, P. K. Varshney. *A pairwise key pre-distribution scheme for wireless sensor networks.* In Proceedings of 10th ACM conference on Computer and communications Security (CCS'03), October 2003.
11. J. Lee, D. R. Stinson. *Deterministic key predistribution schemes for distributed sensor networks.* In Proceedings of 11th Annual Workshop on Selected Areas in Cryptography (SAC'04), August 2004.
12. R. Wei, J. Wu. *Product construction of key distribution schemes for sensor networks.* In Proceedings of 11th Annual Workshop on Selected Areas in Cryptography (SAC'04), August 2004.
13. B. Yener, S. A. Camtepe. *Combinatorial design of key distribution mechanisms for wireless sensor networks.* In Proceedings of 9th European Symposium on Research in Computer Security (ESORICS'04), September 2004.
14. W. Du, J. Deng, Y. S. Han, S. Chen, P. K. Varshney. *A key management scheme for wireless sensor networks using deployment knowledge.* In Proceedings of IEEE INFOCOM'04, March 2004.
15. A. Perrig, R. Szewczyk, V. Wen, D. Cullar, J. D. Tygar. *SPINS: Security protocols for sensor networks.* In Proceedings of 7th International Conference on Mobile Computing and Networking (MOBICOM'01), July 2001.
16. R. Anderson, H. Chan, A. Perrig. *Key infection: smart trust for smart dust.* In Proceedings of 12th IEEE International Conference on Network Protocols (ICNP'04), October 2004.
17. J. Zachari. *A decentralized approach to secure group membership testing in distributed sensor networks.* In Proceedings of 2003 Military Communications Conference (MILCOM 2003), October 2003.
18. Y. W. Law, R. Corin, S. Etalle, P. H. Hartel. *A formally verified decentralized key management architecture for wireless sensor networks.* In Proceedings of 2003 Personal Wireless Communications (PWC'03), IFIP WG 6.8 - Mobile and Wireless Communications. September 2003.

19. J. N. Al-Karaki, A. E. Kamal. *Routing techniques in wireless sensor networks: a survey.* To appear in IEEE Wireless Communications.
20. C. Karlof, D. Wagner. *Secure routing in wireless sensor networks: attacks and countermeasures.* In Proceedings of 1st IEEE International Workshop on Sensor Network Protocols and Applications, May 2003.
21. J. Newsome, E. Shi, D. Song, A Perrig. *The sybil attack in sensor networks: analysis & defenses.* In Proceedings of 3rd IEEE International Workshop on Information Processing in Sensor Networks (IPSN'04), April 2004.
22. R. Di Pietro, L. V. Mancini, Y. W. Law, S. Etalle, P. Havinga. *LKHW: A directed diffusion-based secure multicast scheme for wireless sensor networks.* In Proceedings of 32nd International Conference on Parallel Processing Workshops (ICPP'03), October 2003.
23. J. Deng, R. Han, S. Mishra. *A performance evaluation of intrusion-tolerant routing in wireless sensor networks.* In Proceedings of 2nd IEEE International Workshop on Information Processing in Sensor Networks (IPSN'03), April 2003.
24. Q. Fang, J. Gao, L. J. Guibas. *Locating and bypassing routing holes in sensor networks.* In Proceedings of IEEE INFOCOM'04, March 2004.
25. D. Wagner. *Resilient aggregation in sensor networks.* In Proceedings of 2nd ACM workshop on Security of Ad Hoc and Sensor Networks (SANS'04), October 2004.
26. B. Przydatek, D. Song, A. Perrig. *SIA: Secure information aggregation in sensor networks.* In Proceedings of 1st International Conference on Embedded Networked Sensor Systems (SenSys'03), November 2003.
27. W. Du, J. Deng, Y. S. Han, P. K. Varshney. *A witness-based approach for data fusion assurance in wireless sensor networks.* In Proceedings of GLOBECOM'03, December 2003.
28. F. Ye, H. Luo, S. Lu, L. Zhang. *Statistical en-route filtering of injected false data in sensor networks.* In Proceedings of IEEE INFOCOM'04, March 2004.
29. Y. Zhang, W. Lee. *Intrusion detection techniques for mobile wireless networks.* ACM/Kluwer Wireless Networks Journal, 9(5):545-556, September 2003.
30. C. Hsin, M. Liu. *A Distributed monitoring mechanism for wireless sensor networks.* In Proceedings of 2002 ACM Workshop on Wireless Security (WiSe'02), September 2002.
31. S. S. Doumit, D. P. Agrawal. *Self-organized criticaly & stochastic learning based intrusion detection system for wireless sensor networks.* In Proceedings of 2003 Military Communications Conference (MILCOM'03), October 2003.
32. A. Seshandri, A. Perrig, L. Van Doorn, P. Khosla. *SWATT: software-based attestation for embedded devices.* In Proceedings of 2004 IEEE Symposium on Security and Privacy (S&P'04), May 2004.
33. C. Ozturk, Y. Zhang, W. Trappe, M. Ott. *Source-location privacy for networks of energy-constrained sensors.* In Proceedings of 2nd IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems (WSTFEUS'04), May 2004.
34. H. Qi, Y. Xu, X. Wang. *Mobile-agent-based collaborative signal and information processing in sensor networks.* Proceedings of the IEEE, 91(8)1172-1183, August 2003.
35. A. Perrig, J. Stankovic, D. Wagner. *Security in wireless sensor networks.* Communications of the ACM, 47(6)53-57, June 2004.

Dependable Transaction for Electronic Commerce

Hao Wang[1], Heqing Guo[1], Manshan Lin[1], Jianfei Yin[1], Qi He[2], and Jun Zhang[2]

[1] School of Computer Science & Engineering, South China University of Technology, Guangzhou, China 510640
guozhou@scut.edu.cn
{iswanghao, lmshill, yjhhome}@hotmail.com
[2] Computer Engineering School, Nanyang Technological University, Singapore 639798
{qihe0001, jzhang}@ntu.edu.sg

Abstract. Electronic transaction becomes common practice in real world business. This paper focuses on the issue of dependability in critical transactions like electronic payment, electronic contract signing. Recent fair protocols can recover transactions from network crashes, but cannot survive local system crashes. A two-party dependable transaction protocol is proposed. During the protocol, both parties can recover the transaction from network and local system failures in a transparent way, which means that after the recovery, outcome messages would be just the same as those from a successful run of the transaction.

1 Introduction

Electronic transaction becomes common practice in real world business. When the transaction between organizations is executed on network, they may face risks of broken fairness in case of network failures, local systems failures [3], cheating behavior of either involved organization, and so on. So it is very important for them to follow some kind of transaction protocol assuring *dependability*. Dependability assures fairness for involved parties and recoverability from failures. Fairness means that when the electronic transaction terminates, either both parties get their expected items, or neither does. A Trusted Third Party (TTP) is involved as Pagnia and Garner [6] have proved that no definite fairness can be achieved without a TTP.

We first set up the application scenario for our transaction protocol: company B (the client, denoted as Bob) is going to buy some electronic goods from company A (the merchant, denoted as Alice) and they have settled on the goods and the price. Now they need to finish the exchange of Bob's check with Alice's goods. Bob's check is composed of his bank-certified account information, goods information and can be validated only after signed by his signature. With that signed check, Alice can get her money paid from Bob's bank.

1.1 Related Work

1.1.1 Fair Protocol Capable of Recovery from Network Crashes

In 1996, Asokan et al. [1] introduces the idea of optimistic approach and presents fair protocols with offline TTP, in which TTP intervenes only when an error occurs (net-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 691 – 700, 2005.

work error or malicious party's cheating). But the recovered messages are different from those produced by the sender or the recipient, which make the protocols suffer from bad publicity and weak fairness, as the recovered messages may lose some functionalities of the original ones. *Invisible TTP* is first introduced by Micali [5] to solve this problem. The TTP can generate exactly the same evidences as the sender or the recipient. In this way, judging the outcome evidences and received items cannot decide whether the TTP has been involved, so that the recovery is done in a transparent way.

Using *convertible signatures* (CS) is the recently focused approach to realize transparent recovery. It means to firstly send a partial committed signature that can be converted into a full signature (that is a normal signature) by both the TTP and the signer. Recently, Park et al. [7] present a very efficient protocol in which the output evidences are standard RSA signatures and the partial signature is non-interactively verifiable. But very soon, Dodis and Reyzin [2] break the scheme by proving the TTP can obtain Alice's entire secret key with only her registration information. In the same paper, they propose a new CS scheme (*DR signature scheme*) to produce an equally efficient but more secure protocol.

But all these protocols have not considered cases of systems crashes and assumed that local systems of Alice, Bob and TTP are all stable.

1.1.2 Recovery Methods for Local Systems Crashes

Liu et al. [3] have proposed the *Semantics-based Message Logging* (SbML method) to enable recovery of local systems crashes. The SbML is a logging method balanced between pessimistic logging (log all messages before sending out) and optimistic logging (message processing and logging is separated). Involved parties can define their critical points (called *point-of-no-return*) in the protocol run and message will be logged before they enter the defined points.

This logging method works in protocols with online TTP. But when it comes to offline TTP and invisible TTP, fairness after crashes can be potentially broken. Cases of broken fairness are as following:

Case 1.1: after Alice sends out the first message, her system crashes; when Bob get the message, he can invoke the recover sub-protocol to get the final expected messages; if Alice fails to recover her system before TTP's recovered messages arrive, her fairness will be broken. So simply using their logging method is not enough to guarantee fairness.

Case 1.2: the offline TTP has not logged the variables: *recovered* and *aborted*, if TTP crashes after a successful abort operation requested by Alice; at this time, Alice has quitted the transaction since her request has been confirmed; but if Bob submit a recover request after TTP recovers, TTP will recover the transaction and send proper recovered messages to Alice and Bob; in this case, the message cannot arrive Alice, so fairness for Alice is broken.

1.2 Our Work

In this paper we first define the property of *Dependability* of transaction protocol. Then we present a transaction protocol based on DR signature scheme. To enable transparent recovery of crashes of network and local systems, we adapt the Seman-

tics-based Message Logging method and introduce a new inquiry sub-protocol. Finally we prove that the transaction protocol is dependable.

The remainder of the paper is structured as follows. In Section 2, we define the dependability property of electronic transactions. Section 3 presents the transaction protocol in payment scenario. Section 4 analyzes the protocol in details. Some concluding remarks are given in Section 5.

2 Dependability of Electronic Transactions

Markowitch et al.[4] study many former fairness definitions and present a well-knitted definition. Recently, Wang and Guo [20] present a set of new requirements for fair protocols with invisible TTP. Based on that, we extract 5 properties of transaction protocols and we say a protocol is dependable if it satisfies all these properties.

Definition 1. Effectiveness
A transaction protocol is *effective* if there exists a successful exchange of both parties' expected items.

Definition 2. Fairness
A transaction protocol is *fair* if when the protocol run ends, either both parties get their expected items or neither of them gets anything useful.

Definition 3. Timeliness
A transaction protocol is *timely* if the protocol can be completed in a finite amount of time while preserving fairness for both exchangers.

Definition 4. Non-repudiability
A transaction protocol is *non-repudiable* if when the exchange succeeds, either payer or payee cannot deny (partially or totally) his/her participation.

Definition 5. Transparent recoverability
A transaction protocol is *transparent recoverable* if after a successful exchange, the result evidences of origin/receipt and exchanged items are indistinguishable in respect to whether TTP has been involved.

With all these properties' definitions, we can define the dependability as following:

Definition 6. Dependability
A transaction protocol is *dependable* if it assures effectiveness, fairness, timeliness, non-repudiability and transparent recoverability.

3 A Dependable Payment Protocol with Transparent Recovery

In this section, we present a dependable protocol in the payment scenario described in Section 1. The protocol uses the DR signature as an important cryptographic tool. So we first briefly describe this signature scheme. Then with assumptions clearly presented, all the five parts of the protocol is described in details.

3.1 Dodis-Reyzin Convertible Signature Scheme

The DR signature is based on a recent widely-used RSA-like signature scheme called *gap Diffie-Hellman (GDH) signature* and the corresponding *GDH groups* (see [2] section 4 for detailed description).

GDH Signature. Assume G is a multiplicative group of prime order p. Key generation algorithm of the GDH signature scheme picks a GDH group of order p, and random $g \in G, x \in Z_p$. It computes $h = g^x$, and set the public key to be (g, h) (G, p is public accessible), and the secret key to be x. To sign a message m, one computes $\sigma = H(m)^x$, where $H(m)$ is a random oracle. To verify σ, one outputs $V_{DDH}(g, h, H(m), \sigma)$, that is, test if $\log_g h = \log_{H(m)} \sigma$ (outputting 1 means being equal). One can easily find that a secure zero-knowledge proof can accomplish this test.

DR Signature. This CS signature scheme contains one register procedure and several signing/verifying algorithms.

Register Procedure. Signer (say Alice) chooses random $g \in G, x, x_1 \in Z_p$, computes $x_2 = x - x_1 \bmod p, h = g^x, h_1 = g^{x_1}$, and sets her public key $pk = (g, h)$, secret key $sk = (x, x_1)$, partial public key $ppk = h_1$, partial secret key $psk = x_2$, then she sends the pk, ppk, psk to the TTP, the TTP will check whether $h = h_1 g^{x_2}$ so that it can finish the signature conversion.

Signing/Verifying Algorithms of Full Signature. They are just the signing/verifying algorithms of normal GDH signature: $FS(m) = \sigma = H(m)^x$, $Ver(m, \sigma) = V_{DDH}(g, h, H(m), \sigma)$.

Signing/Verifying Algorithms of Partial Signature. Similar with former ones but using the public key h_1: $PS(m) = \sigma' = H(m)^{x_1}$, $PVer(m, \sigma') = V_{DDH}(g, h_1, H(m), \sigma')$.

Converting Algorithm. The TTP run this algorithm $Convert(m, \sigma')$ to convert $PS(m)$ to $FS(m)$: it will first check whether $PVer(m, \sigma') = 1$, if holds, it outputs $FS(m) = \sigma' H(m)^{x_2}$.

Dodis and Reyzin have proved the DR signature scheme is just as secure as the normal GDH signature scheme ([2] Theorem 3).

3.2 The Protocol

Based on the application scenario set in Section 1, we first state our protocol's assumptions as following:

Communication Network. We assume the communication channel between Alice and Bob is unreliable and channels between exchangers (Alice/Bob) and TTP are resilient. Messages in a resilient channel can be delayed but will eventually arrive. On the contrary, messages in unreliable network may be lost.

Cryptographic Tools. Encryption tools including symmetric encryption, asymmetric encryption and normal signature is secure. In addition, the adopted signature scheme is message recovery.

Honest TTP. The TTP should send a valid and honest reply to every request. Honest means that when the TTP is involved, if a recover decision is made, Alice gets the payment and Bob gets the goods; if a abort decision is made, Alice and Bob get the abort confirmation and they cannot recover the exchange in any future time.

Local Systems. Local systems of Alice, Bob and TTP are recoverable with proper message logging including logging before *point-of-no-return* [3].

To describe the protocol, we need to use several notations concerning the necessary cryptographic tools:

- $E_k()/D_k()$: a symmetric-key encryption/decryption function under key k
- $E_X()/D_X()$: a public-key encryption/decryption function under pk_X
- $S_X()$: ordinary signature function of X
- k: the key used to cipher goods
- pk_X/sk_X: public/secret key of X
- $cipher = E_k(goods)$: the cipher of goods under k
- $X \rightarrow Y$: transmission from entity X to Y
- $h()$: a collision resistant one-way hash function
- *goods*: goods destined to B
- *check*: the check destined for A, it contains transaction identity, goods identity, price information, B's account information, etc
- l: a label that uniquely identifies a protocol run
- f: a flag indicating the purpose of a message

Registration Sub-protocol. To participate in a payment protocol, both Alice and Bob need to run the register procedure with the TTP as required by DR signature. Note that it will not affect the security if they share a same g. Bob also need to send his check for the TTP to verify its validity.

Main Protocol. After Alice and Bob settle the price and the goods, they can follow the main protocol. Note that they both make their own messages logged on stable storage before run the protocol:

Step 1, Alice sends encrypted goods (*cipher*) with the key k encrypted by the TTP's public key ($E_{TTP}(k)$), her partial signature on them (a=*cipher*, $E_{TTP}(k)$, PS_A(a)= σ_A') to initiate the payment process.

Step 2, if Bob decides to give up or he doesn't receive Alice's message in time, he can simply quit and retain fairness. When he receives the message, he will first run $PVer(a, \sigma_A')$, if it equals 1, he will send his *check* and his partial signature on it ($PS_B(check)$= σ_B') to Alice. Otherwise, he quits the protocol.

Step 3, if Alice decides to give up or she doesn't receive Bob's message in time, she can invoke the *abort* sub-protocol to prevent a later resolution by the TTP. When she receive the message, she will first run $PVer(check, \sigma_B')$, if it equals 1, she will log the message and the state information, then send k and her full signature on a (FS_A(a)= σ_A) to Bob. Otherwise, she also invokes the *abort* sub-protocol.

Step 4, if Bob detects that his channel with Alice is broken or doesn't receive the message in time, he can invoke the *recover* sub-protocol. When he receive the mes-

sage, he will check whether k can decrypt the *cipher* and the *goods* is satisfactory, also he will run $Ver(a, \sigma_A)$, if all these checking pass, he will log the message and the state information, then send his *check* and his full signature on it ($FS_A(check) = \sigma_B$) to Alice. Otherwise, he will invoke the *recover* sub-protocol.

Step 5, if Alice detects that her channel with Bob is broken or doesn't receive the message in time, she can invoke the *recover* sub-protocol. When she receives the message, she will run $Ver(check, \sigma_B)$, if it equals 1, she will accept the check. Otherwise, she will invoke the *recover* sub-protocol.

Main Protocol

A: $log(B, l, a, cipher, k)$
B: $log(A, l, check)$
$A \rightarrow B$: $f_{\mathbf{EOO}}, B, l, h(k), cipher, E_{TTP}(l, k), PS_A(a)$
B: **if not** $Ver(a, PS_A(a))$ **then** stop
else $log(A, l, h(k), cipher, E_{TTP}(l, k), PS_A(a))$
$B \rightarrow A$: $f_{\mathbf{EOR}}, A, l, PS_B(b)$
A: **if** times out **then abort**
elseif not $Ver(b, PS_B(b))$ **then abort**
else $log(B, l, PS_B(b))$
$A \rightarrow B$: $f_{\mathbf{NRO}}, B, l, k, FS_A(a)$
B: **if** times out **then** call **recover**[*X*:=*B*,*Y*:=*A*]
else $log(A, l, k, FS_A(a))$
$B \rightarrow A$: $f_{\mathbf{NRR}}, A, l, FS_B(b)$
A: **if** *A* times out **then** call **recover**[*X*:=*A*,*Y*:=*B*]

Recover Sub-protocol. Whenever necessary, Alice/Bob (noted by *X*) will invoke the *recover* protocol to let the TTP decide whether finish or abort the payment process.

Step 1, *X* sends to the TTP $E_{TTP}(k)$, $PS_A(a) = \sigma_A'$, *check*, $PS_B(check) = \sigma_B'$ to initiate a recover process. Because of the resilient channel between *X* and the TTP, this message will eventually arrives the TTP.

Step 2, when the TTP receive the message, it will first check whether the protocol has already been recovered or aborted, if so, it will stop because it is sure that both parties have got the recovered items or the abort confirmation. Then it will decrypt $E_{TTP}(k)$ with its secret key sk_{TTP}, if succeeds, it will run $PVer(a, \sigma_A')$ and $PVer(check, \sigma_B')$. If both equals 1, the TTP will run $Convert(a, \sigma_A')$ and $Convert(check, \sigma_B')$. After all these operations succeed, TTP will log the message and the variable *recovered*, then send the $FS_A(check) = \sigma_B$ to Alice and $FS_A(a) = \sigma_A$ & k to Bob. If either checking fails, it will abort the protocol and send confirmations to Alice and Bob.

Recover Sub-protocol

$X \rightarrow TTP$: $f_{\mathrm{RecX}}, Y, l, h(cipher), h(k), E_{TTP}(k), PS_A(a), PS_B(b)$
TTP: $log(f_{\mathrm{RecX}}, A, B, l, h(cipher), h(k), E_{TTP}(k), PS_A(a), PS_B(b))$
if $h(k) \neq h(D_{TTP}(E_{TTP}(k)))$ **or** *aborted* **or** *recovered* **then** stop
else if $PVer(a, PS_A(a)) \neq 1$ or $PVer(a, PS_A(a)) \neq 1$ **then** stop

else *recovered*=true
$Convert(PS_A(a), x_{2A})$ and $Convert(PS_B(b), x_{2B})$
$log(A, B, l, recovered, FS_A(a), k, FS_B(b))$
$TTP \rightarrow A$: $f_{\mathbf{NRR}}$, A, l, $FS_A(a)$
$TTP \rightarrow B$: $f_{\mathbf{NRO}}$, B, l, k, $FS_B(b)$

Inquiry Sub-protocol. After recovering from local system crashes, Alice/Bob (denoted as *X*) can invoke the *inquiry* sub-protocol to check the current status of the transaction and get what s/he deserves.

Step 1, *X* sends an inquiry request to the TTP. Because of the resilient channel between *X* and the TTP, this message will eventually arrives the TTP.
Step 2, on the inquiry request, TTP will check the current status of the protocol according to the label *l*. If no record is available, that means that protocol has not been submitted to TTP and *X* can directly recover the protocol run with *Y*. So TTP will just need to return a *null* message to X. If the protocol has been recovered, TTP will send the recovered message to *X*, that is, FS_A(a), *k* (for Bob) or $FS_A(check)$ (for Alice). If the protocol has been aborted, TTP will send the abort confirmation to *X*.

Inquiry Sub-protocol
$X \rightarrow TTP$: $f_{\mathrm{Inq}X}$, $\mathbf{Inq_X}$
TTP: **if** *aborted* **then**
$TTP \rightarrow X$: $f_{\mathbf{Con}a}$, A, B, l, $\mathbf{Con}_a$
TTP: **elseif** *recovered* **then**
if *X*=*A* then
$TTP \rightarrow A$: $f_{\mathbf{NRR}}$, A, l, $FS_A(a)$
else
$TTP \rightarrow B$: $f_{\mathbf{NRO}}$, B, l, k, $FS_B(b)$
else
$TTP \rightarrow X$: **null**

Abort Sub-protocol. In step 2 of the main protocol, Alice can invoke this sub-protocol to make the TTP abort this payment protocol run.

Step 1, Alice sends an abort request to the TTP. Because of the resilient channel between *X* and the TTP, this message will eventually arrives the TTP.
Step 2, if the protocol has not been recovered or aborted, the TTP will abort the protocol and log the message and the variable *aborted*, then send confirmations ($\mathbf{Con}_a$) to both parties.

Abort Sub-protocol
$X \rightarrow TTP$: $f_{\mathbf{Abort}}$, l, B, **abort**
TTP: **if** *aborted* **or** *recovered* **then** stop
else *aborted*=true
$log(A, B, l, aborted)$
$TTP \rightarrow A$: $f_{\mathbf{Con}a}$, A, B, l, $\mathbf{Con}_a$
$TTP \rightarrow B$: $f_{\mathbf{Con}a}$, A, B, l, $\mathbf{Con}_a$

4 Analysis of the Protocol

Following is the analysis with respect to the dependability definition in Section 2.

Claim 1. *Assuming the channel between Alice and Bob is unreliable and adopted cryptographic tools are secure, the protocol satisfies the effectiveness requirement.*

Proof: When both Alice and Bob are honest, thus they will follow the protocol to send messages. If the probability of successful transmission in the unreliable channel is δ, then the probability of successful execution of one main protocol run will roughly be δ^4. Even it's small, but it means successful execution without TTP's involvement is still possible. Thus the protocol satisfies the effectiveness requirement.

Claim 2. *Assuming the channels between the TTP and Alice/Bob are resilient, adopted cryptographic tools are secure and the TTP is honest, the protocol satisfies the fairness requirement.*

Proof: The fairness can be proved considering 3 aspects: fairness for Alice, fairness for Bob and recovered fairness after TTP crashes.

- *Fairness for Alice* Assuming Alice is honest, then risks she may faces include:

1) She did not receive any message or the message is invalid in step 3. She can request abort to prevent that Bob may call a recovery later. If Bob's recovery request arrives to the TTP before her abort request, the TTP still will send the recovered goods and evidence to her. Thus will not affect her benefit.
2) She did not receive any message or the message is invalid in step 5. She can submit a recovery request, because the TTP is honest, the exchange will be forced to complete. If Bob sent a recovery request during this period, the result will be the same; if Bob sent an abort request which arrived before Alice's recovery request, the exchange will be aborted by the TTP, and no party can gain advantage.
3) Local system crashes. After Alice recovers from local system crash, she can instantly invoke inquiry sub-protocol to check the current status; if she has submitted abort or recover request before her crash, she will get proper messages (abort confirmation or recovered messages) from TTP; if Bob has submitted recover request before or during her crash, she will get recovered messages from TTP; if no involvement before or during her crash, she can simply contact Bob to continue the transaction. So her fairness is assured.

- *Fairness for Bob* Assuming Bob is honest, then risks he may faces include:

1) He did not receive any message or the message is invalid in step 2. He can simply stop without any risk. And at this time, Alice cannot call recovery.
2) He did not receive any message or the message is invalid in step 4. He can request recovery and the exchange will be forced to complete. If Alice request recovery at the same time, the result will be the same.
3) Local system crashes. After Bob recovers from local system crash, he can instantly invoke inquiry sub-protocol to check the current status; if he has submitted recover request before his crash, he will get recovered messages from TTP;

if Alice has submitted abort or recover request before or during his crash, he will get proper messages (abort confirmation or recovered messages) from TTP; if no involvement before or during his crash, he can simply contact Alice to continue the transaction. So his fairness is assured.

- *Recovered fairness after TTP crashes* Cases of TTP crashes include:

1) Alice has submitted abort request before TTP crashes, and TTP has sent both parties the abort confirmation. Because TTP has logged request message and the variable *aborted*, so after TTP recovers the information about this protocol run, the TTP will deny any later recovery request by either Alice or Bob.
2) Alice/Bob has submitted recover request before TTP crashes, and TTP has sent both parties the recovered messages. Because TTP has logged the request message and the variable recovered, so after TTP recovers the information about this protocol run, the TTP can re-run the recovery operations (if necessary) and will ignore Alice's later abort request.
3) Alice/Bob has submitted abort/recover request during TTP crashes. Alice/Bob can re-submit request after TTP's recovery or TTP can actively broadcast the crashes information so that all requesting parties can re-submit their requests.

Claim 3. *Assuming the channels between the TTP and Alice/Bob are resilient, adopted cryptographic tools are secure and the TTP is honest, the protocol satisfies timeliness requirement.*

Proof: Alice can conclude the protocol in one of the two ways:

1) requesting abort before sending the message of step 3.
2) requesting recovery in any other time.

Bob can conclude the protocol in one of the three ways:

1) stopping at any time before sending the message of step 2.
2) requesting recovery in any other time.

With the channel assumption, the abort confirmation or the recovered information will arrive to both parties in a finite amount of time. And all these conclusions, as discussed in the proof of claim 2, will not hurt either party's interests. So the timeliness is guaranteed.

Claim 4. *Assuming the channels between the TTP and Alice/Bob are resilient, adopted cryptographic tools are secure, the TTP is honest, the protocol satisfies non-repudiation requirement.*

Proof: When the exchange succeeds, either by following the main protocol or recovered by the TTP (including recovered message after inquiry), Alice will get $FS_A(check)= \sigma_B$, and Bob will get $FS_A(a)= \sigma_A$ & k. So Alice can convince outside parties that Bob has received goods and claim her money from Bob's bank. Similarly, Bob can prove that Alice has sent goods.

Claim 5. *Assuming the channels between the TTP and Alice/Bob are resilient, adopted cryptographic tools are secure, the TTP is honest, the protocol guarantees transparent recoverability.*

Proof: Either the TTP is involved or not, the resulting message (FS_B(*check*), FS_A(a) and *k*) are just the same, so the protocol is transparent recoverable.

With all these claims, we can easily see that the protocol is dependable:

Theorem 1. *Assuming the channels between the TTP and Alice/Bob are resilient, adopted cryptographic tools are secure and the TTP is honest, the protocol is dependable.*

5 Conclusions

In this paper, we produce a dependable transaction protocol with transparent recoverability. We have shown that the protocol are practical as it has high recoverability and can survive relatively unreliable network. To be more precisely about effect of every factor in the protocol like network/system reliability, honesty of both parties and etc, we are building an agent-based platform for analysis and verification. Then we can see how dependable protocol can be applied in different environments.

References

1. N. Asokan, M. Schunter, and M. Waidner. Optimistic protocols for fair exchange. In *Proceedings of the fourh ACM Conference on Computer and Communications Security*, 1997.
2. Y. Dodis, L. Reyzin. Breaking and repairing optimistic fair exchange from PODC 2003. In *Proceedings of the 2003 ACM workshop on Digital rights management*, 2003.
3. P. Liu, P. Ning, and S. Jajodia. Avoiding loss of fairness owing to process crashes in fair data exchange protocols. In *Proceedings of the IEEE International Conference on Dependable Systems and Networks, Workshop on Dependability despite Malicious Faults*, 2000.
4. O. Markowitch, S. Kremer and D. Gollmann. On Fairness in Exchange Protocols. In *Proceedings of Information Security and Cryptology (ICISC 2002). LNCS 2587*, Springer-Verlag, 2002.
5. S. Micali. Certified e-mail with invisible post offices. Available from author: an invited presentation at the RSA'97 conference, 1997.
6. H. Pagnia and F. C. Gartner. On the impossibility of fair exchange without a trusted third party. *Tech. Rep. TUD-BS-1999-02 (March)*, Darmstadt University of Technology, 1999.
7. J. M. Park, E. K. P. Chong, H. J. Siegel. Constructing fair-exchange protocols for E-commerce via distributed computation of RSA signatures. *Proceedings of the twenty-second annual symposium on Principles of distributed computing*. 2003.
8. H. Wang and H. Guo. Fair Payment Protocols for E-Commerce. In *Proceedings of Fourth IFIP Conference on e-Commerce, e-Business, and e-Government (I3E'04)*. Building the E-Society: E-Commerce, E-Business and E-Government, Kluwer academic publishers, 2004.

On the Security of a Certified E-Mail Scheme with Temporal Authentication

Min-Hua Shao[1,*], Jianying Zhou[2], and Guilin Wang[2]

[1] Institute of Information Management, National Chiao Tung University
1001 Ta Hsueh Road, Hsinchu 300, Taiwan
mhshao@alumni.nccu.edu.tw
[2] Infocomm Security Department, Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613
{jyzhou, glwang}@i2r.a-star.edu.sg
http://www.i2r.a-star.edu.sg/icsd/

Abstract. Certified e-mail is a value-added service for standard e-mail systems, in which the intended recipient gets the mail content if and only if the mail originator receives a non-repudiation evidence that the message has been received by the recipient. As far as security is concerned, fairness is one of the most important requirements. Recently, Galdi and Giordano (2004) presented an optimistic protocol for certified e-mail with temporal authentication. In this paper, we analyze their protocol and demonstrate that it cannot achieve true fairness and has some other weaknesses. We further propose the improvements to avoid those security problems.

1 Introduction

The lack of evidence for message receipt is a missing piece of the infrastructure required for the more professional use of email [14]. Certified e-mail uses the notion of a signed receipt and strengthens the binding between the evidence and the mail being certified. In other words, the main purpose of a certified e-mail scheme is to achieve the fair exchange of a message and a receipt in the sense that either the sender obtains a receipt from the receiver and the receiver accesses the content of the e-mail simultaneously, or neither party gets the expected item. Although fairness is probably the most important one, there are other properties on the application of certified e-mail. The following security properties are defined in [12,13] and extended in [7].

- *Fairness*: The protocol should be *fair* in the sense that either each party receives the expected item or neither party receives any useful information about the other's item.
- *Non-repudiation*: Neither the sender nor the receiver of a message is able to deny the transmission.

* The author's work was done during her visit to Institute for Infocomm Research, Singapore, and funded by the National Science Council of Taiwan under the contract of NSC 93-2917-I-009-001.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 701 – 710, 2005.

- *Timeliness*: Both the sender and the receiver of a message should have the ability to reach the end of a protocol run in a finite amount of time unilaterally without losing fairness.
- *Authenticity*: The players should be guaranteed of their reciprocal identity.
- *Confidentiality*: None but the intended parties can get access to the plaintext items sent during the protocol.
- *Integrity*: Message transmission should be protected against unauthorized operations in order to guaranty the correctness and authenticity of data.
- *Temporal Authentication*: The sender can obtain the evidence to prove the time at which the message was sent.
- *Sending Receipt*: The sender can obtain the evidence to prove that he/she started the process of sending a certified e-mail.

Certified e-mail has been discussed for years, and there are two major classes of schemes to address the certified mail problem: schemes that require the existence of a trusted third party (TTP), and schemes that don't require the existence of a TTP. Oppliger showed clearly that the second class, i.e., either based on a simultaneous secret exchange or trusted system is inappropriate to provide certified mail services for the Internet [14]. Therefore, the use of TTPs seems advantageous and various types of TTPs can be considered according to their involvement in the certified e-mail protocol: schemes with *in-line* TTPs [3], schemes with *on-line* TTPs [1,5,14] and schemes with *off-line* TTPs [2,6,9].

An in-line TTP, i.e. acting as a delivery authority, involves in each message's transmission during the protocol. The main advantage of in-line TTPs for certified mail is to ensure strong fairness since the TTP collects all information necessary before forwarding them to the concerned entities; and further, the in-line TTP has full control over the message flows and likely provides the sender anonymity services. However, it also implies a communication and computation bottleneck due to the heavy involvement of the TTP.

An improvement to reduce the TTP's involvement is the use of an on-line TTP. The on-line TTP is actively involved during each session of the certified e-mail protocol but not during each message's transmission. Its task may only deal with signaling information, such as cryptographic keys and/or receipts sent back to the originator [14]. In academic literature, there is often an emphasis on reducing the role and the expense of a TTP. Protocols with a light-weight TTP have been proposed. For example, Abadi et al. proposed an efficient certified e-mail scheme with a light on-line TTP [1]. A key feature of their scheme is not to deploy any public-key infrastructure; and further, Imamoto and Sakurai [8] revised their scheme in order to provide the non-repudiation of origin service.

A big step towards more efficient solutions was the introduction of off-line TTPs. That is, an off-line TTP involves in a protocol only in case of an incorrect behavior of a dishonest entity (for example, the recipient claims having not received the message or the originator claims having not received the receipt), or in case of a network error. Considering most of the time no problem will occur, this approach using an off-line TTP is also called the *optimistic* approach.

Galdi and Giordano proposed an improved optimistic protocol for certified e-mail at TrustBus 2004 [7]. Their effort is to introduce a feature of "temporal authentica-

tion" into certified e-mail along a four-message optimistic protocol. Galdi-Giordano's certified e-mail scheme (GG scheme, for short) is effective against misbehavior of one of the players in some cases. However, we demonstrate in this paper that it suffers from a few severe security problems, and some of the security properties mentioned above cannot be satisfied. For example, the receiver can get the e-mail content by *replay attacks* even though he/she did not give the sender a receipt of the message. In this paper, we give a thorough security analysis of the GG scheme and further propose the improvements to avoid these problems.

The rest of this paper is organized as follows. We introduce the notation in Section 2, and briefly review the GG scheme in Section 3. We point out the vulnerabilities and propose solutions in Section 4. We end the paper with conclusions in Section 5.

2 Notation

In this paper, we use the same notation used in the original paper [7]. For completeness and readability, we summarize the model and all cryptographic symbols below.

–*Alice*, *Bob*, *Ted*, *Sam*: four different participating entities in which Alice is the message sender, Bob is the message receiver, Ted acts as an off-line *Trusted Third Party* (TTP), and Sam plays as an online *Trusted Stamping Server* (TSS).
–m_{subj}: the message subject associated with the message m.
–$PK_X(m)$: the encryption of the message m using the public key of the player X, where $X \in \{\mathrm{A}(lice),\mathrm{B}(ob),\mathrm{T}(ed),\mathrm{S}(am)\}$.
–$Sig_X(m)$: the signature of player X on message m.
– $\overline{PK}_x(m,r)$: the encryption of the message m, obtained by using the public key of the Player X and random string r.
–$X \rightarrow Y$: m: player X sends the message m to player Y.
–$x||y$: the concatenation of strings x and y.
–$h(\cdot)$: a collision resistant one-way hash function.

3 The GG Scheme

We first sketch the GG scheme proposed by Galdi and Giordano in [7]. In this scheme, the *basic protocol* is the core of the certified e-mail scheme that ensures timelines and message verifiability. It consists of three messages exchanged between the sender Alice and the receiver Bob in the normal situation. The *extension of the basic protocol* is provided that introduces an on-line time stamping server and add a single message due to the temporal authentication. In addition, the *recovery procedures* are launched in the abnormal situation to achieve fairness for all participants under the help of the TTP's involvement. We now review these three protocols in more detail below.

(1) The basic protocol. Assume that the sender Alice wants to deliver a message m to the receiver Bob with a guarantee that Bob can access the message m if and only if

Alice obtains a receipt from Bob. To this end, they run the following three-message optimistic protocol for fair exchange of certified e-mail.

(b1). Alice → Bob: m_1

where $m_1 = \langle env, Sig_A(env) \rangle$,

$env = \langle ID_A, ID_B, PK_B(\langle m_{subj}, h(\langle m, r \rangle) \rangle), \overline{PK}_T(PK_B(msg), r) \rangle$, $msg = \langle m_{subj}, m \rangle$

(b2). Bob → Alice: m_2

where $m_2 = Sig_B(m_1)$

(b3). Alice → Bob: m_3

where $m_3 = \langle \langle PK_B(msg), r \rangle, Sig_A(\langle PK_B(msg), r \rangle) \rangle$

Clearly, the key idea in the basic protocol is to use an electronic envelope to lock the message. In the first message flow (b1), m_1 is composed of two parts besides the players' identities. The hash value $h(\langle m, r \rangle)$ in the first part will be used to verify that the message received corresponds to the one for which the receipt has been sent. The second part that is a cipher of the actual content of the email by Ted's public key for recovery procedures. In the second message flow (b2), m_2 is the message receipt for *m* that ensures the non-repudiation of receipt for Alice. Finally, in the third message flow (b3), Bob must make sure that the receipt sent to Alice corresponds to the received message. That is, Bob needs to confirm the integrity between *m* in m_3 and *m* of $h(\langle m, r \rangle)$ in m_1. There are three items needed to verify: (1) m_{subj} received in the first message m_1 matches the one received in m_3; (2) the new hash value of $\langle m, r \rangle$ retrieved in m_3 is in correspondence with the one in m_1; (3) the ciphertext $\overline{PK}_T(PK_B(msg), r)$ drawn from m_1 and the new one produced from m_3 by using Ted's public key are the same.

(2) The extension protocol. Due to temporal authentication the time stamping server, Sam, is involved in the protocol to provide a time certification of the message *m*, notated as *t(m)*. More specifically, Sam sends a copy of the message m_2 to Alice during the second message flow (e2) in order to obtain a sender's receipt.

(e1). Alice → Sam: m_1

where $m_1 = \langle env, Sig_A(env) \rangle$,

$env = \langle ID_A, ID_B, PK_B(\langle m_{subj}, h(\langle m, r \rangle) \rangle), \overline{PK}_T(PK_B(msg), r) \rangle$, $msg = \langle m_{subj}, m \rangle$

(e2). Sam → Bob and/or Alice: m_2

where $m_2 = \langle \langle m_1, t(m_1) \rangle, Sig_S(\langle m_1, t(m_1) \rangle) \rangle$

(e3). Bob → Alice: m_3

where $m_3 = \langle m_2, Sig_B(m_2) \rangle$

(e4). Alice → Bob: m_4

where $m_4 = \langle \langle PK_B(msg), r \rangle, Sig_A(\langle PK_B(msg), r \rangle) \rangle$

(3) The recovery procedures. The recovery procedures will be launched by one of the players once the other player P misbehaves on the message m_i, notated as P failed on m_i. There are two failures discussed in the GG scheme: Alice failed on m_3 and Bob failed on m_2. Regarding Alice's failures, i.e., by sending m_3 to Bob with wrong information or by not sending m_3 to Bob, Bob can obtain the message for which he issued a receipt by the involvement of the TTP. The detailed treatment of the recovery procedures is shown below:

(r1). Bob → Ted: m_1, m_2
(r2). If (ID_A in m_1 = Alice and ID_B in m_1 = Bob) and
(verify Alice's signature $Sig_A(env)$ and Bob's signature $Sig_B(m_1)$)
then Ted → Bob: $enc = PK_B(\langle m_{subj}, m\rangle)$
Ted → Alice: m_2

As for Bob's failures, Bob can only fail to reply m_2 to Alice. In this case the following recovery procedures (r3) and (r4) are used for the basic protocol; others are used for the extension protocol.

(r3). Alice → Ted: m_1
(r4). Ted → Bob: m_1
Ted's verification:
If (no response from Bob)
then Ted → Alice: $rej = \langle\langle REJ, m_1\rangle, Sig_T(\langle REJ, m_1\rangle)\rangle$
If (m_2 from Bob) then Ted → Alice: m_2

(r5). Alice → Sam: m_1
(r6). Sam → Ted: $\langle m_s, t(m_s)\rangle$
where $m_s = \langle m_1, t(m_1)\rangle$
(r7). Ted → Bob: m_s
Ted's verification:
If (no response from Bob)
then Ted → Alice: $rej = \langle\langle REJ, m_t\rangle, Sig_T(\langle REJ, m_t\rangle)\rangle$
If (m_2 from Bob) then Ted → Alice: m_2

4 Security Analysis of the GG Scheme

4.1 Vulnerabilities

V-1. Replay Attack
Replay attack is one of active attacks that an adversary records a communication session and replays the entire session, or a portion thereof, at some later point in time. Unfortunately, the sender in the GG scheme may suffer severely from a replay attack. That is, the receiver can collude with some party Cindy to obtain the content of the

message from the TTP without providing a valid receipt to the original sender. Our attack scenario on the basic protocol of the GG scheme is illustrated in Figure 1. The detailed treatment of replay attack on Figure 1 is described below.

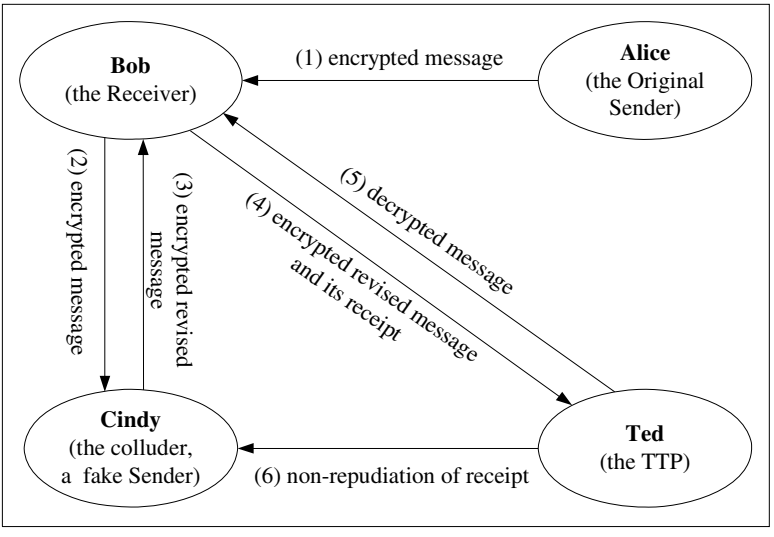


Fig. 1. Replay attack scenario on the basic protocol

(1). Alice → Bob: m_1

where $m_1 = \langle env, Sig_A(env) \rangle$,

$env = \langle ID_A, ID_B, PK_B(\langle m_{subj}, h(\langle m, r \rangle) \rangle), \overline{PK}_T(PK_B(msg), r) \rangle$, $msg = \langle m_{subj}, m \rangle$

(2-3). Bob colludes with Cindy. Cindy creates a revised version $\hat{m}_1$ of the message m_1 where ID_A is replaced with ID_C, and then generates her signature on the revised message env' in order to disguise as the sender.

Bob → Cindy: m_1

Cindy → Bob: $\hat{m}_1$

where $\hat{m}_1 = \langle env', \underline{Sig_C(env')} \rangle$,

$env' = \langle \underline{ID_C}, ID_B, PK_B(\langle m_{subj}, h(\langle m, r \rangle) \rangle), \overline{PK}_T(PK_B(msg), r) \rangle$

(4). Bob produces a false receipt $\hat{m}_2$ on the message $\hat{m}_1$ and then make a claim "Cindy failed on $\hat{m}_1$" to Ted.

Bob → Ted: $\hat{m}_1$, $\hat{m}_2$

Where $\hat{m}_2 = Sig_B(\hat{m}_1)$

(5-6). Ted follows the recovery procedure (r2). Bob will obtain the content of the message and Cindy, the conspirator, can get the receipt.

Ted → Bob: $enc = PK_B(\langle m_{subj}, m\rangle)$

Ted → Cindy: $\hat{m}_2$

Similarly, the same weakness towards replay attack is appeared in the extension protocol in the GG scheme. The attacking scenario is depicted in Figure 2 and the description of the figure is as follows.

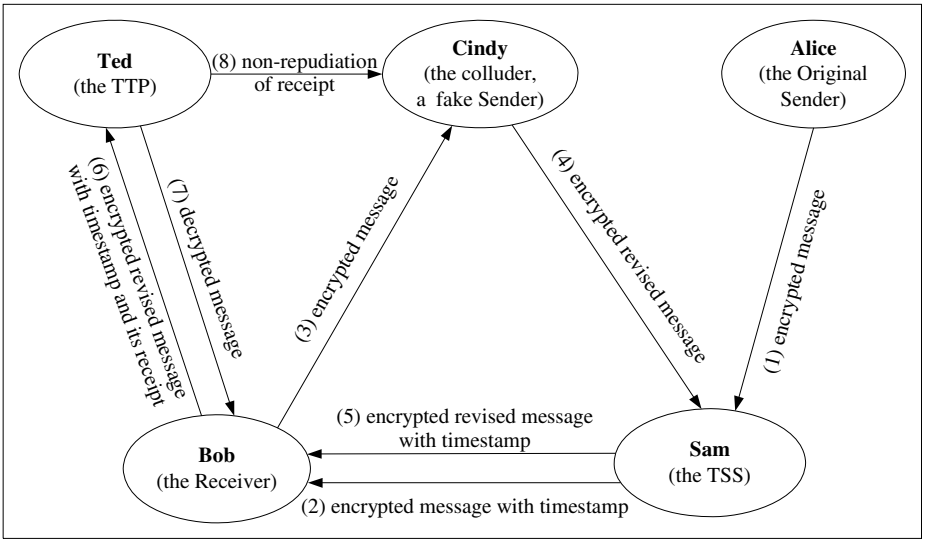


Fig. 2. Replay attack scenario on the extension protocol

(1). Alice → Sam: m_1

where $m_1 = \langle env, Sig_A(env)\rangle$,

$env = \langle ID_A, ID_B, PK_B(\langle m_{subj}, h(\langle m, r\rangle)\rangle), \overline{PK}_T(PK_B(msg), r)\rangle$, $msg = \langle m_{subj}, m\rangle$

(2). Sam → Bob: m_2

where $m_2 = \langle\langle m_1, t(m_1)\rangle, Sig_S(\langle m_1, t(m_1)\rangle)\rangle$

(3-4). Bob colludes with Cindy. Cindy creates a revised version $\hat{m}_1$ of the message m_1 in order to disguise as the sender.

Bob → Cindy: m_1

Cindy → Sam: $\hat{m}_1$

where $\hat{m}_1 = \langle env', \underline{Sig_C(env')}\rangle$,

$env' = \langle \underline{ID_C}, ID_B, PK_B(\langle m_{subj}, h(\langle m, r\rangle)\rangle), \overline{PK}_T(PK_B(msg), r)\rangle$, $msg = \langle m_{subj}, m\rangle$

(5). Sam → Bob: $\hat{m}_2$

where $\hat{m}_2 = \langle\langle \hat{m}_1, t(\hat{m}_1)\rangle, Sig_S(\langle \hat{m}_1, t(\hat{m}_1)\rangle)\rangle$

(6). Bob produces a false receipt $\hat{m}_3$ and then gives $\hat{m}_1$ and $\hat{m}_3$ to Ted for recovery.

Bob → Ted: $\hat{m}_1$, $\hat{m}_3$

where $\hat{m}_3 = \langle \hat{m}_2, Sig_B(\hat{m}_2)\rangle$

(7-8). Ted follows the recovery procedure (r2) after successful verification. Bob can get the content of the message without providing the valid receipt to Alice.

Ted → Bob: $enc = PK_B(\langle m_{subj}, m\rangle)$

Ted → Cindy: $\hat{m}_3$

The above replay attack demonstrates that fairness cannot be preserved in the GG protocol.

V-2. Incomplete Recovery Data on "Alice Failed on m_3"

According to the recovery procedures (r1 and r2) of the GG scheme, in case Alice fails on message m_3 Ted will compute the message *enc* from m_1 and send *enc* to Bob and m_2 to Alice after verifying the correctness of the message m_1 and m_2 provided by Bob. Here the problem is how to prove that the message *enc* is consistent with the non-repudiation receipt m_2. The receiver Bob in the GG scheme is designate to take the responsibility. However, it is beyond Bob's capability due to insufficient data. That is, Bob is short of the random string "*r*", and thus he cannot generate the hash of $\langle m, r\rangle$ from *enc* to compare it with the hash value $h(\langle m, r\rangle)$ from m_1.

4.2 Improvements

I-1. Protection Against Replay Attacks

A basic mechanism to prevent replay attacks is the *challenge-response* technique, in which, one entity (the claimant) proves its identity to another entity (the verifier) by demonstrating knowledge of a secret known to be associated with that entity [11]. This can be done by providing a response to a time-variant challenge that consists of three main classes of time-variant parameters: random numbers, sequence numbers, and timestamps. The weakness of the GG scheme against replay attacks is due to the inability of the TTP in detection of the real initiator of the message m_1 in the recovery procedures. Therefore, the identities of the involved parties (Alice, Bob, and TTP) and timestamp should be considered. The revised *env* in the message m_1 is $\langle ID_A, ID_B, ID_T, PK_B(\langle m_{subj}, h(\langle m, r\rangle)\rangle), \overline{PK_T}(\langle ID_A, ID_B, ID_T\rangle, t_d, PK_B(msg), r)\rangle$. Here, the identities of the involved parties $\langle ID_A, ID_B, ID_T\rangle$ will be effective against such an attack; and further, the timestamp t_d is used to provide the TTP with the deadline for dealing with the recovery procedures.

I-2. Provision of Complete Recovery Data on "Alice Failed on m_3"

The key to verify whether the delivered message corresponds to the non-repudiation of receipt m_2 and is also the promised one in m_1 is the hash value $h(\langle m, r\rangle)$. That

means, both *m* and *r* are required in verification. Therefore, a supplement to recovery data on "Alice failed on m_3" is the random string *r* contained in the recovery message *enc*. That is, Ted will compute the message $enc = \langle\langle PK_B(msg), r\rangle, Sig_T(\langle PK_B(msg), r\rangle)\rangle$ from m_1 and send *enc* to Bob and m_2 to Alice after the correct verification. Then, Bob will be able to verify that the receipt sent to Alice corresponds to the received message.

I-3. Specification of Encryption Algorithms

In the GG scheme, there is no clear specification on the public encryption algorithms PK_B and PK_T, except the authors stressed that PK_T is required to be a randomized encryption algorithm. So it seems the GG scheme works well with (a) any secure randomized encryption algorithm PK_T and (b) any secure encryption algorithm PK_B. However, this is not the fact. First, the random number *r* is needed to check $h(\langle m, r\rangle)$. In order to guarantee the TTP can recover *r*, it is required that from the ciphertext $\overline{PK}_T(PK_B(msg), r)$, the TTP can recover not only the message *m* but also *r*. This requirement is satisfied by the OAEP series of encryption schemes [4], but not by the Cramer-Shoup cryptosystem and the ElGamal encryption scheme. Similarly, we also need to assume PK_B is a deterministic encryption algorithm or a randomized encryption algorithm with the above mentioned property. If this is not true, a verifier (e.g., a judge) cannot verify the non-repudiation evidences.

5 Conclusion

The binding between the irrefutable *evidence* and the electronic mail being delivered is the purpose of certified email. The evidence will be a proof-of-delivery that the message was delivered to the recipient. A desirable requirement for a certified e-mail protocol is fairness.

In this paper, we briefly reviewed an optimistic scheme for certified e-mail proposed by Galdi and Giordano. Their scheme is effective against the failures of the participants in most cases. However, we found that it cannot achieve true fairness, i.e., in case of collusion. We further proposed the improvements to avoid such an attack.

References

1. M. Abadi, N. Glew, B. Horne, and B. Pinkas. Certified email with a light on-line trusted third party: Design and implementation. *Proceedings of 2002 International World Wide Web Conference*, pp. 387-395, ACM Press, 2002.
2. G. Ateniese, B.de Medeiros, and M.T. Goodrich. TRICERT: A distributed certified e-mail scheme. *Proceedings of 2001 Symposium on Network and Distributed Systems Security*, Internet Society, 2001.
3. Bahreman and J.D. Tygar. Certified electronic mail. *Proceedings of 1994 Symposium on Network and Distributed System Security*, pp. 3-19, Internet Society, 1994.
4. F. Bao, G. Wang, J. Zhou, and H. Zhu. Analysis and improvement of Micali's fair contract signing protocol. *Proceedings of 2004 Australasian Conference on Information Security and Privacy*, LNCS 3108, pp. 176-187, Springer-Verlag, 2004.

5. R. Deng, L. Gong, A. Lazar, and W. Wang. Practical protocol for certified electronic mail. *Journal of Network and Systems Management*, Vol. 4, No. 3, pp. 279-297, 1996.
6. J.L. Ferrer-Gomila, M. Payeras-Capella, and L. Huguet-Rotger. An efficient protocol for certified electronic mail. *Proceedings of 2000 Information Security Workshop*, LNCS 1975, pp. 237-248, Springer-Verlag, 2000.
7. C. Galdi and R. Giordano. Certified e-mail with temporal authentication: An improved optimistic protocol. *Proceedings of 2004 International Conference on Trust and Privacy in Digital Business*, LNCS 3184, pp. 181-190, Springer-Verlag, 2004.
8. K. Imamoto and K. Sakurai. A certified e-mail system with receiver's selective usage of delivery authority. *Proceedings of Indocrypt 2002*, LNCS 2551, pp. 326-338, Springer-Verlag, 2002.
9. S. Kremer and O. Markowitch. Selective receipt in certified e-mail. *Proceedings of Indocrypt 2001*, LNCS 2247, pp. 136-148, Springer-Verlag, 2001.
10. S. Kremer, O. Markowitch, and J. Zhou. An intensive survey of fair non-repudiation protocol. *Computer Communications*, Vol. 25, No.17, pp. 1606-1621, Elsevier, 2002.
11. A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone. *Handbook of applied cryptography*. CRC Press, ISBN: 0-8493-8523-7, October 1996.
12. J.R.M. Monteiro and R. Dahab. An attack on a protocol for certified delivery. *Proceedings of 2002 Information Security Conference*, LNCS 2851, pp. 428-426, Springer-Verlag, 2002.
13. J.A. Onieva, J. Zhou, and J. Lopez. Enhancing certified email service for timeliness and multicast. *Proceedings of 4th International Network Conference*, pp. 327-336, Plymouth, UK, 2004.
14. R. Oppliger. Certified mail: The next challenge for secure messaging. *Communications of the ACM*, Vol. 47, No. 8, August 2004.
15. B. Schneier and J. Riordan. A certified e-mail protocol. *Proceedings of 1997 Annual Computer Security Applications Conference*, pp. 232-238, IEEE computer Society Press, 1997.

Security Flaws in Several Group Signatures Proposed by Popescu

Guilin Wang[1] and Sihan Qing[2]

[1] Institute for Infocomm Research (I^{2}R),
21 Heng Mui Keng Terrace, Singapore 119613
glwang@i2r.a-star.edu.sg
[2] ERCIST, Institute of Software,
Chinese Academy of Sciences, Beijing 100080
qsihan@yahoo.com

Abstract. In resent years, Popescu et al. proposed several group signature schemes in [8, 9, 10, 11] that based on the Okamoto-Shiraishi assumption. Their schemes are claimed to be secure. However, we identify several security flaws in their schemes and then show that these schemes are all *insecure*. By exploiting those flaws, anybody (not necessarily a group member) can forge valid group signatures on arbitrary messages of his/her choice. In other words, these schemes are universally forgeable.

Keywords: group signature, digital signature, information security.

1 Introduction

Group signatures have been improved a lot as being first introduced by Chaum and van Heyst in [6] since 1991. They have many practical applications such as e-voting, e-bidding, e-cash, and fingerprinting systems etc. A group signature scheme allows each group member of a given group to sign messages anonymously on behalf of the group. And in case of disputes, a designated group manager can open a group signature and then identify the signer of it.

Following the first work by Chaum and van Heyst, a number of new group signature schemes and improvements have been proposed. Camenisch and Stadler [3] proposed the first group signature scheme for large groups, in which the lengths of the group public key and signatures are independent of the group size. In [4, 5], Camenisch and Michels constructed an efficient group signature scheme based on strong RSA assumption. In 1999, Ateniese and Tsudik [2] pointed out some obstacles that stand in the way of real world applications of group signatures, such as coalition attacks and member deletion. Later on, Ateniese et al. presented a practical and provably secure coalition-resistant group signature scheme in [1]. To deal with exposure of group members' secret keys and deletion of group members, Song [12] proposed forward-secure group signature schemes which support membership revocation.

Based on Okamoto-Shiraishi assumption [7], Popescu et al. recently proposed several group signature schemes. The authors first constructed two standard

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 711–718, 2005.

schemes in [8, 9], and then extended them to a group blind signature [10] and a scheme with revocation [11]. Compared with Song's schemes, Popescu's scheme in [11] has an advantage: The system life time does not need to be divided into a predefined number of time periods. In literatures [8, 9, 10, 11], the authors claimed that their schemes satisfy all the security requirements on group signatures (see Section 2 for details). However, this is not true.

In this paper, some serious security flaws in Popescu's schemes are successfully identified. Exploiting these flaws, an attacker can mount universally forging attacks without any secret. In other words, our attacks allow anybody (not necessarily a group member) to forge valid group signatures on arbitrary messages of his/her choice. This implies that these schemes are all insecure. Since these four schemes have similar constructions, we only review the latest scheme proposed in [11] (For short, the PNBM scheme), and piont out the related security flaws. Similar attacks also apply to other three schemes [8, 9, 10].

The rest of this paper is organized as follows. Section 2 introduces the Okamoto-Shiraishi assumption [7], and the security requirements on a group signature scheme. We then review and analyze the PNBM scheme [11] in Section 3 and 4, respectively. Finally, some concluding remarks are given in Section 5.

2 Assumption and Security Requirements

In this section, we briefly review the Okamoto-Shiraishi assumption on which all Popescu's schemes are based and list the security requirements on group signatures.

Okamoto-Shiraishi Assumption [7]. *Let e be an integer, $e \geq 4$. Given as inputs an RSA modulus $n = pq$ and an element $C \in \mathbb{Z}_n^*$, it is hard to find two integers X and δ such that $X^e \equiv C + \delta \bmod n$ and $\delta \in [a, b]$, where a and b are two integers satisfying $0 \leq a < b < n^{2/3}$.*

In a group signature scheme, a user registers with a *group manager* and then becomes a group member by getting a membership certificate. With the membership certificate, a group member can sign messages on behalf of the group in an anonymous and unlinkable manner. To check the validity of a group signature, however, a verifier only needs to know the unique group public key. In event of later disputes, the group manager can open a valid group signature and then find out the signer's identity.

In a secure group signature scheme, a valid signature can only be generated by using a valid membership certificate, and a new membership certificate cannot be created by any party other than the group manager. More rigorously, a *secure* group signature scheme must satisfy the following six properties [6, 3, 1, 2]:

1. **Unforgeability:** Only group members are able to sign messages on behalf of the group.
2. **Anonymity:** Given a valid signature of some message, identifying the actual signer is computationally hard for everyone but the group manager.

3. **Unlinkability:** Deciding whether two different valid signatures were computed by the same group member is computationally hard.
4. **Exculpability:** Neither a group member nor the group manager can sign on behalf of other group members.
5. **Traceability:** The group manager is always able to open a valid signature and identify the actual signer.
6. **Coalition-resistance:** A colluding subset (even a whole set of the entire group) of group members cannot generate a valid signature such that the group manager cannot link it to one of the colluding group members.

3 Review of the PNBM Scheme

This section reviews the PNBM group signature scheme proposed by Popescu et al. in [11]. The whole scheme consists of six components.

3.1 SETUP

To setup a system, the group manager performs the following steps:

(1) Select two random safe primes p and q, i.e., there exist two primes p' and q' such that $p = 2p' + 1$ and $q = 2q' + 1$. Then, the group manager sets his RSA modulus $n = pq$. Let l_n denote the bit-length of n.
(2) Pick $\bar{G} = \langle \bar{g} \rangle$ of order n in which computing discrete logarithms is infeasible. For example, $\bar{G}$ can be a subgroup of $\mathbb{Z}_{\bar{p}}^*$ for a large prime $\bar{p}$ such that $n|(\bar{p}-1)$.
(3) Choose a public exponent e satisfying $e > 4$, and $\gcd(e, \varphi(n)) = 1$.
(4) Select an element g of order $2p'q'$ in $\mathbb{Z}_n^*$. Let $G = \langle g \rangle$, and l_G denote the bit-length of the order of G, i.e., $|G| = |\mathrm{ord}(g)| = l_G$ [1].
(5) Select an element $C \in \mathbb{Z}_n^*$ and an element $h \in_R G$ whose discrete logarithm to the base g must not be known.
(6) Pick a secret value $x \in_R \mathbb{Z}_n^*$ and computes $y = g^x \bmod n$.
(7) Publish a collision-resistant hash function $H : \{0,1\}^* \rightarrow \{0,1\}^k$, and set security parameters $\epsilon > 1, l_1, l_2$.
(8) Finally, the public key is $PK = (n, e, g, \bar{g}, y, h, C, l_n, l_1, l_2, k, \epsilon, H)$ and the secret key is $SK = (p', q', x)$. In practice, components of PK must be verifiable to prevent framing attacks (refer to [4] for more detail).

An example for choosing the parameters is given by (see §5 of [9] or §2.1 of [11]): $l_n = 1200$, $l_1 = 350$, $l_2 = 240$, $k = 160$, $\epsilon = 5/4$, and $e = 5$.

[1] In [11], this step is specified as follows: Select g an element of $\mathbb{Z}_n^*$ of order n. Let $G = \langle g \rangle$ be a cyclic subgroup of $\mathbb{Z}_n^*$ of order l_G. We note their specification is incorrect. Firstly, no element in $\mathbb{Z}_n^*$ has an order n, since $2p'q'$ is the maximum order of an element in $\mathbb{Z}_n^*$. Secondly, l_G should denote the bit-length of the order of G, not the order itself. So we correct these errors in our description.

3.2 JOIN

Suppose now that a user U_i wants to join the group. It is also assumed that the communication between the group member and the group manager is secure, i.e., private and authentic. A membership certificate in the PNBM group signature scheme consists of a pair of integers (X, δ) satisfying $X^e \equiv C + \delta \bmod n$ and $\delta \in [2^{l_1}, 2^{l_1} + 2^{l_2} - 1]$. To obtain his membership certificate, each user U_i must perform the following protocol with the group manager.

(1) The user U_i selects a random element $x_i \in [2^{l_1}, 2^{l_1} + 2^{l_2} - 1]$, and computes $ID_i = g^{x_i} \bmod n$.
(2) The user U_i must prove to the group manager that he knows $\log_g ID_i$ and that this value is in the interval $(2^{l_1} - 2^{\epsilon(l_2+k)+1}, 2^{l_1} + 2^{\epsilon(l_2+k)+1})$.
(3) Then, the user U_i chooses a random number $r \in \mathbb{Z}_n^*$ and computes $z = r^e(C + x_i) \bmod n$. He sends z to the group manager.
(4) The group manager computes $v = z^{1/e} \bmod n = r(C + x_i)^{1/e} \bmod n$ and sends v to the user U_i.
(5) The user U_i computes $A_i = v/r = (C + x_i)^{1/e} \bmod n$. The pair (A_i, x_i) is the membership certificate of the user U_i.

Consequently, at the end of the protocol, the group manager does not know the membership certificate (A_i, x_i) of the user U_i. The group manager creates a new entry in the group database to store ID_i.

3.3 SIGN

With a membership certificate (A_i, x_i), a group member U_i can generate his group signature on any message $m \in \{0,1\}^*$ as follows.

(1) Select two random integers $w \in_R \{0,1\}^{l_2}$ and $r \in \mathbb{Z}_n^*$, and then compute: $A = A_i h^w \bmod n$, $B = g^w \bmod n$, $D = g^{x_i} y^w \bmod n$, $E = \bar{g}^r$, and $F = E^{b_s^{x_i}} \bmod n$. (Note that b_s is called the current revocation base, which is issued by the group manager in the REVOKE protocol. See the detail later.)
(2) Pick five random numbers $r_1 \in_R \{0,1\}^{\epsilon(l_2+k)}$, $r_2 \in_R \{0,1\}^{\epsilon(l_G+l_1+k)}$, $r_3 \in_R \{0,1\}^{\epsilon(l_G+k)}$, $r_4 \in_R \{0,1\}^{\epsilon(l_2+k)}$, $r_5 \in_R \{0,1\}^{\epsilon(l_2+k)}$, and then compute: $d_1 = B^{r_1}/g^{r_2} \bmod n$, $d_2 = g^{x_i^2} D^{r_4}/y^{r_5} \bmod n$, $d_3 = g^{r_3} \bmod n$, and $d_4 = g^{r_1} y^{r_3} \bmod n$.
(3) Evaluate hash value $c = H(m||g||h||y||A||B||D||E||F||d_1||d_2||d_3||d_4)$.
(4) Calculate $s_1 = r_1 - c(x_i - 2^{l_1})$, $s_2 = r_2 - cx_i w$, $s_3 = r_3 - cw$, $s_4 = r_4 + x_i + c2^{l_1}$, $s_5 = r_5 + x_i w + c2^{l_1}$ (all in $\mathbb{Z}$).
(5) Release $(c, s_1, s_2, s_3, s_4, s_5, A, B, D)$ as the group signature for message m.
(6) The user U_i proves in zero-knowledge that the double discrete logarithm of F with bases E and b_s, respectively, is the same as the discrete logarithm of D's representation to the bases g and h [2].

[2] Note that there exist such protocols though not very efficient. For example, the one proposed in [3].

Since D is computed as $D = g^{x_i} y^w \bmod n$, the resulting proof of knowledge is verifiable if and only if the same x_i is used in the construction of both F and D.

3.4 VERIFY

Upon receiving an alleged group signature $(c, s_1, s_2, s_3, s_4, s_5, A, B, D)$ on a message m, a verifier can check its validity as follows:

(1) Compute $d'_1 = B^{s_1 - c2^{l_1}} / g^{s_2} \bmod n$, $d'_2 = D^{s_4 - c2^{l_1}} / y^{s_5 - c2^{l_1}} \bmod n$, $d'_3 = B^c g^{s_3} \bmod n$, and $d'_4 = D^c g^{s_1 - c2^{l_1}} y^{s_3} \bmod n$.
(2) Evaluate hash value $c' = H(m||g||h||y||A||B||D||E||F||d'_1||d'_2||d'_3||d'_4)$.
(3) Check whether $c \equiv c'$ and $s_1 \times s_2 \times s_3 \times s_4 \times s_5 \in \left\{-2^{l_2+k}, ..., 2^{\epsilon(l_2+k)}\right\} \times \left\{-2^{l_G+l_1+k}, ..., 2^{\epsilon(l_G+l_1+k)}\right\} \times \left\{-2^{l_G+k}, ..., 2^{\epsilon(l_G+k)}\right\} \times \left\{-2^{l_2+k}, ..., 2^{\epsilon(l_2+k)}\right\} \times \left\{-2^{l_2+k}, ..., 2^{\epsilon(l_2+k)}\right\}$.
(4) For each $V_{s,j} \in CRL$, check if $F \neq E^{V_{s,j}} \bmod n$.
(5) Check the proof of equality of double discrete logarithm fo F and the discrete logarithm of D's representation to the bases g and h.
(6) Accept the group signature $(c, s_1, s_2, s_3, s_4, s_5, A, B, D)$ if and only if all the above three checks hold.

3.5 OPEN

When a group signature $(c, s_1, s_2, s_3, s_4, s_5, A, B, D)$ on a message m is given, the group manager can find out which member issued this signature by first checking its correctness via the VERIFY protocol. If the signature is not correct, he stops. Otherwise, the group manager performs the following steps to identify the signer.

(1) Recover $ID_i = D/B^x \bmod n$, and use the identity information ID_i to find the true signer U_i.
(2) Prove in zero-knowledge that $\log_g y = \log_B (D/ID_i \bmod n)$ [3, 9].

3.6 REVOKE

We begin by assuming, as usual, that a Certificate Revocation List (CRL) is a structure available at all time from a number of well-known public repositories or severs. A CRL is also assumed to be signed and timestamped by its issuer which may be a universally trusted CA, or the group manager. In addition, revocability means that a group signature produced using the SIGN algorithm by a revoked member must be rejected using the VERIFY algorithm.

We use s to denote the index of the current CRL issue where there are l group members to be revoked. The following REVOKE algorithm is executed by the group manager whenever a member or a collection of members leaves or is expelled.

(1) Choose a random number $b_s \in_R QR(n)$ of order $p'q'$. This value b_s becomes the *current revocation base*.
(2) For each U_j $(1 \leq j \leq l)$, compute $V_{s,j} = b_s^{x_j} \bmod n$.
(3) The actual revocation list is then published $CRL = \{b_s, V_{s,j} | 1 \leq j \leq l\}$.

4 Security Flaws in the PNBM Scheme

Popescu et al. claimed that their above scheme (and other schemes) satisfies all the security requirements listed in Section 2. However, we find that this is not true.

4.1 REVOKE Algorithm

First of all, we note that the REVOKE algorithm proposed in the PNBM scheme does not work in the normal framework of group signatures.

The reason is that in PNBM scheme, to issue each $V_{s,j}$ for all group members to be revoked, the group manager needs to know the value of x_i. However, x_i is U_i' member secret which cannot be revealed to anyone including the group manager. If the group manager knows the value of x_i, it can recover the certificate A_i by computing $A_i = (C + x_i)^{1/e} \bmod n$. In this condition, the group manager can mount a *framing attack*, i.e., he can use the membership certificate (A_i, x_i) to generate valid signatures on behalf of the group member U_i. Apparently, this weakness is intolerable in any group signature scheme since *exculpability* is not satisfied any more. This is a design error in their scheme.

Furthermore, according to the analysis presented below, the PNBM scheme [11] (as well as the other three schemes) is not secure even in the situation where member revocation is not supported.

4.2 Security Parameters

Note that in the Camenisch-Michels scheme [4], the security parameters l_1 and l_2 are set as $l_1 = 860$, and $l_2 = 600$. While in the PNBM scheme, the authors suggested to set the security parameters as $l_1 = 350$, and $l_2 = 240$ (see Section 2.1 of [11]). With much shorter exponents, the PNBM scheme may be more efficient[3]. However, the security parameters l_1 and l_2 should be selected as larger numbers. Especially, the difference between these two parameters should be guaranteed big enough. Otherwise, the schemes are vulnerable to some forging membership certificate attacks [13]. Usually, the following condition is required [1]:

$$l_1 > \epsilon(l_2 + k) + 2.$$

4.3 Cheating in the JOIN Protocol

In the third step of the JOIN protocol, user U_i is not required to prove that z and ID_i committed the same secret value of x_i. Therefore, a dishonest user U_i can replace x_i in z with a random number $\bar{x}_i \in_R [2^{l_1}, 2^{l_1} + 2^{l_2} - 1]$. That is, by choosing a random number $r \in_R \mathbb{Z}_n^*$, U_i prepares a value of $\bar{z}$ as

$$\bar{z} = r^e(C + \bar{x}_i) \bmod n,$$

[3] However, the PNBM scheme is not much efficient, since the signer has to execute a zero-knowledge protocol to show that he knows a double discrete logarithm. As we mentioned above, this is time-expensive.

and sends $\bar{z}$ to the group manager. Then, according to the JOIN protocol, the group manager sends back U_i the value of $\bar{v}$ which satisfies $\bar{v} = \bar{z}^{1/e} \bmod n$. Finally, U_i gets a valid membership certificate $(\bar{A}_i, \bar{x}_i)$ by computing

$$\bar{A}_i = \bar{v}/r \bmod n.$$

Using this valid certificate, U_i can generate valid group signatures at will. In the event of disputes, however, the group manager cannot open the signatures generated by such certificates. Because U_i uses $\bar{x}_i$ in stead of x_i in the SIGN protocol, and $\bar{x}_i$ has no any relationship with ID_i, the U_i's identity.

4.4 Universal Forgery

We note that in the SIGN protocol, the value of $A = A_i h^w \bmod n$ is not used in essence (except it is embedded in the hash value of c). In other words, what the SIGN protocol proves is that the signer knows some secrets such that the values of B and D are prepared properly. But the critical fact whether A and D commit the same secret x_i is not proved. Therefore, anybody (not necessarily a group member) can generate a valid group signature for any message m of his choice as group member does.

To this end, an attacker first picks two random numbers $\bar{A}_i \in_R \mathbb{Z}_n^*$, and $\bar{x}_i \in_R [2^{l_1}, 2^{l_1} + 2^{l_2} - 1]$. Then, he can generate group signatures on any messages according to the procedures described in the SIGN protocol. It is easy to check that the resulting signatures are valid, i.e., they satisfies the VERIFY protocol.

5 Concluding Remarks

In this paper, we identified four security flaws in the group signature scheme with revocation by Popescu et al. [11]. Except the problem in the REVOKE algorithm, other security flaws can also be used to break the three schemes proposed in [8, 9, 10] by Popescu, since these schemes share similar constructions. Therefore, our results showed that all these schemes are completely *insecure*, and that the scheme in [11] does not support member revocation in essence. From our discussions presented above, we know that these security flaws mainly result from the insecurity of the JOIN and SIGN protocols, i.e., they are not designed securely. To rectify these schemes, the JOIN and SIGN protocols should be carefully re-designed. The designing rule of thumb is that provably secure protocols are preferable and convincing. If member deletion is a necessary function in the system, a new REVOKE algorithm has to be proposed, too.

Acknowledgements: The authors would like to thank Dr. Tieyan Li, and the anonymous referees for their helpful suggestions on the improvement of this paper.

References

1. G. Ateniese, J. Camenisch , M. Joye, and G. Tsudik. A practical and provably secure coalition-resistant group signature scheme. In: *Crypto'2000, LNCS 1880*, pp. 255-270. Springer-Verlag, 2000.
2. G. Ateniese and G. Tsudik. Some open issues and new directions in group signature schemes. In: *Financial Cryptography* (*FC'99*), *LNCS 1648*, pp. 196-211. Springer-Verlag, 1999.
3. J. Camenisch and M. Stadler. Effient group signature schemes for large groups. In: *Crypto'97, LNCS 1294*, pp. 410-424. Springer-Verlag, 1997.
4. J. Camenisch and M. Michels. A group signature scheme with improved efficiency. In: *ASIACRYPT'98, LNCS 1514*, pp. 160-174. Springer-Verlag, 1998.
5. J. Camenisch and M. Michels. A group signature scheme based on an RSA-variant. *Technical Report RS-98-27*, BRICS, University of Aarhus, November 1998. An earlier version appears in [4].
6. D. Chaum and E. van Heyst. Group Signatures. In: *Eurocrypt'91, LNCS 950*, pp. 257-265. Springer-Verlag, 1992.
7. T. Okamoto and A. Shiraishi. A fast signature scheme based on quadratic inequalities. In: *Proceedings of IEEE Symposium on Security and Privacy*, pp. 123-132. IEEE Computer Society, 1985.
8. C. Popescu. Group signature schemes based on the difficulty of computation of approximate e-th roots. In: *Proc. of Protocols for Multimedia Systems* (*PROMS 2000*), pp. 325-331. Cracow, Poland, 2000.
9. C. Popescu. An efficient group signature scheme for large groups. *Studies in Informatics and Control Journal*, Vol. 10, No. 1: 7-14, March 2001. `http://www.ici.ro/ici/revista/sic2001-1/art1.htm`
10. C. Popescu. A group blind signature scheme based on the Okamoto-Shiraishi assumption. In: *Romanian Journal of Information Science and Technology*, Vol. 4, No. 5, 2002.
11. C. Popescu, D. Noje, B. Bede, and I. Mang. A group signature scheme with revocation. In: *Proc. of 4th EURSIP Conference focused on Video/Image Processing and Multimedia Communications* (*EC-VIP-MC 2003*), pp. 245-250. 2-5 July 2003, Zagreb, Coratia. IEEE Computer Society, available via IEEExplore `http://ieeexplore.ieee.org/`.
12. D.X. Song. Practical forward secure group signature schemes. In: *ACM CCS'01*, pp. 225-234. ACM press, 2001.
13. G. Wang, F. Bao, J. Zhou, and R.H. Deng. Security remarks on a group signature scheme with member deletion. In: *Information and Communications Security (ICICS 2003), LNCS 2836*, pp. 72-83. Springer-Verlag, 2003.

A Simple Acceptance/Rejection Criterium for Sequence Generators in Symmetric Cryptography*

Amparo Fúster-Sabater[1] and Pino Caballero-Gil[2]

[1] Instituto de Física Aplicada, C.S.I.C.,
Serrano 144, 28006 Madrid (Spain)
amparo@iec.csic.es
[2] DEIOC, University of La Laguna,
38271 La Laguna, Tenerife, Spain
pcaballe@ull.es

Abstract. A simple method of checking the degree of balancedness in key-stream generators of cryptographic application has been developed. The procedure is based exclusively on the handling of bit-strings by means of logic operations and can be applied to standard generators proposed and published in the open literature (combinational generators, multiple clocking generators, irregularly clocked generators). The requirements of time and memory complexity are negligible. The method here developed is believed to be a first selective criterium for acceptance/rejection of this type of generators with application in symmetric cryptography.

Keywords: Confidentiality, stream cipher, bit-string algorithm, cryptography.

1 Introduction

Transmission of sensitive information between two interested parties needs several security requirements (confidentiality, integrity, non repudiation, authentication ...) that can be satisfied by means of design, assessment and implementation of cryptographic algorithms and security protocols.

Confidentiality makes use of an encryption function currently called *cipher* that converts the *plaintext* into the *ciphertext*. Ciphers are usually divided into two large classes: stream ciphers and block-ciphers. Stream ciphers are very fast (in fact, the fastest among the encryption procedures) so they are implemented in many technological applications e.g. algorithms A5 in GSM communications [5] or the encryption system E0 used in the Bluetooth specifications [1]. Stream

* Research supported by Ministerio de Educación y Ciencia (Spain) under grant SEG2004-02418 and SEG2004-04352-C04-03.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 719–728, 2005.

ciphers try to imitate the ultimate one-time pad cipher and are supposed to be good pseudorandom generators capable of stretching a short secret seed (the secret key) into a long sequence of seemingly random bits. This key-stream is then XORed with the plaintext in order to obtain the ciphertext. Most generators producing key-stream sequence are based on Linear Feedback Shift Registers (LFSRs) [4]. The pseudorandom output sequence is a periodic sequence generated as the image of a nonlinear Boolean function in the LFSR stages.

Balancedness in the output sequence is a necessary although never sufficient condition that every cryptographic generator must satisfy. Roughly speaking, a binary sequence is balanced if it has approximately the same number of $1's$ as $0's$. Due to the long period of the generated sequence ($T \simeq 10^{38}$ bits in current cryptographic applications), it is unfeasible to produce an entire cycle of such a sequence and then count the number of $1's$ and $0's$. Therefore, in practical design of binary generators, portions of the output sequence are chosen randomly and statistical tests (frequency test or monobit test [8]) are applied to all these subsequences. Nevertheless, passing the previous tests merely provides *probabilistic evidence* that the generator produces a balanced sequence (see [9], [10]).

In the present work, balancedness of LFSR-based generators has been treated in a *deterministic way*. In fact, a simple binary model allows one to compute the exact number of $1's$ (number of $0's$) in the output sequence without producing the whole sequence. The general expression of the number of $1's$ is obtained as a function of the generator parameters. The number of $0's$ is just the sequence period minus the number of $1's$. In this way, the degree of balancedness of the sequence can be perfectly checked: the obtained number of $1's$ ($0's$) is compared with the value required for this sequence to be balanced (half the period $\pm$ a tolerance interval). In case of non-accordance, the LFSR-based generator must be rejected. Thus, the procedure here developed can be considered as a first selective criterium for acceptance/rejection of this type of generators. At any rate, generators satisfying the balancedness requirement must be subjected to further testing.

The computational method is based exclusively on the handling of binary strings by means of the logic operation OR. Indeed, the general expression of the number of digits is just an interpretation of such binary strings. The procedure can be applied to cryptographic generators proposed and published in the open literature. Some illustrative examples including combination generators, filter generators, multiple clocking or irregularly clocked generators complete the work.

2 Fundamentals and Basic Concepts

Several basic concepts and definitions to be used throughout the paper will be presented in the following subsections.

Definition 1. *A minterm of L binary variables* (m_0, m_1, ..., m_{L-1}) *is a monomial of the L variables, where each variable can be in its true or complementary form. For L variables, there exist* 2^L *minterms, each minterm being expressed as the logic product of the L (appropriately complemented) variables.*

If a minterm includes the variables m_i ... m_j in their true form and the rest of variables are in complementary form, then such a minterm is denoted by $M_{i...j}$.

Since a binary sequence generator is characterized by a Boolean function, the representation of such functions is considered.

2.1 Representation of Boolean Functions

Two different representations of Boolean functions are introduced.

1. *Boolean functions in Algebraic Normal Form:* Any L-variable Boolean function can be uniquely expressed in Algebraic Normal Form (A.N.F.) or Müller expansion ([7], [11]) by means of the exclusive-OR sum of logic products in the L variables. A simple example of Boolean function in A.N.F. is:

$$F(m_0,\, m_1,\, ...,\, m_{L-1}) = m_0\, m_1\, m_{L-1} \oplus m_1\, m_{L-2} \oplus m_{L-1},$$

   where the concatenation of variables represents the logic product and the symbol $\oplus$ the exclusive-OR logic operation.
2. *Boolean functions in terms of their minterms:* Any L-variable Boolean function can be canonically expressed as a linear combination of its minterms ([7], [11]). A simple example of Boolean function in terms of its minterms is:

$$F'(m_0,\, m_1,\, ...,\, m_{L-1}) = M_{012} \oplus M_{0\,L-2} \oplus M_{2\,L-2\,L-1}.$$

In this work, both representations of Boolean functions will be systematically addressed.

2.2 LFSR-Based Generators

A binary LFSR is an electronic device with N memory elements (stages), cyclic shifting and linear feedback [4]. In the sequel, only maximum-length LFSRs will be considered (LFSRs whose output sequences have maximum period of value $2^N - 1$). In fact, these LFSRs cycle through their $2^N - 1$ different states since the zero state is excluded.

Definition 2. *An LFSR-based generator is a nonlinear Boolean function* $F : GF(2)^L \rightarrow GF(2)$ *in A.N.F., whose input variables* m_i $(i = 0, ..., L-1)$ *are the binary contents of the LFSR stages.*

At each new clock pulse, the new binary contents of the stages will be the new input variables of the function F. In this way, the generator produces the successive bits of the output sequence.

Definition 3. *A minterm function is a minterm of L variables expressed in A.N.F. Every minterm function can be easily obtained by expanding out the corresponding minterm.*

Let $M_{i...j}$ be a minterm with d indexes $(1 \leq d \leq L)$. According to [7], its corresponding minterm function is perfectly characterized:

- Such a minterm function has $\binom{L-d}{0}$ terms of order d, $\binom{L-d}{1}$ terms of order $d+1$, $\binom{L-d}{2}$ terms of order $d+2$, ..., $\binom{L-d}{L-d}$ terms of order L, so in total the number of terms is:
$$\text{No. of terms} = 2^{L-d}. \tag{1}$$
- The particular form of the terms of each orden is perfectly determined too, see [7]. In fact, the LFSR stages involved in the least order term are repeated in the rest of terms of greater order.

On the other hand, every minterm function considered as an generator applied to the L stages of an LFSR generates a canonical sequence with a unique 1 and period $T = 2^L - 1$ (see [9]). Finally, let us introduce the minterm function of F notated Φ_F.

Definition 4. *Let Φ_F be the minterm function of F defined as a nonlinear Boolean function that substitutes each term $m_i\, m_j \ldots m_k$ of F for its corresponding minterm function $M_{ij\ldots k}$. The function Φ_F is an involution.*

$$F(m_i) = m_0\, m_1 \oplus m_2 \quad \Rightarrow \quad \Phi_F(m_i) = M_{01} \oplus M_2.$$

The generalization of the previous concepts to several LFSRs is quite immediate. In fact, let A, B, ..., Z be maximum-length LFSRs whose lengths are respectively L_A, L_B, ..., L_Z (supposed $(L_i, L_j) = 1$, $i \neq j$). We denote by a_i $(i = 0, ..., L_A - 1)$, b_j $(j = 0, ..., L_B - 1)$, ..., z_k $(k = 0, ..., L_Z - 1)$ their corresponding stages. The minterms of a nonlinear generator, called *global minterms*, are of the form, e.g. $A_{ij}\, B_{pqr} \ldots Z_s$ that is the logic product of the individual minterms of each LFSR. Therefore, each global minterm depends on L variables, where L is given by:

$$L = L_A + L_B + \ldots + L_Z.$$

A global minterm function is a global minterm of L variables expressed in A.N.F. As before, every global minterm function considered as a generator applied to the stages of the LFSRs [9] generates a canonical sequence with an unique 1 and period $T = (2^{L_A} - 1)(2^{L_B} - 1) \ldots (2^{L_Z} - 1)$.

In brief, every LFSR-based generator can be expressed as a linear combination of its minterms as well as each minterm provides the output sequence with a unique 1. Thus, the basic idea of this work can be summarized as follows:

- The number of minterms in the expression of F equals the number of $1's$ in the output sequence.

As every LFSR-based generator is designed in Algebraic Normal Form, the Boolean function F has first to be converted from its A.N.F. into its minterm expansion.

3 Boolean Function Conversion from A.N.F. to Minterm Expansion

The conversion procedure is carried out as follows:

Input: N_Z (number of LFSRs), L_A, L_B, ..., L_Z (lengths of the LFSRs) and a nonlinear function F given in A.N.F.

For instance, $N_Z = 2$, $L_A = 2$, $L_B = 3$ and $F(a_0, a_1, b_0, b_1, b_2) = a_0\, b_0$.

– *Step 1:* Compute Φ_F

$$\Phi_F = A_0\, B_0.$$

– *Step 2:* Substitute every minterm by its corresponding function in A.N.F.

$$\Phi_F = (a_0\, a_1 \oplus a_0)(b_0\, b_1\, b_2 \oplus b_0\, b_1 \oplus b_0\, b_2 \oplus b_0).$$

– *Step 3:* Compute $F(a_i, b_j) = \Phi_F \circ \Phi_F$

$$F(a_i, b_j) = \Phi_F \circ \Phi_F = A_{01}\, B_{012} \oplus A_{01}\, B_{01} \oplus A_{01}\, B_{02} \oplus A_{01}\, B_0 \oplus$$

$$A_0\, B_{012} \oplus A_0\, B_{01} \oplus A_0\, B_{02} \oplus A_0\, B_0.$$

Output: F expressed in terms of its global minterms.
The number of global minterms in step 3 (or equivalently the number of terms in step 2) coincides with the number of 1's in the generated sequence.

4 An Efficient Algorithm to Compute the Degree of Balancedness in Binary Generators

An algorithm that automates the comparison among the different terms in Φ_F, checks the cancelled terms and computes the number of final terms is presented. Such an algorithm is based on an L-bit string representation.

4.1 Specific Terminology for the Implementation

First of all some simplifications in the notation are introduced. A Greek letter (for example α) will represent the set of indexes $ij \ldots k$ above described to label the minterm functions, i.e. $A_{ij\ldots k} = A_\alpha$. A similar notation will be used for the product of LFSR stages, $a_i\, a_j \ldots a_k = a_{ij\ldots k} = a_\alpha$. According to this notation, the nonlinear function F and its minterm function Φ_F can be written as $F = \sum_{\oplus} a_{\alpha_i}$ and $\Phi_F = \sum_{\oplus} A_{\alpha_i}$, respectively. In subsection 2.2, it was told that the minterm function A_α had in total $2^{L-d(\alpha)}$ terms, $d(\alpha)$ being the number of indexes in α. In order to implement this algorithm, every minterm function A_α is represented by a L-bit string numbered $0, 1, ..., L-1$ from right to left. If the n-th index is in the set α ($n \in \alpha$), then the n-th bit of such a string takes the value 1; otherwise, the value will be 0. Thus, $d(\alpha)$ equals the number of $1's$ in the L-bit string that represents A_α.

Definition 5. *We call maximum common development (m.c.d.) of two minterm functions A_α and A_β, notated $MD(A_\alpha, A_\beta)$, to the minterm function A_χ such that $\chi = \alpha \cup \beta$.*

Under the L-bit string representation of the minterm functions, the m.c.d. can be realized by means of a bit-wise OR operation between the binary strings of both functions. The m.c.d. represents all the terms that A_α and A_β have in common. If two minterm functions A_α and A_β are added, $A_\alpha \oplus A_\beta$, then the terms corresponding to their m.c.d. are cancelled. Thus, the total number of terms in $A_\alpha \oplus A_\beta$ is the number of terms in A_α plus the number of terms in A_β minus twice the number of terms in the m.c.d., that is $2^{L-d(\alpha)} + 2^{L-d(\beta)} - 2 \cdot 2^{L-d(\alpha \cup \beta)}$.

Definition 6. *The auxiliary function H is defined as $H = \sum_i s_i A_{\alpha_i}$, s_i being an integer with sign that specifies how many times the function A_{a_i} is contained in H.*

In a symbolic way, H indicates whether the minterm functions A_{α_i} are added (sign +) or cancelled (sign -) as well as how many times such terms have been added or cancelled. Remark that the m.c.d. can be applied to the functions H too. Keeping in mind all these definitions, we can now describe the algorithm.

4.2 The Algorithm

Let $F = \sum_{\oplus} a_{\alpha_i}$ $(i = 1, ..., N)$ be a nonlinear Boolean function of N terms applied to the stages of the LFSRs. In order to compute the number of $1's$ (notated $\mathcal{U}_F$) in the generated sequence, the following algorithm is introduced:

- *Step 1:* Define the function Φ_F from the N terms a_{α_i} of F. Initialize the function H with a null value, $H_0 = \oslash$.
- *Step 2:* Run this loop from $i = 1$ to $i = N$: update $H_i = H_{i-1} + A_{\alpha_i} - 2 \cdot MD\,(A_{\alpha_i}, H_{i-1})$.
- *Step 3:* From the final form of $H_N = \sum_j s_j A_{\beta_j}$, compute the number of $1's$ in the generated sequence by means of the expression $\mathcal{U}_F = \sum_j s_j \cdot 2^{L-d(\beta_j)}$.

5 Computing the Number of 1's in Standard Binary Generators

The previous section presents a procedure to compute the number of 1's in the sequence obtained from a LFSR-based generator. Now we are going to apply such a procedure to standard generators of cryptographic application.

5.1 Combination Generators

Let A, B, C be three LFSRs of lengths L_A, L_B, L_C respectively. The LFSR-combination generator is chosen:

$$F = \sum_{i=1}^{3} m_{\alpha_i} = a_0 b_0 \oplus b_0 c_0 \oplus c_0, \tag{2}$$

which corresponds to the Geffe's generator [10]. In order to fix the L-bit strings, very low values are assigned to the lengths of the LFSRs: $L_A = 2, L_B = 3, L_C = 5$, with $((L_i, L_j) = 1,\ i \neq j)$, thus $L = 10$. According to the previous section, we proceed:

Step 1.- $\Phi_F = \sum_{i=1}^{3} M_{\alpha_i} = A_0\, B_0 \oplus B_0\, C_0 \oplus C_0$ and the minterms of Φ_F in 10-bit string format are:

$$\begin{aligned} M_{\alpha_1} &= A_0\, B_0 = 00000\ 001\ 01 \\ M_{\alpha_2} &= B_0\, C_0 = 00001\ 001\ 00 \\ M_{\alpha_3} &= C_0 \quad\ = 00001\ 000\ 00. \end{aligned}$$

H is initialized $H_0 = \oslash$.

Step 2.- For $i = 1, \ldots, 3$: realize $H_i = H_{i-1} + A_{\alpha_i} - 2 \cdot MD\,(A_{\alpha_i}, H_{i-1})$. The final value H_3 obtained at the end of the loop is:

$$H_3 = H_2 + M_{\alpha_3} - 2 \cdot MD\,(M_{\alpha_3}, H_2) =$$

$$00000\ 001\ 01\ +\ 00001\ 000\ 00\ -\ [1]\,00001\ 001\ 00.$$

Step 3.- Calculation of the number of $1's$ from H_3

$$\begin{aligned} 00000\ 001\ 01 \quad &\text{implies} \quad 2^{L_A-1}\, 2^{L_B-1}\, (2^{L_C} - 1) \ \text{ones} \\ 00001\ 000\ 00 \quad &\text{implies} \quad (2^{L_A} - 1)\, (2^{L_B} - 1)\, 2^{L_C-1} \ \text{ones} \\ -\,[1]\ 00001\ 001\ 00 \quad &\text{implies} \quad -\,(2^{L_A} - 1)\, 2^{L_B-1}\, 2^{L_C-1} \ \text{ones}. \end{aligned}$$

Thus,

$$\mathcal{U}_F = 2^{L_A-1}\, 2^{L_B-1}\, (2^{L_C} - 1) + (2^{L_A} - 1)\, (2^{L_B-1} - 1)\, 2^{L_C-1}. \qquad (3)$$

For lengths of the LFSRs in a cryptographic range $L_i \approx 64$ the number of $1's$ in the output sequence is $\simeq T/2$. Consequently, the generated sequence is quasi balanced.

The application of this procedure gives us a general expression for the number of $1's$ in the sequence produced by a Geffe's generator. From very low values of L_i, a general expression is achieved that can be applied to LFSR lengths in a range of practical interest.

5.2 Filter Generators: Straight Application to TOYOCRYPT-HR1

For a unique LFSR and from the previous algorithm, it can be noticed that if a minterm function in Φ, for instance M_{α_N},includes a unique index non-repeated, then at the N-th loop, the final updating of H is:

$$H_N = H_{N-1} + M_{\alpha_N} - 2 \cdot MD\,(M_{\alpha_N}, H_{N-1})$$

Nevertheless, the terms $2 \cdot MD\,(M_{\alpha_N}, H_{N-1})$ and H_{N-1} in the second member contains exactly the same binary strings except for the new index. Thus, regarding the number of $1's$, we have for every pair of quasi-equal strings:

$$2^{L-k} - 2 \cdot 2^{L-k-1} = 2^{L-k} - 2^{L-k} = 0$$

for any k. Therefore, the final number of $1's$ in the sequence generated is due exclusively to the contribution of the term M_{α_N}. Thus,

$$\mathcal{U}_F = 2^{L-1}. \tag{4}$$

Consequently, the output sequence of such generators will be balanced. That is the case of the TOYOCRYPT-HR1 keystream generator for the stream cipher TOYOCRYPT-HS1 [3]. In fact, this pseudorandom binary generator employs a LFSR of length $L = 128$ where the last stage, notated a_{127}, appears as a unique term with non-repeated index in the generating function F. In this practical example and according to the previous algorithm the balancedness of the output sequence is always guaranteed.

5.3 Massey-Rueppel Generator (Multiple Clocking)

Let A, B be two LFSRs of lengths L_A, L_B respectively, where $(L_A, L_B) = 1$ and $L_A < L_B$. The combining function [9] is:

$$F = a_0 b_0 + a_1 b_1 + ... + a_{L_A-1} b_{L_A-1}. \tag{5}$$

The LFSR B is clocked at a different speed factor from that one of LFSR A. For instance, B is clocked d times faster than A. In practice, $(d, 2^{L_B-1}) = 1$. This scheme is equivalent to the following:

- The same LFSR A.
- The same LFSR B repeated L_A times. The initial states of these LFSRs are separated a distance d.

Modifying the characteristic polynomial or the initial states of the LFSRs does not affect the number of $1's$ in the output sequence but the location of such $1's$. Thus, the Massey-Rueppel generator is a particular example of a broader family of binary sequence generators where all of them produce sequences with the same number of $1's$.

Remark now the possible cancellations in the terms of the function Φ_F:

- For terms $a_i B_i$, there are no cancelled terms so that every product provides 2^{L_B-1} terms in Φ_F.
- For terms $a_{ij} B_i$, $a_{ij} B_j$, there are $(2-1)\, 2^{L_B-1}$cancelled terms and 2^{L_B-1} terms left.
- For terms $a_{ijk} B_i$, $a_{ijk} B_j$, $a_{ijk} B_k$, there are $(3-1)\, 2^{L_B-1}$cancelled terms and 2^{L_B-1} terms left,

and so on ... Thus, the number of non-cancelled terms in Φ_F or equivalently the number of $1's$ in the sequence generated is:

$$\mathcal{U}_F = \sum_{i=1}^{L_A} \binom{L_A}{i} 2^{L_B-1} = (2^{L_A} - 1)\, 2^{L_B-1}, \tag{6}$$

which gives us a general expression for the number of 1's in the sequence obtained from a broad family of multi-clock generators. For L_A, L_B in a cryptographic range, the number of $1's$ in the sequence generated is $\simeq T/2$ and the output sequence is quasi balanced.

5.4 The Shrinking Generator (Irregularly Clocked Generator)

The shrinking generator is a binary sequence generator [2] composed of two LFSRs : a control register, LFSR A, that decimates the sequence produced by the other register, LFSR B. We denote by L_A, L_B their corresponding lengths where $(L_A, L_B) = 1$ and $L_A < L_B$.

The sequence produced by the LFSR A, that is $\{a_i\}$, controls the bits of the sequence produced by the LFSR B, that is $\{b_i\}$, which are included in the output sequence $\{c_j\}$ (the shrunken sequence). The decimation rule P is described as follows:

1. If $a_i = 1 \Longrightarrow c_j = b_i$
2. If $a_i = 0 \Longrightarrow b_i$ is discarded.

In brief, the sequence produced by the shrinking generator is an irregular decimation of the sequence generated by B controlled by the bits of A. The period of the shrunken sequence [2] is $T = (2^{L_B} - 1)2^{(L_A - 1)}$. The application of the previous algorithm to the shrinking generator can be realized as follows. The generating function in terms of a Boolean function is:

$$F = a_0 b_0, \tag{7}$$

as the decimation affects the period of the output sequence but never the number of $1's$. In order to fix the L-bit strings, very low values are assigned to the lengths of the LFSRs: $L_A = 2, L_B = 3$ with $((L_A, L_B) = 1)$, thus $L = 5$. According to the previous section, we proceed:

Step 1.- $\Phi_F = A_0\, B_0$ and the minterms of Φ_F in 5-bit string format are:

$$M_{\alpha_1} = A_0\, B_0 = 001\ 01$$

H is initialized $H_0 = \oslash$.
Step 2.- For $i = 1$: realize $H_i = H_{i-1} + M_{\alpha_i} - 2 \cdot MD\,(M_{\alpha_i}, H_{i-1})$.
The final value H_1 obtained at the end of the loop is:

$$H_1 = H_0 + M_{\alpha_1} - 2 \cdot MD\,(M_{\alpha_1}, H_0)\,.$$

$$H_1 = \oslash + 001\ 01 - \oslash = 001\ 01.$$

Step 3.- Calculation of the number of $1's$ from H_1. Thus,

$$\mathcal{U}_F = 2^{L_A - 1}\, 2^{L_B - 1}. \tag{8}$$

For lengths of the LFSRs in a cryptographic range $L_i \approx 64$ the number of $1's$ in the output sequence is $\simeq T/2$. Consequently, the generated sequence is quasi balanced.

6 Conclusions

An easy and efficient method of computing the degree of balancedness in the output sequence of key-stream generators has been presented. From the handling of bit-strings of low length, it is possible to derive general expressions for the number of $1's$ ($0's$) in the output sequence for any kind of LFSR-based generator found in the literature. Then, the obtained values are compared with the expected values for a sequence to be balanced. The result of this comparison implies the assessment (acceptance/rejection) of such a sequence generator. Remark that the particular form of these generators allows one the application of the computational procedure with negligible time and memory complexity. The method here developed can be considered as a first selective criterium for acceptance/rejection of this type of generators of cryptographic application.

References

1. Bluetooth, *Specifications of the Bluetooth system,* Version 1.1, February 2001, available at http://www.bluetooth.com/
2. D. Coppersmith, H. Krawczyk, Y. Mansour, "The Shrinking Generator", CRYPTO'93, LNCS Springer Verlag, Vol. 773, 1994, pp. 22-39.
3. CRYPTREC project- cryptographic evaluation for Japanese Electronic Government, www.ipa.go.jp/security/enc/CRYPTREC/index-e.html
4. S.W. Golomb, Shift Register-Sequences, Aegean Park Press, Laguna Hill, 1982.
5. GSM, *Global Systems for Mobile Communications,* available at http://cryptome.org/gsm-a512.htm
6. R. Lidl and H. Niederreiter, Introduction to Finite Fields and Their Applications, Cambridge: Cambridge University Press, 1986.
7. D. Mange, Analysis and Synthesis of Logic Systems, Artech House, INC., Norwood 1986.
8. A.J. Menezes *et al.*, Handbook of Applied Cryptography, New York:CRC Press, 1997.
9. R.A. Rueppel, Analysis and Design of Stream Ciphers, New York:Springer-Verlag, 1986.
10. G.J. Simmons (ed.), Contemporary Cryptology: The Science of Information Integrity, New York:IEEE Press, 1991.
11. L. Wang, A.E. Almani, Fast conversion algorithm for very large Boolean functions, Electron. Lett. 36 (2000) 1370-1371.

Secure Electronic Payments in Heterogeneous Networking: New Authentication Protocols Approach

J. Torres, A. Izquierdo, A. Ribagorda, and A. Alcaide

Universidad Carlos III de Madrid, Spain
{jtmarque, aizquier, Arturo, aalcaid}e@inf.uc3m.es

Abstract. Recent research efforts have been addressed towards maintain the heterogeneous networking transparent under the powerful all-IP concept. As example, current global standarization initiative specifies the 3G cellular system to wireless LAN inter-working. On the other hand, smart cards are presented as enough powerful devices capable to perform a strong authentication at lower layers of the protocol stack. Our work proposes a novel model reference and a scenario of applicability for secure electronic payment in this environment. Impact on the trust relations are assesed and a set of authentication requirements are provided. Finally, a new approach based on end-to-end layer 2 authentication protocols is adjusted to this proposal, considering the most interesting improvements in the authentication mechanisms applicable to this context.

1 Introduction

In then recent years the word heterogeneity has been wide applied and from diverse perspectives in the publications and works on information and communication technologies topics. Research efforts have been dealt towards maintain the heterogeneous networking transparent and hidden under the powerful all-IP concept. By means of that, technologies of different nature, such as satellite communications, backbone internet, wired and wireless LANs, and cellular networks could cooperate together replying with a wide spectrum of solutions and added-value services to the business model demands. Consequently, this fact has an impact on the development growth of multi-mode terminals capable to inter-work to a variety of access networks.

Over that landscape, our work pays special attention to the recent initiative carried out by the 3G Partnership Project in order to specify the 3G cellular system to wireless LAN inter-working [1], and more specifically to the security aspects [2].

The considerations to deploy this case of heterogeneous networking could be summarised in:

a) Since WLAN terminals (STA) are being massively integrated in normal-life and in a diversity of environments, they are far for being considered exclusively for sophisticated users. Decrease of the devices size and power consumption in contrast to the increase of the computation capacity, improvement of security protocol performance [3, 4] and multi-mode radio links capabilities [5].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 729 – 738, 2005.

b) Combining both technologies [6], coverage and data rates suffer an important growth comparing to the capacity of the next future cellular access networks. As target, the WLAN is seen as a complement to the 3G systems to deliver enhanced value to the end-user. Obviously, roaming solutions, vertical and horizontal handoffs among others require special considerations. The known security gaps, mentioned above in this document, could be the weakness aspect of this solution, therefore is object of several studies and part of this work review them.

c) Business model are pending to be well-defined in order to include in this scenario the Wireless Internet Service Provider, WISP functionalities. New revenue from access and services could be disputed between the three parties: WLAN operator, WISP and cellular operator.

d) A global standardization process [1] supports this initiative integrating the interests of vendors, operators, manufacturers and service providers. This in-progress standard proposal could be considered as a step in the roadmap to the 4G systems.

Firstly, our work proposes a novel model reference and an example of applicability for secure electronic payment in heterogeneous networking environment. The detailed description of such scenario is presented in the next section. Afterwards, section 3 is devoted to the impact of the new participant entities on the Trust Model and consequently to trust relations. In order to guarantee the security levels in section 4 a set of authentication requirements are derived. In section 5 a new authentication protocols approach is introduced in order to complain such requirements and includes a network architecture proposed in our study.

2 Heterogeneous Networking

2.1 Reference Model

Our work takes as start point the reference model defined in [1] and more concretely the scenario[1] 3 with non-roaming features. Afterwards it redefines an enhanced model adding a new entity with authentication capabilities in user-side, here named Supplicant Equipment (SE). In our reference model in Fig. 1, SE exclusively interacts with the system through the WLAN User Equipment (WLAN-UE) by means of the Reference Point At, R_{At}. The User Equipment might exchange messages with the SE over the R_{At} in order to initiate the session and gain access to the rest of the system, establish data traffic transmission/reception and finalize the session. Details of these low-level messages are out the scope of this work.

The 3GPP AAA server retrieves authentication information and subscriber profile (including subscriber's authorization information) from the HLR/HSS. Afterwards, it communicates authorization information to the WLAN and Packet Data Gateway, PDG. It should inform to PDG about the authorized W-APN, necessary keying material for tunnel establishment and user data traffics.

[1] 3GP Project envisages a gradual complexity of the potential scenarios. Scenario 3 is included in Release 6. This standarization process will be completed with a total of 6 scenarios.

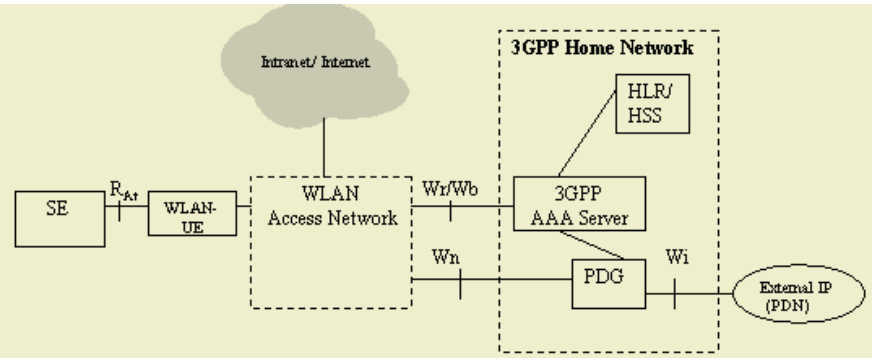


Fig. 1. Reference Model

The WLAN routing functionalities address packets through PDG and vice versa, providing WLAN UE with 3G Packet Switched based services. External IP network services are accessed via PDG by means of a authorisation process, service selection (e.g. Wireless-Access Point Name selection, W-APN access point for a concrete service) and subscription checking. PDG functionalities includes IP address resolution and allocation.

Some of the envisaged functions for the WLAN-UE are i) associating to an I-WLAN, ii) WLAN access authentication based on EAP methods, iii) building an appropriate NAI, iv)obtain a local IP address. And specifically for the scenario 3 v) building an appropriate W-APN, vi) request the resolution of a W-APN to a PDG address and establish a secure tunnel to the PDG, vii) obtain a remote IP address in order to access services provided in the operators PS domain.

From our perspective, and such as is mentioned the set of WLAN-UE functions must include the provision to SE of relay access to the rest of the system.

2.2 Secure Electronic Payments Scenario

In figure 2, an example of scenario of applicability based on our reference model is shown. In this scenario, a smart card featured as a credit/debit card (assuming the presence of a Cardholder) plays the role of Supplicant Equipment, SE. The referred WLAN-UE could be materialized by a Wireless Point of Sale, WPOS, featured as a mobile internet device and provisioned with a SIM/USIM registered with the HLR at the 3G Home PLMN. In our context, the stated 3GPPP AAA server (e.g. RADIUS Server) should be under the Payment Operator control, since it is the responsible entity for authenticating the WPOS in its domain. Agreements between payment operator and cellular operator should be envisaged. This aspect tightly concerns to the conception of the business models.

Although the WPOS is located at merchant facilities, the security of the payment procedure is payment operator's responsibility; therefore Access Point (AP) and a potential WLAN AAA Proxy should remain transparent in the end-to-end authentication process. However, the need of a payment card in this realistic scenario introduces complexity. Thus, for security reasons we propose that the WPOS should be remain transparent in the complete payment process, playing the role as tunnelled

entity avoiding risks of potential attacks from a manipulated point of sale terminal. This novel perspective is clearly different of the current envisaged process for smart cards payments [7]. The performance for electronic payment through a WPOS could be derived in the following summarised manner:

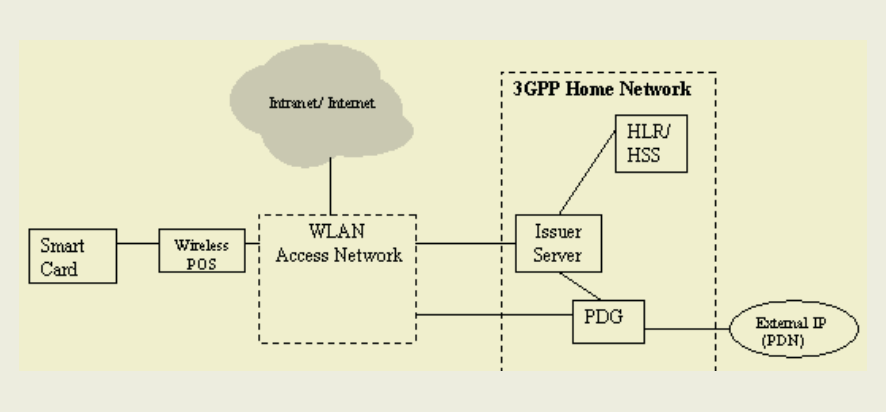


Fig. 2. Proposed Scenario

Once the Cardholder inserts the debit/credit smart card in the card reader, the Cardholder Client System integrated in the WPOS initialises the chip card and the smart card application, then it communicates with them in order to obtain a cryptogram from the smart card for validation.

In the course of communicating with the chip card several exchanges occur. The chip card may request that the Cardholder enter a PIN. If so then the Cardholder Client System notifies the Cardholder using the display and activates or connects, if necessary, the PIN pad or key entry device to permit and capture PIN entry. The application on the chip card returns one or several cryptograms for validation.

While sending messages (commands) and receiving responses from the card in the previous steps, the Cardholder Client System was gathering the information necessary to compose the Authentication Response which it now sends to the Issuer Server (issuer bank domain). Finally, the Issuer Server validates cryptogram sent from Cardholder Client System software.

In our opinion, at the glance of this performance, the inclusion of improved authentication procedures should be required. Our work proposes partially redefine the functionalities of these entities in this address, in order to secure the payment in our scenario. Therefore, once the WPOS device is authenticated by payment operator then it is secure to accept a payment card. As mentioned, the smart card exploit the Supplicant functionality in a generic authentication context. Thus, to perform the payment the smart card by means of the credentials inside referred to the cardholder should be authenticated directly by the issuer bank server and vice versa. In our proposed scenario, the 3GPP AAA server is represented by such Issuer Server.

The envisaged functions for the smart card are: i) store and execute the application, ii) communicate with Cardholder Client System software (WPOS), iii) validate cardholder PIN and, iv) generate and return cryptogram upon request.

On the other side, the functions provided by Cardholder Client System should be: i) communicate with issuer server, with smart card and with Cardholder for Card insertion/removal and PIN entry, ii) exchange of authentication messages both smart card and issuer server, and iii) relay of the cryptogram from smart card to the issuer server.

The Issuer Server (authentication server) functions could be summarized: i) securely stores keys needed for cryptogram validation, ii) collect the necessary data to perform smart card authentication processing and initiate this procedure, iii) validate cryptogram sent from Cardholder Client System software at Hardware Security Module (HSM).

3 Trust Model

Given the scenario proposed in this contribution is possible to determine a high-level conceptual model where the involved entities and relations between them are evaluated in terms of trust. For simplicity, four entities (commonly just three entities [8]) have been considered in order to highlight the relevant issues in the context of this work. In figure 3, the novel Trust Model is represented.

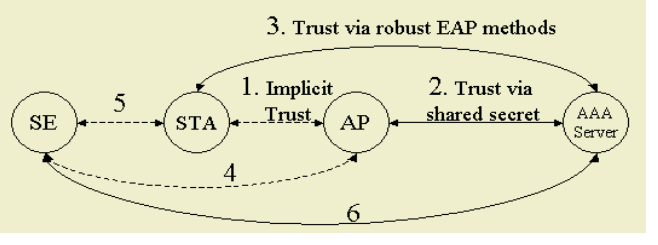


Fig. 3. Trust Model

The essential difference comparing to [8] is the Supplicant Equipment (smart card in our scenario). New trust relations, interfaces 4, 5 and 6 have been established. For detailed description of the numbered interfaces 1 and 2 see [8]. In [9] recent improvements on the original proposal referred to interface 3 can be found.

The relation trust in the interfaces 5 and 6 is a target issue of this work and in the next section the security requirements and specifically the authentication requirements are described in order to obtain a significant level of trust. Once this occurs, an implicit trust is attributed to interface 4.

Given the Trust Model depicted in figure 3, the trust relation at interfaces 4, 5 and 6 is studied. The analysis of the trust relation between the involved entities carried out by this work addresses to the next considerations:

a) Since the STA (WPOS) is not a device located at the user domain (commonly referred as Issuer Domain) but it should belong to the merchant domain (commonly referred as Acquirer Domain), and more over, the cardholder could enter her/his card

in the unknown WPOS world wide, the trust relation represented by the interface 5 should be assessed. The WPOS could be manipulated in order to perform different types of attacks against the smart card (e.g. MitM, DoS, etc.). Thus, the proposal of this work is to define the WPOS as a transparent mean for relaying, avoiding the interference during the smart card authentication process. As mentioned, the WPOS should have been previously recognized as a trust terminal after a correct authentication process (implicitly including AP and other intermediate equipments) by the AAA Server (Issuer Server) and finally the cardholder might trust in a secure electronic payment. Tunnelling authentication should be investigated and the communication between SE (Smart Card) and STA (messages/ commands exchange) should be redefine in this sense.

b) The trust relation 6 between the Suplicant Equipment (SE) and the AAA Server is featured by the fact that SE is directly or indirectly associated to a concrete card-issuer bank and therefore its AAA Server (Issuer Server) might obtain by first-hand the Primary Account Number (PAN) for the card to be authenticated from an out of band channel, such as its own cardholder database or from other allowed environment. Thus, tunnelling authentication mechanisms between SE and AS with transparent relaying on the intermediate entities should be investigated.

c) The trust relation 4 between SE and AP is considered as an implicit trust if we establish an end-to-end (SE -- AAA Server) authentication procedure as mentioned in b) and trust in the relation AP--AAA Server is guaranteed. Obviously, it is necessary to comment in this point that the value of this trust relation irrespective of the WLAN-UE --AP trust relation.

The new perspective derived from presented Trust Relations represents an impact not only on technological aspects but moreover on the business model as well. The cellular operators could be obtain important benefits, serving the major part of this infrastructure, offering a secure networking service. Furthermore, other additional services could be derived from this layout.

4 Authentication Requirements

In order to make effective the reviwed Trust Model over the proposed scenario, a set of correct authentication mechanisms are required, as described before, enforcing secure payment considering the smart card (SE) and Issuer Server (AS) the extremes of an end-to-end authentication procedure, minimizing the potential attacks from hosts in the transaction route. Based on this goal, our work detected a lack of authentication requirements to be suited to this scenario, in order to obtain major security guaranties.

Requirement 1: end-to-end security over this heterogeneous architecture with multiple nature hosts, capable to be origin or bridge of attacks, attempting against the security guarantee of the whole system. As is claimed, smart card and Authentication Server and should be the end-entities in this authentication scheme.

Requirement 2: mutual authentication once mechanisms for session key establishment have been provided. An attack in this scenario could be seen from both smart card perspective and the Authentication Server one. For instance, the server could be supplanted in order to obtain critical information from the smart card, therefore the server should be correctly authenticated by the smart card, representing to the cardholder. On the other side, the smart card could be used in a fraudulent manner by a malicious user. The repercussion of that in bank card is more than relevant. The mechanisms for session keys establishment should guarantee the freshness of them.

Requirement 3: Protected and flexible cipher-suite negotiation. This could be considered as a general purpose requirement since is defined to cover the different protocol implementations/versions and the end-to-end participant capabilities. But more specifically in our work, this requirement is introduced in order to adjust the cipher-suite complexity or security degree depending on the payment usage context (e.g. wired/wireless, micropayment, mobility context , etc.)

Requirement 4: scalable key management capabilities. Obviously, the approach proposed is an direct and strong impact on the server capability in order to attend requests coming from an important number of the on-line smart cards.

5 New Authentication Protocols Approach

The conception of this inter-working system is addressed towards an all-IP environment, therefore and as is stated in [1] the mechanisms in order to guarantee the security must deployed when feasible along the whole protocol stack taking into account both signalling and data traffic planes. These mechanisms must be irrespective of others implemented in upper or lower layers.

Our research is focused to the layer 2 authentication, but not preclude additional secure mechanisms adopted at IP level (where mobile-IP could be considered, from the WLAN-UE perspective) or transport level, remaining out the scope of this work. Motivation on this focus is briefly stated:

- security strengthening, irrespective of adopted solutions in upper/lower layers.
- to guarantee a strong authentication in absence of IP connectivity. Current limitations in smart cards show them as enough powerful devices capable to perform a strong authentication at lower layers but insufficient to perform protocols at IP and upper layers.
- payment transaction session (messages exhange) is relatively single from the payment card point of view. Likely solutions at IP level are not necessary (other financial application such as SW downloading, customisations, etc. should be desirable in the next future [10]).
- an implementation of an authentication infrastructure [11] with efficient key distribution mechanisms [12] would be easily well-suited to our model. Wireless LAN roaming aspects will be issues to be included in further work.
- important initiatives based on EAP methods have been covered by IETF, IEEE and interested parties.

Far from a detailed analysis of the potential specific protocols to be implemented, our work aims to feature a new layer 2 authentication protocols approach on the basis of the reference model in figure 1 and as result a novel protocol stack architecture is proposed and depicted in figure 4.

One of the approaches of this work is to exploit the smart card capabilities as embedded device in terms of lower layer protocol implementations and secure storage, highlighting the differences with the conventional magnetic band card to perform electronic payments. A better performance including IP-like techniques in the next generation smart cards is foreseeable. As mentioned before, our work foresees an autonomous authentication smart card procedure before the cardholder enters the PIN. Therefore, could be considered as an embedded device authentication based on the credentials stored in the card. The user authentication is completed with something that he/she knows.

It is not in the scope of our work to define a specific EAP-type for the architecture in fig. 4, but establish the authentication relationship between the participant entities. However, for our new approach, the recent standardization initiative [13] for EAP support in smart card should be considered. Its goal is defining universal ISO 7816 interface, supporting most of EAP authentication protocols. Components such as a logical network interface that directly processes EAP messages and allows EAP profiles (types) definition, an operating system interface to provide identity management, storage of cryptographic material, and management interface are envisaged. As result of this effort, in [14] an specific type, EAP-SC defines an EAP method and multiplexing model for the support of smart card based authentication methods. EAP-SC provides for encapsulation of other EAP methods such as EAP-TLS among others.

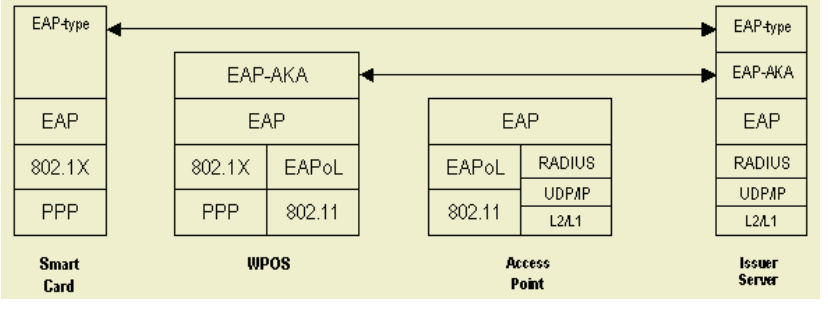


Fig. 4. Procotol Stack Architecture

From the point of view of the authentication method, the wireless POS behaves as standard WLAN User Equipment. In particular, the protocols EAP-SIM [14] and EAP-AKA [15] should be considered. Although these protocols have been included in the 3GPP specifications, weakness are identified in [15, 16] and improvements should be investigated. In [17,18] some signalling UMTS/GPRS messages are embeeded within EAP messages facilitating fast handover. In [9] the authors propose to combine the well-suited to this heterogeneous environment performance of EAP-AKA with strong authentication of TLS protocol.

6 Conclusion

In a non-roaming 3G cellular system and wireless LAN interworking scenario, the smart cards has been presented as enough powerful devices capable to perform a strong authentication at lower layers of the protocol stack. A novel reference model and a scenario of applicability for secure electronic payment in this environment has been introduced. Impact on the trust relations has been assesed and a set of authentication requirements has been provided in orden to guarantee the security process. Finally as result, a feasible new approach based on end-to-end layer 2 protocols is well-adjusted to this proposal, considering the most interesting improvements in the authentication mechanisms applicable to this context. Further work will consider roaming scenarios and the interaction with other authentication protocols at upper layer, expecting a best performance of smart cards.

References

1. 3GPP TS 23.234 v6.2.0: 3GPP system to Wireless Local Area Network (WLAN) Interworking; System Description, September 2004
2. 3GPP TS 33.234 v6.3.0; 3GPP system to Wireless Local Area Network (WLAN) Interworking Security System, December 2004
3. Argyroudis, P.G., Verma, R., Tewari, H., O'Mahony, D.: Performance Analysis of Cryptographic Protocols on Handheld Devices. NCA 2004: 169-174
4. Gupta, V., and Gupta, S.: Experiments in wireless internet security. Proc. IEEE Wireless Communications and Networking Conference, WCNC, March 2002, Vol. 1, 859-863
5. IST Project Brain. Broadband Radio Access for IP based Networks (IST-1999-10050), 2001
6. IST Project Wine. Wireless Internet Networks. (IST-1999-10028), 2001
7. 3D –Secure Functional Specification, Chip Card Specification v1.0, Visa Corp., August 2001
8. IEEE Standard for Information technology, 802.11i-2004 Amendment to IEEE Std 802.11i/D7.0: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Specification for Enhanced Security, 2004
9. Kamborakis, G., Rouskas, A., Hormentzas, G., Gritzalis, S.: Advanced SSL/TLS-based authentication for secure WLAN-3G interworking, IEEE Proc. Communications, Vol. 152, No. 5 October 2004
10. Lagosanto, L.: Java Card evolution toward service-oriented architectures Fifth e-Smart Conference, Sophia Antipolis, France, September 2004
11. Shi M., Shen X., Mark, J.W.: IEEE802.11 Roaming and Authentication in Wireless LAN/cellular mobile networks, IEEE Wireless Communications, August 2004
12. Salgarelli, L., Buddhikot, M., Garay, J., Patel, S., Miller, S.: Efficiency authentication and key distribution in wireless IP networks , IEEE Wireless Communications, December 2003
13. WLAN Smart Card Consortium Specifications, www.wlansmartcard.org, Sept. 2004
14. Urien, P. et al., EAP Smart Card Protocol, IETF Internet Draft <draft-urien-eap-smartcard-type-00>, September 2004
15. Haverinen, H. and Salowey, J: EAP-SIM Authentication, IETF Internet Draft <draft-haverinen-pppext-eap-sim-16>, December 2004

16. Arkko, J. and Haverinen, H.: EAP-AKA authentication, <draft-arkko-pppext-eap-aka-15>, December 2004
17. Salkintzis, A.K.: Interworking techniques and architectures for WLAN/3G integration toward 4G mobile data networks, IEEE Wireless Communications,Volume: 11 , Issue: 3, 50 – 61, June 2004
18. Salkintzis, A., C. Fors and R. S. Pazhyannur, WLAN-GPRS Integrationfor Next Generation Mobile Data Networks," IEEE Wireless Communications, vol. 9, no. 5, pp. 112-124, Oct. 2002

Software Reliability Measurement Use Software Reliability Growth Model in Testing

Hye-Jung Jung[1] and Hae-Sool Yang[2]

[1] Department of Information StatisticsPyongTack University
PyongTack-City, Kyonggi, 450-701, Korea
jhjung@ptuniv.ac.kr
[2] Graduate School of Venture, HoSeo Univ. Bae-Bang myon, A-San,
Chung-Nam,336-795, South Korea
hsyang@office.hoseo.ac.kr

Abstract. In this paper, we study the software reliability measurement method of reliability testing metrics. Software reliability is very important. But, it is very difficult to test for software reliability measurement. So we describes the software reliability metrics for ISO/IEC 9126, and we introduce Gamma-Lomax software reliability model for multiple error debugging. And we calculate the software reliability measure(reliability metrics, parameter estimation etc). We introduce the measurement method of software reliability quality of software product.

Keywords: Software reliability quality metrics, ISO/IEC 9126, Gamma-Lomax Software Reliability Growth Model, Quality Metrics.

1 Introduction

Software system have been applied into the various fields of science, business, society, our life and its importance and necessary has been raised and recognized. The software system environments have been enlarged, that is the range of applied fields become wide and the technology changed rapidly. As a result software system seem more complex. Also, software reliability is most important part of business in the late 20^{th} century. It is a time to study on testing the software developing techniques and research on the software reliability for the security of the software products. We have to measure the software reliability quality of the product produced and use that information to improve the process producing it. Software reliability quality assurance for software development needs to be bound up with good measurement.
The rest of this paper is arranged as follows: Chapter 2 describes the commonly used software reliability attributes and currently available metrics for measuring software product quality.

Chapter 3, we introduce a Gamma-Lomax software reliability growth model(SRGM) for the multiple software errors debugging in the software testing stage. This model is the modification of Jelinski-Moranda model incorporated de-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 739 – 747, 2005.

pendence among the interfailure times for multiple software errors. And we evaluate the software reliability measures by the method of maximum likelihood estimation.

Chapter 4, we introduce the concepts of the software reliability testing method. Also, we introduce the result of the simulation study for software reliability measure.

2 Software Reliability Metrics

Software reliability is important part of the chief industries in the late 20th century. Integral to the development of software is the process of detecting, locating, and correcting bugs. In spite of these effective more common types of software nonperformance include the failure to conform to specifications or standards. Software testing tools are available that incorporate proprietary testing algorithms and metrics that can be used to measure the performance and conformance of software. But, development of standard testing tools and metrics for software testing could go a long way toward addressing some of the software testing problems that plague the software industry. Improved tools for software testing could increase the value of software in a number of ways:

(1) reduce the cost of software development and testing
(2) reduce the time required to develop new software products
(3) improve the performance, interoperability, and conformance of software.

MaCall, Richards, and Walters(1977) attempted to assess quality attributes for software. Boehm(1978) introduced several additional quality attributes. As software changed and improved and the demands on software increased, a new metrics of software quality attributes was needed. In 1991, the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) adopted ISO/IEC 9126 as the standard for software quality. The ISO/IEC 9126 is now widely accepted. ISO/IEC 9126[1] consists of the external metrics, internal metrics. It categorises software quality attributes into six characters (functionality, reliability, usability, efficiency, maintainability, and portability), which are further subdivided into subcharacteristics. The subcharacteristics can be measured by internal or external metrics. An reliability character in external metrics should be able to measure related to the software reliability model. Software reliability models enable to make estimations and predictions for the reliabilities of software systems by applying the software failure data and experiences obtained from the previous experiments for the similar software systems to the developing target software system. The reliability measurement consist of the maturity metrics, fault tolerance metrics, recoverability metrics, reasbility compliance metrics. The maturity measurement of the software product to avoid failure as a result of faults in the software. Fault tolerence measurement of the software product to maintain a specified level of performance in cases of software faults of its specified interface. The recoverability measurement of the software product to re-establish a specified level of performance and recover the data directly affected in the case of a failure. The reliability compliance capability of the software product to adhere to standards, conventions or regulations relating to reliability. Testing was seen as a necessary process to prove to the final user that the product worked. An improved instrastructure for software testing has the potential to affect software developers and users by removing bugs before the software product is released and

detecting bugs earlier in the software development process and locating the source of bugs faster and with more precision. The product analyzed in this paper is a data through the survey. Respondents were asked to answer the satifaction of the software product quality on the seven categories rating scale: “Very Very Dissatified”, “Very Dissatified”, “Dissatisfied”, “Neutral”, “Satisfied”, “Very Satisfied”, “Very Very Satisfied” Also, Respondents were asked to answer the importance of the software product quality on the same method. Characteristics and subcharacteristics of software product quality have been measured with the same rating scale. In order to anaslysis the difference influence of characteristics, and subcharacteristics on overall satisfication and importance due to user difference was collected including application domain, training, age, similar software experience, and respondent’s job type(developer, user). We analysis the result using the SPSS/PC statistical package program.

3 GAMMA-LOMAX Software Reliability Model

In 1970’s, researchers have studied the modeling of software reliability. The Jelinski-Moranda(1972)[9] model for software reliability growth model(SRGM) is one of the most commonly cited models. The main property of the model is that the intensity between two consecutive failure is constant. And the distribution of interfailure time T_i at the i-th testing stage is given by

$$f(t_i)= (N-i+1)\ \phi \exp((-N-i+1)\phi t_i) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.1)$$

where the parameter $\lambda_I=(N-i+1)\phi$ is a failure rate, and N is the number of latent software errors before the testing starts, and ϕ is the failure intensity contributed by each failure. The parameter ϕ and N in Jelinski-Moranda model was estimated by maximizing the likelihood function.

The software reliability growth model suggested by Littlewood and Verrall(1973)[10] is perhaps the most well known Bayesian model. The Littlewood-Verrall model assumes that interfailure times are exponential distributed random variable with the density function.

$$f(t_i|\ \lambda_i)= \lambda_i\ \exp(-\lambda_i t_i) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.2)$$

where λ_i is an unknown parameter whose uncertainty is due to the randomness of the testing and the random location of the software errors. And the probability density function of λ_i is

$$f(\lambda_i \mid \alpha, \varphi(i))= \frac{(\varphi(i)^{\alpha}\ \lambda_i^{\ \alpha-1} \exp(-\varphi(i)\lambda_i))}{(\Gamma(\alpha))} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.3)$$

where $\varphi(i)$ is depending on the number of detected error. Usually, $\varphi(i)$ describes the quality of the test and it is a monotone increasing function of i.

Langberg and Singpurwalla(1985)[12] presented a shock model interpretation of software failure and justified the Jelinski-Moranda model.

In the Langberg and Singpurwalla Bayesian model the parameters in the Jelinski-Moranda model are treated as random variables.

Nayak(1986)[13] proposed a model for incorporating dependence among the detection times of software reliability measures. Many software reliability model have been studied . But they treated with the case of software reliability growth model for single error debugging at each testing stage until now. Jung[20] studied the software reliability model with the multiple errors debugging.

In this section, we introduce the software reliability growth modeling and the parameters in the introduced software reliability growth model for the multiple errors debugging at each testing stage. This model is condition under the incorporated dependence for multiple errors.

The model introduced by Jung(1994)[20] for describing multiple errors at each testing stage assumes that

(1) The multiple software errors at the i-th testing stage occur n_i times for i=1,2,…,m , and the last testing stage, m is unknown and fixed constant, and $N=\sum_{i=1}^{m} n_i$

(2) All of the detected software errors in each testing stage are perfectly debugged at once before the next testing stage begins

(3) The multiple software interfailure times, $T_i = X_{(i)} - X_{(i-1)}$ i=1,2,….,m have independent gamma distribution with parameters n_i and $\lambda_i = (m-i+1)\phi$, where $0= X_{(0)} \le X_{(1)} \le ,\ldots, \le X_{(m)}$ are the ordered failure times and the software failure intensity ϕ is unknown

The test environments change the software failure rates of all testing stages by a common parameter η, so that the random multiple software interfailure times T_1, T_2, …, T_m.

Of a software system are independently gamma-distributed with multiple software failure rates $\eta\lambda_1$, $\eta\lambda_2$,…, $\eta\lambda_m$ respectly, under the software test environments in assumptions (1), (2), and (3). Assuming that η is a gamma random variable as G(a,b) then the unconditional joint probability density function $f(t_1, t_2, \ldots, t_m)$ of interfailure times $T_1, T_2, \ldots, T_m$ occurring $n_1, n_2, \ldots, n_m$ times respectively is given by

$$
\begin{aligned}
&GL_m,(a, b; \lambda_1, \lambda_2, \ldots,\lambda_m ; n_1, \ldots, n_m) \\
&=(\prod_{i=1}^{m} t_i^{n_i-1} \lambda_i^{n_i} /\Gamma(n_i)) \\
&\times(b\, \Gamma(\sum_{i=1}^{m} n_i +a)/\Gamma(a)) \\
&\times(1 / \sum_{i=1}^{m} \lambda_I t_i+b)^{\sum_{i=1m} n_i +a} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.4)
\end{aligned}
$$

and Jung[20] introduce this distribution as a Gamma-Lomax Software Reliability Growth Model which is an extension case of the multivariate Lomax distribution of Nayak[13].

The marginal probability density function of the multiple software interfailure times $T_1, T_2, \ldots,T_m$ until r(≤m)-th testing stage is $GL_r,(a, b; \lambda_1, \lambda_2, \ldots,\lambda_r ; n_1, \ldots, n_r)$ where λ_i is the failure rate of the i-th testing stage.

We will use the method of maximum likelihood in order to estimate the unknown parameters m, ϕ, n_{r+1} of the software reliability. The number of multiple software errors detected until r-th testing stage is denoted by $\sum_{i=1}^{r} n_i$, which will be called sample size. Suppose that $\Theta=\{ t_1, t_2, \ldots,t_r \}$ is a tested data set of the multiple software interfailure times.

The parameters m, ϕ, n_{r+1} of the software reliability can be estimated by miximizing the likelihood function. Let the observed r ordered multiple software failure times be $x_{(1)}, x_{(2)}, \ldots, x_{(r)}$ and let τ be the specified time $\tau \in (x_{(r)}, x_{(r+1)},)$.

And, we let $t_i,= x_{(i)} - x_{(r-1)}$, i=1,2,….r and $T_{r+1},> \tau - x_{(r)},$. We estimate the parameters as follows procedures;

[ML1]
We calculate the likelihood function of m and ϕ given Θ from equation (3.4).

$$
\begin{aligned}
&L_1(m, \phi \mid \Theta) \\
&=(\prod_{i=1}^{m} ((m-i+1)\phi)^{n_i} t_i^{n_i-1})/\Gamma(n_i)) \\
&\times(b^a \Gamma(\sum_{i=1}^{r} n_i + a))/\Gamma(a) \\
&\times(1/\sum_{i=1}^{m}(m-i+1)\phi t_i + b)^{(\sum i=1r\, ni + a)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.5)
\end{aligned}
$$

[ML2]
We obtain the conditional probability density function of t given Θ.
[ML3]
We combine (3.5) and conditional probability density function in [ML2] and get the likelihood function of m, ϕ, and $n_{r+1.}$

$$
\begin{aligned}
&L_2(m, \phi, n_{r+1} \mid \Theta) \\
&=(\prod_{i=1}^{m} ((m-i+1)\phi)^{n_i} t_i^{n_i-1})/\Gamma(n_i)) \\
&\times((b^a/\Gamma(a)) \times(((m-r)\phi)^{nr+1}/\Gamma(n_{r+1})) \\
&\times((t^{nr+1-1}\Gamma(\sum_{i=1}^{r+1} n_i + a))/((\sum_{i=1}^{r} \\
&(m-i+1)\phi t_i + b + (m-r)\phi t)^{(\sum i=1r+1\, ni + a)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.6)
\end{aligned}
$$

Hence we have the logarithm in the above likelihood function of (3.6) by

$$L(m, \phi, n_{r+1}) = \mathrm{lm}\, L(m, \phi, n_{r+1})$$

[ML4]
We take the partial derivative of the log likelihood function with respect to m, ϕ, n_{r+1}, respectively, and equate them to zeros. Then we have the normal equation (3.7).

$$
\begin{aligned}
&\partial L(m, \phi, n_{r+1})/\partial m = 0 \\
&\partial L(m, \phi, n_{r+1})/\partial \phi = 0 \\
&\partial L(m, \phi, n_{r+1})/\partial n_{r+1} = 0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.7)
\end{aligned}
$$

[ML5]

We arrange the simultaneous linear equation induced from (3.7) with respect to to m, ϕ, n_{r+1}, respectively.

$$
\begin{aligned}
&(\sum_{i=1}^{r+1} (n_{r+1})/(m-i+1)) \\
&=((\sum_{i=1}^{r+1} n_i)(\sum_{i=1}^{r} t_i + t))/((\sum_{i=1}^{r} (m-i+1)\, t_i + (m-r)t)) \\
&\ldots\ldots\ldots\ldots \qquad \ldots\ldots (3.8) \\
&\phi = \exp(\Gamma'(\sum_{i=1}^{r+1} n_i + a)/\Gamma(\sum_{i=1}^{r+1} n_i + a) \\
&\Gamma'(n_{r+1}/\Gamma(n_{r+1})) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.9)
\end{aligned}
$$

where

$$\Gamma'(x)/\Gamma(x) = -\gamma + (1-(1/x)) + ((1/2) - (1/(x+1))) + ((1/3) - (1/(x+2))) + \ldots + ((1/n) - (1/(x+n-1))) + \ldots,$$

and

$$-\gamma = \gamma'(1) = \int \exp(-x) \ln(x)\, dx.$$

Where the range of intergral is 0 to infinity.

According to the maximum likelihood estimation procedure [ML1] to [ML5], the approximations of estimates m, ϕ, n_{r+1}, can be obtained by solving the numerical algorithms, respectively.

Therefore substituting these estimates into the multiple software reliability (3.7), (3.8), and (3.9), we can obtain the maximum likelihood estimates of the multiple software reliability by the result of Zehna.

4 Reliability Metrics

An external software reliability metric of ISO/IEC 9126 should be able to measure attributes related to the software failure times. Software reliability consists of the maturity metrics, fault tolerence metrics, recoverability metrics, reliability compliance metrics. Also, Maturity metrics consist of the estimated latent failure density , estimated latent fault density , failure density etc.

But, we can calculate this measure by using the software failure time distribution . It is difficult to get the software failure time data. So we try to the survey for measuring of software reliability metrics.

We present our result the number of failure during one month. For example, the questions as follows;

The product analyzed in this paper is a data through the survey. Respondents were asked to answer the satifaction of the software product quality on the seven categories rating scale: “Very Very Dissatified”, “Very Dissatified”, “Dissatisfied”, “Neutral”, “Satisfied”, “Very Satisfied”, “Very Very Satisfied”

How many failures were detected during one month?
How many fault were detected during one month?
How many failure condition are resolved?
How often the software product cause the break down ?

Also, we predict the quality of software product by using the satisfaction survey.
If the causes of failure occur according to a Poisson distribution. The times between failures are independent exponential variables, and the data constitutes a random sample of sample size from exponential with parameter $1/\phi$.

A variable or distribution would arise if one discusses the number of occurrences in some time interval. For each interval the number of failures follows a poisson distribution with parameter $\mu = \phi t$.

For example , specifically the Goel & Okumoto model assumes that the failure process is modeled by Nonhomogeneous Poisson Process model with mean value function m(t) given by

$$m(t)=a(1-\exp(-\phi t)),\ a>0,\ b>0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(4.1)$$

where a and b are parameters to be determined using collected failure data.

Note that for Goel & Okumoto model[4], we have

$$m(\infty)=a, \text{ and } m(0)=0.$$

Since $m(\infty)$ is the expected number of faults which will eventually be detected, the parameter a is the final number of faults that can be detected by the testing process.
A company examine the as follows .

Table 1. Frequency Table

Number of failure	Frequency	Probability
0	8	0.21052632
1	15	0.39473684
2	7	0.18421053
3	7	0.18421053
5	1	0.02631579

The value of ϕ is unknown, but sample mean is a good estimate of $E(T)=1/\phi$.

We estimate the mean, parameter a, and ϕ after the one month test period.

Table 2. Mean and Estimation of parameters after the thirty test day

Mean	1.447368421
ϕ	0.048245614
a	71.91313439

We estimated the mean, parameter a, and ϕ using the homogeneous poisson distribution. So we can calculate the value of mean value function m(t) and hazard rates r(t) on based true mean value M(t).

Table 3. Mean value and hazard rate

M(t)	55
m(t)	54.1640112
r(t)	0.433445678

We apply the simple homogeneous poisson software reliability model. By using the mean ϕ, we can calculate the software reliability by software GOMMA-LOMAX model.

5 Concluding Remarks

In this paper, we present the software reliability growth model and calculate the reliability and estimate the parameters in consideration of the multiple errors debugging at each testing stage. We introduce a Gamma-Lomax software reliability growth model for multiple software error debugging at each testing stage on the modification of Jelinski-Moranda model to have incorporated dependence between the failure times. And we calculate the parameters by maximum likelihood estimation method. Recently, software quality assurance activities for the development procedure concerned with the software development project have been highly concerned as well as the software product itself. However the measurement of the software reliability is very difficult. We suggest that the software reliability model should be applied for the standardization and the software quality measurement activities of the software product. We are going to study parameter estimation of software reliability models who propose in research that see forward.

Acknowledgments

This research was supported by the MIC(Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Assessment)

References

[1] ISO/IEC 9126 "Information Technology -Software Quality Characteristics and metrics-Part 1,2,3.
[2] ISO/IEC14598 "Information Technology -Software Product Evaluation-Part 1~6.
[3] ISO/IEC 12119 "Information Technology - Software Package - Quality requirement and testing".
[4] A.L. Goel, and K. Okumoto, “Time-dependent error detection rate model for software reliability and other performance measures”, IEEE Trans. Reliability, R-28, pp206-211, 1979.
[5] A.L. Goel, “Software reliability model : assumptions, limitations, and applicability”, IEEE Trans Software Eng, SE-11, pp 1411-1423, 1985.
[6] L.J. Bain, ‘Statistical Analysis of Reliability and Life-Testing Models’, MARCEL, DEKKER, 1991.
[7] Catuneanu, V.M. et al, “Optimal software release policies using SLUMT”, Microelectronics and Reliability, 28, pp. 547-549, 1988.
[8] Catuneanu, V.M. et al, “ Optimal software release time with learning rate and testing effort”, IEEE Trans. Reliability, 1991.
[9] Z. Jelinski, and P.B. Moranda, ”Software reliability research, in Statistical Computer Performance Evaluation”, Ed. W. Freiberger, Academic Press, New York, pp.465-497, 1972 .
[10] B. Littlewood, and J. L. Verrall, “ A Bayesian reliability growth model for computer software”, Applied Statistics, 22, pp.332-346, 1973.

[11] K. Okumoto, and A.L. Goel, " Optimum release time for software systems based on reliability and cost criteria", J. Systems and Software, 1, pp.315-318, 1980.
[12] N. Langberg, and N.D. Singpurwalla, "A Unification of some software reliability models", SIAM J. Scientific and Statistical Computation,6, pp.781-790, 1985.
[13] T. K. Nayak, "Software Reliability:Statistical Modeling and Estimation", IEEE Trans. On Reliability, 35, pp.566-570, 1986.
[14] M. L. Shooman, "Software Reliability : measurement and models", Proc. Ann. Reliability and Maintainability Symp, pp.485-491 , 1975.
[15] S. Yamada, and S. Osaki, "Cost-reliability optimal release policies for software systems", IEEE Trans. Reliability, R-34, pp422-424, 1985.
[16] S. Yamada, and S. Osaki, " Optimal software release polices for nonhomogeneous software error detection rate model", Microelectronics and Reliability, 26, pp.691-702, 1986.
[17] S. Yamada, and S. Osaki, "Cost-reliability optimal release policies for software systems", IEEE Trans. Reliability, R-34, pp422-424, 1985.
[18] S. Yamada, and S. Osaki, " Optimal software release polices for nonhomogeneous software error detection rate model", Microelectronics and Reliability, 26, pp.691-702, 1986.
[19] A.O.C. Elegbede, C.Chu, K.H. Adjallah, & F. Yalaoui, "Reliability Allocation Through Cost Minimization", IEEE Trans on Reliability, 52, 1, pp.96-105, 2003.
[20] H. J. Jung . "Software Reliability Growth Modeling and Estimation for Multiple Errors Debugging", Kyungpook National University, 1994.

Thesaurus Construction Using Class Inheritance

Gui-Jung Kim[1] and Jung-Soo Han[2]

[1] Department of Computer Engineering, Konyang University
26, Nae-Dong, Nonsan, Chungnam, Republic of Korea
gjkim@konyang.ac.kr
[2] Division of Information Communication, Cheonan University
115, Anseo-Dong, Cheonan, Chungnam, Republic of Korea
jshan@cheonan.ac.kr

Abstract. A lot of methodologies have been proposed for the component retrieval. Among them, thesaurus concept has been introduced for the similar component retrieval. In this paper, for the efficient retrieval of component, we classified classes with the concept of the inheritance relation and applied fuzzy logic to thesaurus method, and constructed an object-oriented thesaurus. The proposed method could express the category between concepts automatically, and calculate the fuzzy degree between classes by comparing matching weight with mismatching weight of each class and each category, and construct the thesaurus finally. By using classes of a component in the component retrieval, candidate components could be retrieved according to the priority order by the fuzzy similarity. Also, the retrieval performance was improved greatly by the thesaurus and the decision of the most suitable threshold value through simulation.

1 Introduction

By the effect of object-oriented methodology, a lot of methodologies have been proposed to retrieve components [1-3] and, among them, thesaurus concept has been introduced widely [4-8]. But the existing thesaurus research has several problems. First, since the clarity of query or conceptual association process from the query is ambiguous, it is difficult to define conceptual terms. Second, as components increase, a lot of noises are generated. Third, domain experts are required too much time and effort by extracted features [5, 6]. Like this, using thesaurus requires a lot of retrieval time and the case that we cannot find the information we need in retrieval process even happens [7, 8].

In this paper, we classified components conceptually, using the inheritance relation of classes constituting components, for efficient component retrieval. Since the inheritance relation classifies the concept group, which has been prescribed by only the subjectivity of constructor, automatically, consistency and objectivity of thesaurus can be guaranteed. For this purpose, we create Class Concept Category (CCC) to classify classes. Thesaurus is the synonym table constructed by comparing matching weight with mismatching weight of each class and CCC,

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 748–757, 2005.

using CCC Related Value (CRV), and calculating the fuzzy degree among these. This synonym table is used for query expansion.

In this research, we constructed the fuzzy-based thesaurus performing the component retrieval efficiently and made it possible to retrieve candidate components through query expansion by thesaurus.

2 Related Work

Object-based thesaurus [9] provides the automatic construction strategy for the expression of complex relation existing in objects by applying object-oriented paradigm consisting of concept expression level and instance expression level to thesaurus. But this technique requires the efficient query rebuilding process and the retrieval system that can be practically applied. Adaptive thesaurus [10] proposed the learning thesaurus through neural network on the basis of spreading activation based similarity measurement method. Proposed thesaurus can control the weight efficiently, but requires the method that automatically extracts thesaurus of standardized type and the technique that develops more suitable activity function. Hierarchical thesaurus system [6] proposed the method that the category for hierarchical classification is set up in code and the component is classified by behavioral characteristics. Component characteristic is composed of a pair of terms and extracted form software descriptor (SD)[11,12]. Pairs of terms classified hierarchically according to the characteristic construct the synonym dictionary by using fuzzy thesaurus. But, since it is not component retrieval but class retrieval using member functions and parameters, hierarchical thesaurus system has a disadvantage that, as classes increase, noises grow by the exponential increment of member functions.

3 Thesaurus construction

3.1 Class Concept Category

Class concept is classified by facet item and, in this research, we created Class Concept Category (CCC) to classify classes. Basic hypothesis of this research is that "For CCC constituting the software descriptor of a component, the number of class included in ith CCC implies the relationship between the component and ith CCC". Table 1 shows facet item (CCC) for class. 'Class Type' was classified into 5 CCC on the basis of MFC class library classification to develop Microsoft window application. And 'Component Type', 'User Interaction Style' and 'Using Scope' are classified by the facet classification method of existing object-oriented component [13].

Classified CCC can express the class inheritance relation corresponding to component characteristic. Inherited subclass can be classified into the same category as the super class or the different category with the super class. So, when it is classified into the different category, since the subclass contains the super class characteristic, we made the subclass inherit CCC of super class. But, as

the inheritance is nested, the characteristic of super classes becomes inconsistent; therefore the relation degree of CCC should be decreased according to the inheritance level. For that reason, in this research, we decreased the relation degree of CCC by $(1/2)^h$ according to the inheritance level h. Fig. 1 shows CCC inheritance. Since 'Document' is included in 'Data Structure and File' CCC and inherits 'Command_Target' class, it can inherit 'Window Support' and 'Frame Architecture' CCC. At this time, CCC relevance value is calculated with the number of class occurrence time per each CCC as $(1/2)^h$ according to inheritance level h. That is, numbers of class occurrence time per 3 CCC about 'Document' become 1 in 'Data Structure and File', $(1/2)^1$ in 'Window Support' and $(1/2)^1$ in 'Frame Architecture'.

Table 1. CCC by facet classification

Facet	Facet item (CCC)
Component Type	- File System and Interprocess Communication Service - User Interface Development
Class Type	- Window Support - Data Structure and File - Internet and Network - Exception and Debugging - OLE
Using Scope	- Frame Architecture - Application Processing
User Interaction Style	- Graphic and Drawing Support - Interactive and Dialog

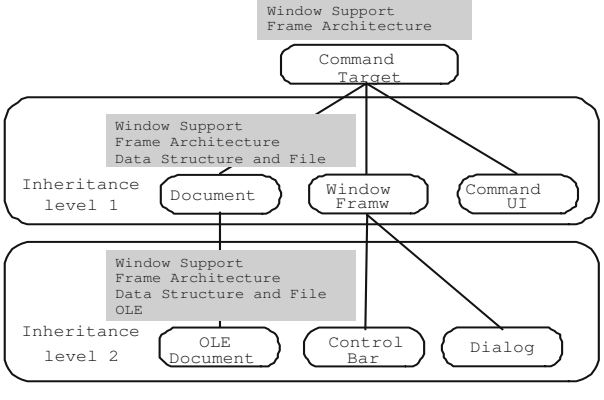


Fig. 1. CCC inheritance

On the basis of this, we defined '*class-CCC Table*'. This method is based on the weight calculation method by category [6], this table is calculated by CCC Relevance Value (CRV).

$$CRV_{i,j} = \frac{occurrence\ of\ i\ class\ in\ jth\ CCC}{occurrence\ of\ Total\ classes\ in\ jth\ CCC} \times \frac{occurrence\ of\ i\ class\ in\ jth\ CCC}{Total\ occurrence\ of\ i\ class\ in\ Total\ CCC} \quad (1)$$

In above equation, by CCC inheritance level, in case of CCC including the class, 'class occurrence number' is 1, and in case of CCC inherited from the super class, it becomes $(1/2)^h$ according to the inheritance level h. By giving the largest value to CCC including the class itself and the differentiated value to inherited CCC, object-oriented inheritance can be represented naturally. CCC relevance of the ith class in the jth CCC is dependent on the percentage of the ith class appearing in the jth CCC. This definition represents $CRV_{i,j}$ value is, irrelevant to the relative size of CCC, dependent on the class occurrence frequency. CRV has the value of range from 0 to 1 and, when no class (i) appears in a CCC(j), $CRV_{i,j}$ becomes 0.

3.2 Thesaurus Construction

The number of CCC occurrence calculation. Class repeated in several CCC means that it corresponds to more than one facet item, since the class has several characteristics. In this case, not only the CCC containing the class but also the CCC inherited have relevance to the corresponding class, *class-CCC* table should be able to represent both these CCC. Fig. 2 shows the class list of each component classified by CCC.

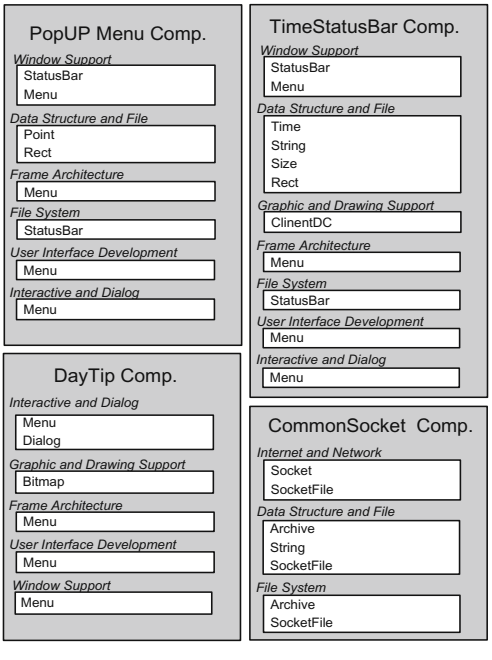


Fig. 2. Component classified by CCC

Table 2. CCC inheritance representation

CCC \ Component	PopUpMenu	TimeStatusBar	CommonSocket	DayTip
Window Support	StatusBar : 1 Menu : 1 Rect : 1/2	StatusBar : 1 Menu : 1 Rect : 1/4 ClientDC : 1/2	—	Menu : 1 Dialog : 1/2
Data Structure	Point : 1 Rect : 1	Time : 1 String : 1 Size : 1 Rect : 1	SocketFile : 1 Archive : 1 String : 1	—
Graphic	Point : 1/2 Rect : 1/2	Rect : 1/2 ClientDC : 1	—	Bitmap : 1
Internet	—	—	Socket : 1 SocketFile : 1	—
Interactive	Menu : 1	Menu : 1	—	Menu : 1 Dialog : 1
Fram Arch.	Menu : 1 StatusBar : 1/2	Menu : 1 StatusBar : 1/2	—	Menu : 1 Dialog : 1/2
File System	StatusBar : 1	StatusBar : 1	SocketFile : 1 Archive : 1	—
Application Processing	—	—	SocketFile : 1/2 Archive : 1/2 String : 1/2	—
User Interaction	Menu : 1	Menu : 1	—	Menu : 1 Dialog : 1

Table 2 represents CCC relevance to the class in figure 2 as a number. '1' is the case that the class is contained by corresponding CCC, '1/2' is the case that the class inherits CCC including super class, and '1/4' represents CCC inherited with two levels. These values are used in calculating CRV as the number of class occurrence time.

$CRV_{l,j}$ calculation for the *j*th CCC and the *l*th class. Table 3 shows the CCC Relevance Value (*CRV*) of components in Fig. 2.

Table 3. CRV

class \ CCC	Window Support	Data Struc.	Graphic	Internet	Interactive	Fram Arch.	File System	App. Processing	User Interface
StatusBar	0.123	—	—	—	—	0.044	0.2	—	—
Menu	0.115	—	—	—	0.188	0.167	—	—	0.188
Point	—	0.074	0.048	—	—	—	—	—	—
Rect	0.011	0.127	0.082	—	—	—	—	—	—
Time	—	0.111	—	—	—	—	—	—	—
String	—	0.148	—	—	—	—	—	0.056	—
Size	—	0.111	—	—	—	—	—	—	—
ClientDC	0.026	—	0.190	—	—	—	—	—	—
Dialog	0.013	—	—	—	0.083	0.019	—	—	0.083
Bitmap	—	—	0.286	—	—	—	—	—	—
Socket	—	—	—	0.5	—	—	—	—	—
SocketFile	—	0.032	—	0.154	—	—	0.071	0.048	—
Archive	—	0.444	—	—	—	—	0.1	0.067	—

Matching and mismatching weight calculation between classes. We explained the calculation process for matching and mismatching weight between classes using CRV. Matching and mismatching weight is the method using fuzzy

logic. In this method, for two classes, A and B, the more similar each CRV is for a CCC, the higher the affinity (i.e. matching weight) is, and the bigger the difference of CRV is, the higher the mismatching weight is. Here, we can get two vector $(a_1, a_2, \ldots, a_j, \ldots, a_c)$ and $(b_1, b_2, \ldots, b_j, \ldots, b_c)$ corresponding to the row of CRV table about class A and B. Vector value a_j and b_j represent CRV of A and B for the jth CCC. Following equation shows the calculation of matching weight for jth CCC of A and B.

$$\textit{matching weight for jth CCC} \quad : \quad \begin{cases} m_j = \dfrac{1}{1+|a_j - b_j|} & , \ a_j \rangle 0 \quad \textit{and} \quad b_j \rangle 0 \\ 0 & , \ a_j = 0 \quad \textit{or} \quad b_j = 0 \\ & , \ a_j = 0 \quad \textit{and} \quad b_j = 0 \end{cases} \tag{2}$$

Only in case that A and B occur in one CCC at the same time, we can get the class matching weight, as CRV is similar to each other, matching weight is close to 1 (i.e. A and B are similar). By processing these steps for all CCC $(1 \leq j \leq Cccc)$,we can get the matching weight of two classes, A and B.

$$\textit{matching weight betwwen class A and B} \quad : \quad M = \sum_{j=1}^{C} m_j \tag{3}$$

Also, mismatching weight for jth CCC of class A and B are calculated as the following.

$$\textit{mismatching weight for jth CCC} \quad : \quad \begin{cases} m^m{}_j = \dfrac{|a_j - b_j|}{1+|a_j - b_j|} & , \ \textit{either} \quad a_j = 0 \quad \textit{or} \quad b_j = 0 \\ 0 & , \ a_j = 0 \quad \textit{and} \quad b_j = 0 \\ & , \ a_j \neq 0 \quad \textit{and} \quad b_j \neq 0 \end{cases} \tag{4}$$

Only in case that A or B occurs in a CCC, we can get the class mismatching weight. As the difference of CRV for a CCC becomes bigger, mismatching weight is close to 1 (i.e. A and B are different). We can obtain the mismatching weight of A and B also by applying this process to all CCC $((1 \leq j \leq Cccc))$.

$$\textit{mismatching weight betwwen class A and B} \quad : \quad M^m = \sum_{j=1}^{C} m^m{}_j \tag{5}$$

The followings represent the matching and mismatching weight of class 'Status Bar' and 'Menu' using CRV table.

Matching weight of 'StatusBar' and 'Menu' :

$$\frac{1}{1+|0.123-0.115|}+\frac{1}{1+|0.044-0.167|}=1.882$$

Mismatching weight of 'StatusBar' and 'Menu' :

$$\frac{0.188}{1+|0-0.188|}+\frac{0.188}{1+|0-0.188|}+\frac{0.2}{1+|0-0.2|}=0.483$$

Matching and mismatching weight calculation between two classes and CCC. Matching and mismatching weight calculation for two classes between CCC is similar to the calculation between classes. CCC matching weight calculates only CRV of two classes, which occur for one CCC or not at the same time, and mismatching weight selects CRV of CCC which only one class between two classes occurs in. Therefore, calculation process is similar to that of matching and mismatching weight calculation between classes. And CCC matching weight is represented by M' and CCC mismatching weight is represented by $M^{m\prime}$. The following shows the calculation of matching and mismatching weight of class 'StatusBar' and 'Menu'.

CCC matching weight of 'StatusBar' and 'Menu':

$$\frac{1}{1+|0.123-0.115|}+\frac{1}{1+|0.044-0.167|}=1.882$$

CCC mismatching weight of 'StatusBar' and 'Menu':

$$\frac{0.188}{1+|0-0.188|}+\frac{0.188}{1+|0-0.188|}+\frac{0.2}{1+|0-0.2|}=0.483$$

Fuzzy synonym value calculation of two classes. Using matching and mismatching weight between classes and between two classes and CCC, we can create the final fuzzy synonym value of two classes.

$$f=\max\left(0,\frac{1}{2}\times\left(\frac{M-M^{m}}{M+M^{m}}+\frac{M'-M^{m'}}{M'+M^{m'}}\right)\right),\qquad 0\le f\le 1 \tag{6}$$

In case that $\frac{1}{2}\times\left(\frac{M-M^{m}}{M+M^{m}}+\frac{M'-M^{m'}}{M'+M^{m'}}\right)$ has the value less than 0, since it means

that mismatching weight is bigger than matching weight, the value is set up as 0 (i.e. it means there is no synonymy). The following shows the final synonym value of 'StatusBar' and 'Menu'.

$$f_{\text{'StatusBar','Menu'}} = \frac{1}{2} \times \left(\frac{1.882 - 0.483}{1.882 + 0.483} + \frac{1.882 - 0.483}{1.882 + 0.483} \right) = 0.592$$

As explained above, by applying the process for all classes in thesaurus, we can create synonym values between all classes.

4 Performance Evaluation

We used Visual C++ class library for simulation. These classes are common library and, since these are not under a bias toward specific category, provide various classes. Simulation environment is composed of 131 components included in 11 CCC and each CCC is composed of minimum 1 class to maximum 22 classes. The average number of classes is about 8 and total 1,023 classes, permitting overlapping, are included in CCC and the number of independent classes is 303.

In this research, we set up the optimum query expansion threshold that can guarantee the maximum component retrieval efficiency through simulation. In order to set up threshold, 20 classes are selected arbitrarily and similar expended sets for selected classes are decided. The similar expended sets means classes that are similar to a class or can be used together functionally and classes that are included in the similar expanded sets are limited to classes in thesaurus. And then, for 20 classes given as query, through comparing the expanded query by thesaurus proposed in this research with the similar expended sets we measured the precision and the recall from 0.6 to 1.0 at 0.05 threshold intervals.

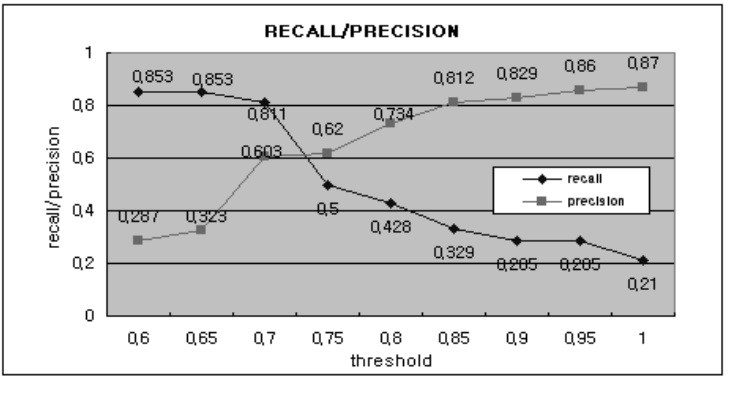


Fig. 3. Precision/Recall comparison by threshold

Fig. 3 show the average of precision and recall according to the retrieval efficiency by threshold for 20 arbitrary classes. In this research, we set up the

threshold value with 0.7 that maintains precision and guarantees recall to the utmost as the query expansion range to ensure the maximum retrieval efficiency.

And we made a comparison between the precision without thesaurus and that with thesaurus proposed in this research at 0.1 recall intervals [14]. This result is shown in table 4. The proposed retrieval with thesaurus compared with the retrieval without thesaurus (No-thesaurus) improved the efficiency by 22.3% $(((0.769 - 0.629)/0.629) * 100)$.

Table 4. Recall and precision

Recall	No-Thesaurus Precision	Proposed-Thesaurus Precision
0.1	0.68	1.00
0.2	0.70	0.95
0.3	0.78	0.92
0.4	0.75	0.82
0.5	0.72	0.86
0.6	0.63	0.72
0.7	0.59	0.71
0.8	0.54	0.65
0.9	0.48	0.57
1.0	0.42	0.49
Average Precision	0.629	0.769
Advanced average percent	-	22.3%

5 Conclusion

In this research, for the efficient component retrieval and the construction of flexible system structure, we constructed the object-oriented thesaurus, which the concept classification using class inheritance relation and the fuzzy logic are applied to. Proposed technique created the class concept category for class classification and constructed the fuzzy-based thesaurus synonym table using class relevance value by CCC for all classes. Comparing matching weight with mismatching weight for class and CCC and calculating fuzzy degree among them constructed fuzzy thesaurus. As a result of simulating the efficiency by adjusting the threshold of the query expansion to reduce noise, in case of synonym value more than 0.7, Average recall and precision were more than 81.1% and 60.3% each and the efficiency of proposed thesaurus compared with that of No-thesaurus improved by 22.3%. Therefore, the proposed technique in this paper can expand the component selection range by retrieving candidate components through the query expansion with thesaurus and, since components are classified by the concept about class library, is more efficient in component assembling.

But, since this technique constructed thesaurus according to the occurrence frequency by CCC, the efficiency might be decreased under the environment where the number of classes in a domain is small or classes are aggregated in the specific domain. Also, since the retrieval result is dependent on the degree

of similarity, it is limited to the accurate definition of functions about components, and therefore the additional information is required to understand more components.

References

1. Frankes and B. William, "Information Retrieval-Data Structure and Algorithms," Prentice-Hall, 1992.
2. R. F. Simmons, "Probability and Fuzzy-Set Application to Information Retrieval," Annual Review of Information Science and Technologyp, pp.117-151, 1985.
3. C. Danilowicz, "Modelling of User Preference and Needs in Boolean Retrieval System," Information Processing and Management, Vol.30, No.3, pp.363-378, 1994. Flyod S.: High speed TCP for Large Congestion Windows, IETF (2003)
4. E. Damiani, M. G. Fugini, "Automatic thesaurus construction supporting fuzzy retrieval of reusable components," Proceeding of ACM SIG-APP Conference on Applied Computing, Feb., 1995.
5. Jae-Hun.Choi, Ji-Suk.Kim, Gi-Hwan.Cho, "A Fuzzy Retrieval System to Facilitate Associated Learning in Problem Banks," The Korea Information Science Society Transactions, Vol.29, No.4, pp.278-288, Apr., 2002.
6. E. Damiani, M. G. Fugini and C. Bellettini, "Aware App-roach to Faceted Classification of Object-Oriented Component," ACM Transaction on Software Engineering and Methodology, Vol.8, No.4, pp.425-472, Oct., 1999.
7. R. Rada, H. Mili, E. Bickenell and M. Blettner, "Development and Application of a Metric on Semantic Nets," IEEE Transaction on System, Mand Cybernetics, Vol.19, No.1, pp. 17-30, 1989.
8. H. Chen, T. Tim and D. Fye, "Automatic Thesaurus Gene-ration for an Electronic Community System," Journal of the American Society for Information Science, Vol.46, No.3, pp.175-193, 1995.
9. Jung-Ae.Kim, Jong-Min.Park, Won-Jung.Kim, Jae-Dong.Yang,, "Relaxing Reference Queries in the Object-Oriented Thesaurus," Proceedings of the Korean Information Science Society Conference, pp.208-210, Oct., 2002.
10. Jong-Pill.Choi, Myeong-Bok.Choi, Min-Koo.Kim, "Adaptive Thesaurus using a Neural Network," The Korea Information Science Society Transactions, Vol.27, No.12, pp.1211-1218, Dec., 2000.
11. Y. S. Maarek, D. M. Berry and G. E. Kaiser, "An information retrieval approach for automatically constructing software library," IEEE Transaction on Software Engineering, Vol.18, No.8, pp.800-813, Aug., 1996.
12. D. Batory and S. O'Malley, "The design and implementation of hierarchical software systems with reusable components," ACM Transaction on Software Engineering and Methodology, Vol.1, No.4, pp.355-398, Oct., 1992.
13. Gui-Jung. Kim, "Fuzzy Logic Based Thesaurus Construction for efficient Component Retrieval," Kyung-Hee University, Korea, pp.129, 2003.
14. B. Y. Ricardo and R. N. Berthier, "Modern Information Retrieval," Addison-Wesley, 2000.

An Object Structure Extraction Technique for Object Reusability Improvement Based on Legacy System Interface

Chang-Mog Lee, Cheol-Jung Yoo, and Ok-Bae Chang

Division of Electronics and Information Engineering, Chonbuk National University, 664-14 1ga Duckjin-Dong Duckjin-Gu Jeonju, South Korea
{cmlee, cjyoo, okjang}@chonbuk.ac.kr

Abstract. This paper suggests a technique, OSET(An Object Structure Extraction Technique for Object Reusability Improvement Based on Legacy System Interface) for reuse and reengineering by analyzing the Legacy System interface to distill the meaningful information from them and disassemble them into object units which are to be integrated into the next generation of systems. The OSET method consists of a procedure with 4 steps: 1) the interface use case analysis step, 2) the interface object dividing step, 3) the object structure modeling step, and 4) the object model integration step. In step 1, the interface structure and information about the interaction between the user and the Legacy System are obtained. In step 2, the interface information is divided into semantic fields. In step 3, the structural and collaborative relationship is studied and modeled among interface objects. Finally, in step 4, the object model integration step, integrates the models and improves the integrated model at a higher level.

1 Introduction

In most companies, the system is developed by using non object-orienting or object-orienting language, but most are still using the legacy application that does not accurately apply the object-orienting concept. This is largely due to the stable nature of the current system that most companies use. In most cases, however, these existing systems cannot cope with rapidly changing business environments.

The purpose of this project is to acquire knowledge about the Legacy System interface, and to create an object-oriented model with this acquired knowledge. Interface refers to the user interface for the end user. That is to say, it is the most commonly-used method for the end user of user applications[1]. In most cases, information in application systems is generated by interaction that communicates with the system through an interface. OSET can naturally classify the interface information generated through the course of interaction between the end user and the system. The organization of this paper is as follows. First in chapter 2, I will examine related studies, and in chapter 3, the structure of OSET's use of reverse engineering is explained. Chapter

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 758 – 768, 2005.

4 shows examples of the course of distilling and classifying interface objects from the Legacy application using OSET. Chapter 5 discusses conclusions and future projects.

2 Related Work

In this chapter, information related to reverse engineering used in distilling object information from the Legacy System interface is prescribed. Reverse engineering is a reverse process that recovers the factors of higher-level conceptual abstractions from the physical systems. This flow from the developed system to the initial course is the analysis step and design in reverse. In table 1, characteristics of reverse engineering methodologies are compared, along with their advantages and disadvantages. These were compared and examined for application in this study.

Table 1. Comparison of kinds of reverse engineering methodology

Classification / Contents	Input data	Standpoint of classification	Target model	Characteristics	Etc.
Kinds of reverse engineering	DB schema	Data	ER diagram	• Key criterion • Inclusion dependency criterion • Combination of inclusion and key dependency	Using methodology such as Batini[2], Navathe[3] Chiang[4], Shivak[5] etc.
	Source code	Process	Reuse module	• Using candidature criterion	using RE2[6], Saleh[7] methodologies
	Form structure	Object	Object-oriented diagram	• Using application system form • Object structure and scenario	using FORE methodology[8]
	Source code	Object	CTF (Compact Trace Format)	• Expressing object interaction movement using CTF schema	Abdelwahab[9]
	Interface structure & interaction information with end user	Object	Object-oriented diagram	• Using interface of legacy system • Grasping object structure • Classification and restoration of object unit	OSET

2.1 Example of Reverse Engineering That Has Database Schema as Its Input Data

This section presents several examples of reverse engineering: 1) A method to transform the relational database schema to the ER diagram, 2) A method to transform the network schema into the ER diagram, 3) A method to transform the hierarchical schema into the ER diagram, and 4) A method to transform the object-oriented database into a binary relationship diagram etc.. In this paragraph, the course to reverse the engineer relational database schema, which to date the legacy system used the most widely, into the ER diagram is described. The method to reverse the engineer

relational database schema into the ER diagram is based on functional dependence and inclusion dependence. Also, it derives a conceptual diagram of the database using tuples of the existing database and catalog information in DBMS (Database Management System). Relations are defined by using a primary key and a foreign key. This definition of relations is obtained through interactions with the user. Then the relational schemas are interpreted and transformed on the basis of the classification of the relations. Cardinality constraints are derived through analyzing functional dependencies[2, 3].

2.2 Example of Reverse Engineering with a Source Program as Its Input Data

This example of reverse engineering with a source program as its input data is provided in order to lead the module structure of the source code and program its logic structure. To do this, the source program is analyzed and divided again into semantic information units. The representative example is the RE2 (Reuse Reengineering) project[6]. The RE2 project outlines the principles that the reverse engineering and reengineering techniques use in the Reuse Reengineering process that include the existing source code. This means that its purpose is to reuse software by setting candidature criterion. Also, the purpose is to create a module that can be reused from existing systems. To do this, data must be abstracted. Although criterion based on this treatment principle was suggested, the setting is indistinct. The Saleh methodology is a reverse engineering methodology suggested by Saleh Boujarwah. The purpose of this method is to obtain a high level abstract of communication software with a particular language that is established in ISO generally referred to as Estelle for system dispersion and protocol. However, this method is dependent on communication software [7]. FORE methodology is similar to this study in creating an object structure because it has the form structure and user interaction information as its input data, but it uses a CRC (Class Responsibility Collaboration) card technique to express the object model. Currently Integrated Modeling Language, known as UML (Unified Modeling Language), is being used, and an object-oriented system is being developed that now follows UML notation in most cases. Therefore, it complicate matters when it is mapped into UML again in order to reuse this object information in the current system because the distilled object model information is not suitable for current conditions, and it contains a possibility of data error possibly generated by that course.

3 The Object Structure Extraction Technique: OSET

The first step of OSET is the interface use case analysis. This step obtains the Legacy application interface field information and user interface information. There are several examples of interaction using the legacy system interface, an action to input data into the input interface, an event creating an action to send the interaction example through the interface, an operation action to treat input data created by the event, an action to return the data treated by the operation action to the user interface, and an action to deliver data from the database to the interface, etc. Fig. 2 shows the algorithm used to analyze one case of interface use. This interaction information and in-

formation about the interface are automatically collected by an agent-based information collector.

The information composing the interface obtainable by the agent-based information collector can be divided into 4 types: 1) interface area, 2) input support control, 3) interface editing field, and 4) interface treatment. The second step of OSET is the interface object point of division. In this step, information about the interface collected by the agent program of the previous step is used. Then, it divides the various types of field information shown in the interface into a semantic information unit object, according to the type of interface information such as area (item, total, grid, event, conditional area etc.), input support control(combo box, server interface, radio button, check button etc.), and interface edition field type (key input field, data

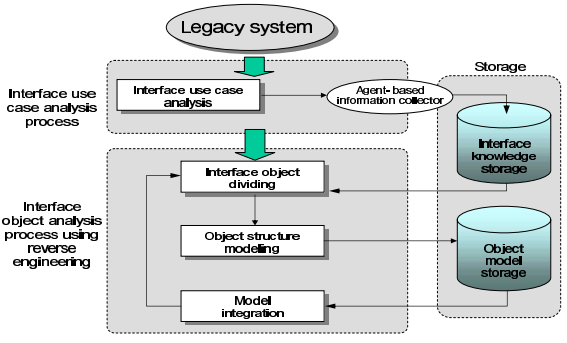


Fig. 1. Structure of TELOR

```
DB  DBList, Sgl_DBList;
 // DBList is connection list having information about all Sgl_DBList
while(!All_Interface_exit()) // repeat on all interfaces
{
   getSgl_Interface(); // obtain single interface
   while(!exit)
   {
getInter_User_System();
// obtain interaction information between user and system through agent system
     Sgl_DBList = saveAll_Interface_Inform();
// save the obtained single interface information
   } // end while
   addList(DBList, Sgl_DBList); // add single DBList to DBList
} // end while
```

Fig. 2. Interface use case analysis algorithm

input field, inquiry and modification field, inquiry field, event generation field, and system control field) etc. Fig. 3 is the algorithm that shows the interface object dividing process.

The third step of OSET is the object structure modeling. This step employs the results from the previous steps. It also derives structural and collaborative relations of the initial level between objects. As a result of this step, the object structure model consisting of an object name, object attribute, and the structural relationships of the derived object are formed. Also, to indicate the object structure model, UML, the standard current modeling language, was used. Fig. 4 is the algorithm that shows the course of the modeling object structure.

```
while (DBList != NULL) // DBList: List of all information obtained by agent
{
Object  InterOb, Sgl_InterOb; // interface object
Object  ComplexOb, Sgl_ComplexOb; // complex object
Object  DBaccessOb, Sgl_DBaccessOb; // DB approaching object
Object  ExtraOb, Sgl_ExtraOb; // additional obejct
Sgl_InterOb = createInterfaceObject(DBList); // creating single interface object
giveNameToObject(Sgl_InterOb.name); // Giving object name using interface name
givePropertyToField(Sgl_InterOb.field); // allotting property to interface field
addList(InterOb, Sgl_InterOb); // add to interface object list
while (Sgl_InterOb.field != NULL) // Search all fields in single interface
{
if (isComplexField(Sgl_InterOb.field)) // search complex field
{
Sgl_ComplexOb = createObject(Sgl_InterOb.field);
// lead complex field into complex object
giveNameToObject(Sgl_ComplexOb.name); // give object name
givePropertyToField(Sgl_ComplexOb); // give property
addList(ComplexOb, Sgl_ComplexOb); // add to object list
} // end if
if (isDBaccess(Sgl_InterOb.field)) // search DB approaching field
{
Sgl_DBaccessOb = createObject(Sgl_InterOb.field);
// lead DB approaching field to DB approaching object
giveNameToObject(Sgl_DBaccessOb.name); // give object name
givePropertyToField(Sgl_DBaccessOb.field); // give property
addList(DBaccessOb, Sgl_DBaccessOb); // add to object list
} // end if
else if (checkObject(Sgl_InterOb.field)) // search additional field
{
Sgl_ExtraOb = createObject(Sgl_InterOb.field);
// lead additional field into additional object
giveNameToObject(Sgl_ExtraOb.name); // give object name
givePropertyToField(Sgl_ExtraOb); // give property
addList(ExtraOb, Sgl_ExtraOb); // add to object list
} // end if
Sgl_InterOb.field = Sgl_InterOb.field.next; // move to the next field
} // end while
DBList = DBList.next; // move to the next single interface
} // end while
```

Fig. 3. Interface object dividing algorithm

The fourth step, the last step of OSET, is the model integration step. The purpose of that is to integrate the model integration step and suggest a higher level integration

```
while (InterOb != NULL) // search all interfaces
{
  while (InterOb.ExtraOb != NULL) // search all total field applicants
  {
    if (isCompute_field(InterOb.ExtraOb)) // verify total field
    {
    Compute_Inform = callAgentSystem(InterOb.ExtraOb);
       // obtain all information relted to totl field from agent
    while (Compute_Inform != NULL)
    {
      operator = searchOperator(Compute_Inform);
      // search for the minimum unit operator
      operand = searchOperand(Compute_Inform);
      // search for the minimum unit operand
      variable = initVariable(operand); // set up variables
      relationOb = searchObject(InterOb, operand);
      // search for object related to operand
      setRelation(variable, relationOb);
      // Set up the relation with object and variables
      method = createMethod(variable, operator);
      // create calculation formula between operator and operand
      Compute_Inform = Compute_Inform.next;
      // move to the next information related to totl field
    } // end while
    setMethodToObject(method, InterOb.ExtraOb);
      // allot method to total field object
    InterOb.ExtraOb = Sgl_ExtraOb.next; // move to the next total field applicant
  } // end if
} // end while
   InformOb = InformOb.next; // move to the next interface
} // end while
```

Fig. 4. Object structure modeling algorithm

```
model = NULL; // set up the initial value of model variables
alter_model = NULL; // set up the initial value of alter_model variables
while (!isSuitableModel(model)) // decide appropriateness of model
{
  model = selectStruc(InterOb); // choose whole interface object structure model
  InterOb = InterOb.next; // move to the next interface
  while (InterOb != NULL) // search all interface objects
  {
    alter_model = selectStruc(InterOb); // choose object structure model of the moved interface
    if (isDifferModel(model, alter_model)) // decide the similarity of two models
    {
      mergeSameOb(model, alter_model); // integrate model based on common object
      mergeByRelation(model, alter_model); // integrate related model
      solveAmbiguousName(model, alter_model); // solve indistinct name
      solveAmbiguousStruc(model, alter_model); // solve indistinct name
      model = saveToAll(); // save the integrated model
    } // end if
   InterOb = InterOb.next; // indicate the next interface
   } // end while
} // end while
```

Fig. 5. Model integration algorithm

model unit models in order to show object information. The final result obtained from the course of OSET is the object structure model. Fig. 5 is the algorithm used for model integration.

4 OSET's Application Example

This section will describe the course of a real example for obtaining an object model from the interface legacy system interface information using OSET. To show this actual example, Sales Force Automation(SFA) used in a merchandising company was selected as the legacy system.

The SFA system is a computer system used to strengthen company sales, and the legacy system is generally used in most companies with distribution networks. Fig. 6 and Fig.7 show a part of the SFA system interface.

4.1 Interface Use Case Analysis Step

The Interface use case analysis step seeks the object structure and method based on the characteristics of the interface. Information about the interface can be divided into two parts: 1) interface structure and 2) user operation and database. Table 2 is an example of information about sales note input of the SFA system. It is the information about the interface structure among the abstraction(virtual) information to be collected based on Fig. 2 which is an acquisition procedure of use case information.

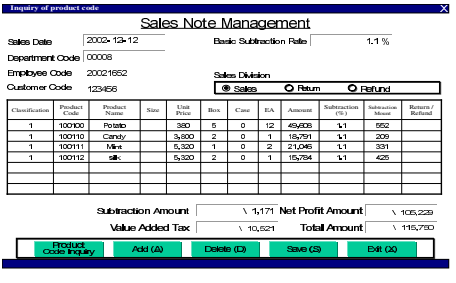


Fig. 6. Sales note management interface

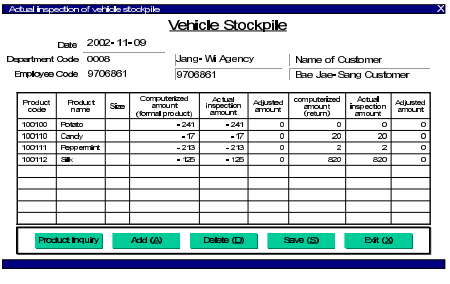


Fig. 7. Actual inspection interface of vehicle stockpile

In the interface structure in Table 2, the name of the caption is the field that plays a role to help read the meaning of the interface field information. The value of the interface field is assigned to the name of the variables during the actual operation of the legacy system. Then the interface field type is roughly divided into input support control/interface field type/types per area/types per event etc. as divided earlier in chapter 3. When this is going on, the field type that includes variables of each field is indicated at the right clause of Table 2. The system field is a part controlled by the system, and the information obtained from the system together with the data is in cluded. The data input field inputs single data, and the element field that the user

cannot further divided. The Grid field is a group field form that interface fields with the same name and same meaning are repeatedly connected with as arrangements. “Total field” means a field in which the value is calculated by a particular calculation. The use case with information about this interface calculation is show in Table 3.

Table 2. Example of information about sales note input interface: Interface structure

Contents			Input support control/ interface field type /area/types per event
Classification	**Caption name**	**Name of variables**	
Name of interface	Sales note input	pan_FSSAT01_v	Interface object
Interface field	sales date	date_v	System field
	department code	dept_code_v	Data input field
	employee code	emple_code_v	Data input field
	customer code	custom_code_v	Data input field
	basic deduction rate	deduct_rate_v	Data input field
	sales division	pan_v, back1_v, back2_v	Radio button
	classification, product code, product name, unit price,. BOX, CASE, amount, deduction, deduction amount	divide_v1, divide_v2, p_code_v1; p_code_v2, … p_name_v1; p_code_v2, …	Grid field
	Deduction amount	d_amount	Total field
	VAT	a_amount	Total field
	Pure sales amount	p_amount	Total field
	Total sales amount	s_amount	Total field
	Product code inquiry	p_code_inq_v	Event generation field
	Add	p_add	Event generation field
	Delete	p_del	Event generation field
	Save	p_sav	Event generation field
	Finish	p_exit	Event generation field

Table 3. Example of information about sales note input interface: interface operation order and the result

Operation order	Action	Correspondent interface field	Result content
1	System provision calculation	date_v	2002-12-12
2	User data input	dept_code_v	00008
3	User data input	emple_code_v	20021652
4	User data input	custom_code_v	123456
5	User data input	deduct_rate_v	1.1%
6	Radio button input	rad_pan_v	rad_pan_v = true
		rad_back1_v	
		rad_back2_v	
7	Action event (Product code inquiry button)	p_code_inq_v	p_code_inq_v = true
8	Database approaching	divide_v1, divide_v2, …	1
		p_code_v1, p_code_v2, …	100100
		p_name_v1, p_name_v2, …	peppermint candy
		p_value_v1, p_value_v2, …	5,320
9	Total	d_amount_v	1,171
10	Total	a_amount_v	10,521
11	Total	p_amount_v	105,229
12	Total	s_amount_v	115,750
:	:	:	:

4.2 Interface Object Dividing Step

This step analyzes the interface information saved in the knowledge storage and divides it into semantic information units. The interface object consists of several factors: interface area, input supporting control, interface editing field, and interface treatment type, etc. as classified in section 3. These factors are object applicants. This course has a process as in Fig. 3 of section 3. The first step creates an object interface of a single unit by using the sales note management interface. One single unit interface object has all fields of sales note management interface as its attribute, and Fig. 8 shows the class diagram of sales note input interface (Class name: pan-FSSAT01). The second step abstracts complex fields (ex, grid field, input supporting control) among the attributes in the sales note input object interface. Using this method, small unit object separation can be derived from a complex large object. Even small unit objects derived this way are assigned an object name. Fig. 9 shows small unit objects (Pan_FSSAT01, Pan_FSSAT01_grid) separated from the sales note input object interface. As the grid field is included here in the example of sales note input interface, the grid field element is separated. As for the object name of a small unit object derived separately, the new name is to be made by mixing the newly abstracted object name (therefore, interface field name) to identify the original object from the small-unit object name.

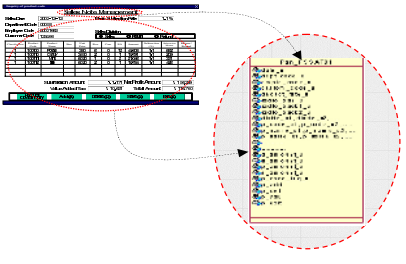


Fig. 8. Sales note input interface object

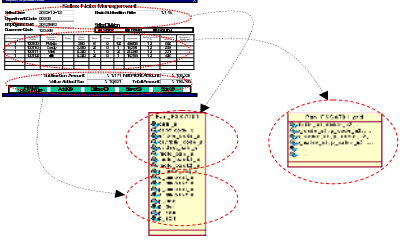


Fig. 9. Sales note input interface object planated by Pan_FSSAT01, Pan_FSSAT01_grid

The third step determines which field approaches the database among fields in each object. In Fig. 9, the field variables to approach to database is divide_v1, divide_v2, …, p_code_v1, p_code_v2, …, p_name_v1, p_name_v2, …, p_value_v1, p_value_v2, … etc. of the Pan_FSSAT01_grid object. These values are delivered from the employee code table(CEMP01TT), product code(CGDS01TT) table, department code(CDEP01TT) table, and customer code (CDCS01TT) table etc. The fourth step abstracts additional applicant objects. It can be considered as a step to check the objects that are not found in the previous step.

4.3 Object Structure Modeling Step

The object structure modeling step is made using the objects that were divided in the previous step. Although the objects divided in the previous step have obtained the

attributes, the method is not defined, even though this method is required for the operation. This step obtains the operation of the object, because it can be the main target of the method as it indicates the movement of the objects. The field shows that this operation phase is the total field. The total field consists of the minimum unit operator and the single field. Method definition can be derived on the basis of the name of the total field because the name of the total field is included with the name of the variables that contain the logic structure of the treatment.

4.4 Object Model Integration Step

In this step, the object structure model derived from several interface objects is integrated into one model. Also, the principle of class structure using generalization/specialization of the object is applied to the integrated object for a better result. Therefore, a more abstracted object exists that can represent this. In this way, the model that has two objects integrated is shown in Fig. 10.

5 Conclusion and Further Study

In this paper, objects were analyzed for semantic information from the interface of the legacy system, and OSET was suggested to perform the abstraction of the integrated modeling.

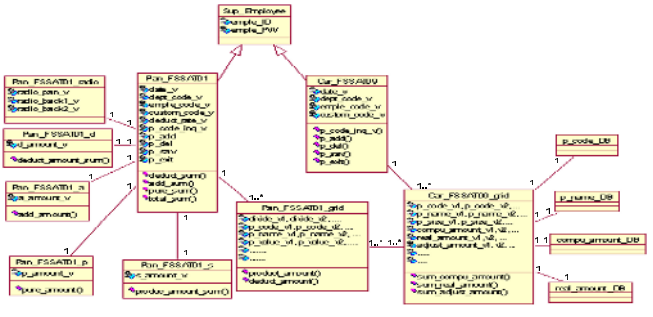


Fig. 10. Integrated object model

The achievements of this study are as follows. First, input data is a reverse engineering study that uses interaction information based on an interface that happens most basically between the user and the system. Second, OSET suggests how to save the knowledge from the interface in knowledge storage using the algorithm step. Third, it treats the data and the process simultaneously. Fourth, it suggests the object analysis model using this interface information and shows examples by applying them to an actual system known as SFA. In the future, to treat system analysis and designer work more easily, an object analysis course automation tool, such as OSET, needs to be developed.

References

1. M. Malki, A. Flory, M. K. Rahmouni, “Static and Dynamic Reverse Engineering of Relational Database Applications: A Form-Driven Methodology”, ACS/IEEE International Conference on Computer Systems and Applications (2001) 191 - 194
2. C. Batini, S. Ceri and S. B. Navathe, Conceptual Database Desing - An Entity Relationship Approach, Benjamin/Cummings, Redwood City (1992)
3. S. B. Navthe and A. M. Awong, “Abstracting Relational and Hierarchical Data with a Semantic Data Model”, Proceedings 6th Int. Conference on the Entity-Relationship Approach, Springer-Verlag (1987) 305 - 336,
4. R. H. L. Chiang, T. M. Barron, and V. C. Storey, "Reverse Engineering of Relational Database: Extraction of an EER Model from a Relational Database", Data Knowledge Engineering, 12(2) : (1994) 107 - 142
5. P. Shoval and N. Shreiber, “Database Reverse Engineering: from the Relational to the Binary Relationship Model”, Data Knowledge Engineering (1993) 293 - 315
6. G. Canfora, A. Cimitile, and M. Munro, “A Reverse Engineering Method for Identifying Reusable Abstract Data Types”, Proceedings Working Conference on Reverse Engineering,, IEEE Computer Society Press (1993) 73 - 82
7. K. Saleh and A. Boujarwah, “Communications software reverse engineering: a semi-automatic approach”, Information and Software Technology (1996) 379 - 390
8. H. Lee and Ch. Yoo, “A Form Driven Object-oriented Reverse Engineering Methodology”, Information Systems, Vol. 25 (2000) 235 - 259
9. Abdelwahab Hamou-Lhadj, Timothy C. Lethbridge “A Metamodel for Dynamic Information Generated from Object-Oriented Systems” Electronic Notes in Theoretical Computer Science, Volume 94 (2004) 59 - 69

Automatic Translation Form Requirements Model into Use Cases Modeling on UML

Haeng-Kon Kim[1] and Youn-Ky Chung[2]

[1] Department of Computer Information & Communication Engineering,
Catholic University of Daegu, Korea
hangkon@cu.ac.kr
[2] Department of Computer Engineering, Kyung Il University, Republic of Korea
ykchung@kiu.ac.kr

Abstract. Given the investment organizations are making in use cases and the increasing use of computers to control safety-critical applications, the research on integrating use cases and safety consideration is highly demanded. However, currently the approaches for incorporating safety analysis into use case modeling are scarce. In this paper, we present an approach to integrating safety requirements into use case modeling. We demonstrate how deductive and inductive safety techniques can be used hand-in-hand with use cases. An application of the proposed approach facilitates early hazard identification and assessment, as well as elicitation and tracing of safety requirement though the entire development process. The proposed approach is illustrated by a realistic case study – a liquid handling workstation.

Keywords: UML, use cases, safety analysis, safety requirements, CBD.

1 Introduction

The Unifying Modeling Language (UML) [1] is gaining increasing popularity and has become de facto industry standard for modeling various systems many of which are safety-critical. UML promotes use case driven development process [1] meaning that use cases are the primary artifacts for establishing the desired behavior of the system, verifying and validating it. Elicitation and integration of safety requirements play a paramount role in development of safety-critical systems. Hence there is a high demand on methods for addressing safety in use case modeling. In this paper we propose an approach to integrating safety requirements in use cases.

It is widely accepted that building in safety early in the development process is more cost-effective and results in more robust design [2,3,4]. Safety requirements result from safety analysis – a range of techniques devoted to identifying *hazards* associated with a system and techniques to eliminate or mitigate them [3,4]. Safety analysis is conducted throughout the whole development process from the requirements conception phase to decommissioning. However, currently the approaches for incorporating safety analysis into use case modeling are scarce.

To make an integration of safety analysis and use cases complete we need to investigate the use of not only deductive techniques but also inductive safety techniques.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 769 – 777, 2005.

Indeed, to conduct safety analysis at the early stages of the development we need deductive techniques able to assess hazards at functional level. We describe a general structure and behavior of control systems which employ manual reconfiguration for error recovery. We give a brief overview of the case study – the liquid handling workstation. We finally present our approach to conducting hazard analysis at the early stages of development by embedding the method into use cases.

2 Dependable Control Systems

2.1 Control Systems

In this paper we focus on modeling of dependable [2] control systems. The basic building blocks of a control system are presented in Fig.1. An *application* is a physical entity whose operation is being monitored and controlled by a *computer-based controller*. The behaviour of the application evolves according to the involved physical processes and control signals provided by the controller. The controller observes behaviour of the application by means of *sensors* – the devices that convert application's physical quantities into electric signals to be input into the computer. On the basis of data obtained from the sensors the controller manipulates states of *actuators* – devices that convert electrical signals from the output of the computer to physical quantities, which control the function of the application.

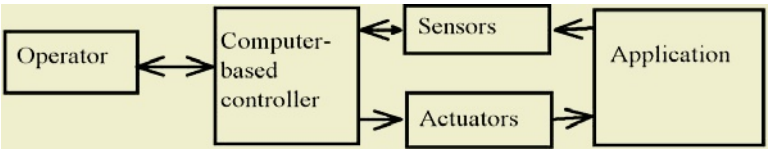


Fig. 1. A general structure of a control system

2.2 Manual Reconfiguration

As we discussed above, system's components are susceptible to various kinds of faults. A fault manifests itself as *error* – an incorrect system's state [11]. The controller should detect these errors and initiate error recovery. In this paper we investigate manual reconfiguration, as it is the most commonly used error recovery mechanism [11]. The manual reconfiguration is performed by an operator. The behaviour of the control systems with manual reconfiguration for error recovery follows a general pattern graphically represented in Fig.2. Initially the system is assumed to be fault free. The operator starts the system functioning and upon successful initialization the system enters an automatic operating mode. Within the automatic mode the system executes a standard control loop provided no error is detected.

Upon detection of an error the system leaves the automatic mode and reverts to the operator's control. An operator assesses damage caused by an error and executes a required error recovery procedure. As a result of error recovery the operator might resume

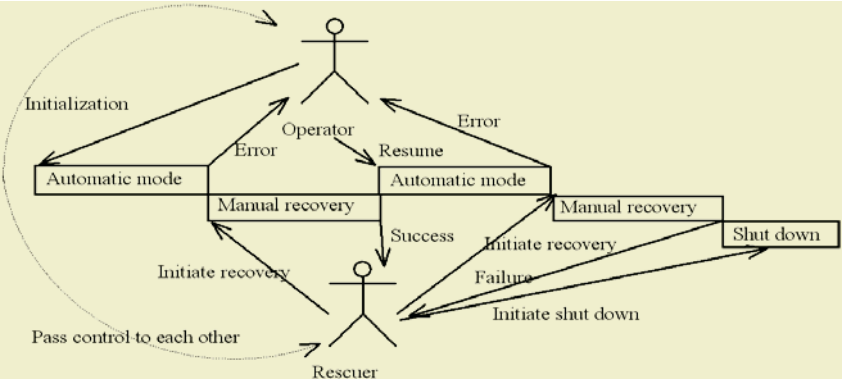


Fig. 2. Illustration of manual reconfiguration

the automatic mode or execute shutdown. The human operator plays two different roles while controlling a system with manual reconfiguration. On one hand, the operator initiates and monitors system functioning in the automatic mode, on the other hand s/he is a trouble-shooter performing error recovery in the manual mode. Hence, the operator should be represented by two actors in the use case diagram: the first is *operator* and the second is *rescuer.* The general form of the use case diagram is given in Fig.3.

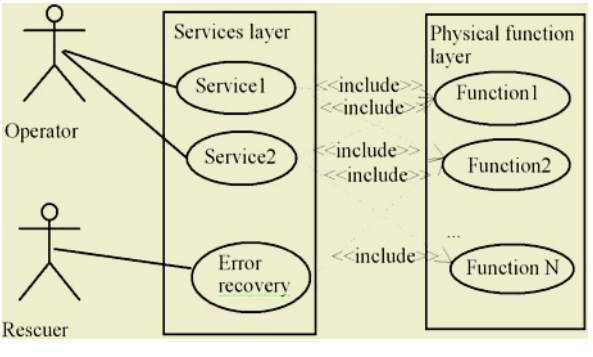


Fig. 3. Schematic representation of use case diagram

3 Use Case Modeling and the Requirement Risk Assessment Method

3.1 The Functional Hazard Assessment Method

Essentially, a use case is a coherent unit of system's functionality [1]. Hence to smoothly incorporate safety analysis in use case modeling we need a technique, which is able to identify hazards on a function level. In this paper we investigate how the Requirement Risk Assessment (RRAM) method can be applied to use cases to iden-

tify hazards and assess them. The RRAM allows us to identify hazards at a functional level and correspondingly derive safety requirements. The RRAM process consists of five steps [5]:

1. Identification of all functions associated with the level under study.
2. Identification and description of failure conditions associated with these functions.
3. Determination of the effects of the failure condition.
4. Classification of failure effects on the system.
5. Assignment of requirements to the failure conditions to be considered at the lower level.

An identification of failure conditions can be done systematically by applying the following guidewords [7]:

- *Loss of function*
- *Function provided when not required*
- *Incorrect operation of function.*

We argue that use case model constitutes a suitable description of system's functionality. Indeed

- use cases clearly define system's functions
- use case diagrams explicitly show interdependencies between use cases by means of associations
- the proposed layered structure of use case diagram clearly identifies levels of abstraction and therefore, allows us to apply the RRAM method iteratively, layer after layer.

Use case Move to X position

Brief description This use case defines reaction on the command **"Move to** X **position".** As a result of the execution of the use case either the operating head is brought to X position and success is reported or failure is reported.

Includes none

Preconditions None

Postconditions The operating head is placed at the position X or failure is reported

Typical flow of events

1. Check that $xmin \leq X \leq$ xmax, if not **X_Failure1** in alternative flow of events
2. Check current x-position.
3. If current x-position equals X then report success of execution. Otherwise move operating head to X position.
4. Check current x-position. If current x-position equals X then report success of execution else **X_Failure2** in alternative flow of events

Alternative flow of events

X_Failure1. Prompt message "Input parameter X is outside of valid range"

X_Failure2. Prompt message "Loss of precision of X movement"

Fig. 4. Example of physical-layer use case description

For the control systems with manual reconfiguration discussed above the RRAM starts from analyzing use cases at the server's layer. By systematically applying guidewords we describe consequences, criticality and mitigation of the corresponding failures for each use case. Such an analysis of individual use cases suffices if the use cases are independent. However, if there are use cases associated with each other or the use cases, which might be executed in parallel, then we also need to consider the effect of multiple failures. In this case various combinations of failures should be explored and hazards induced by them assessed.

Table 1. The template for conducting the RRAM

Name of use case (of a group of use cases) under investigation			
Grouping	Guide word: (Loss / When not required/ Incorrect operation of function)		
	consequences	criticality	mitigation
Independent use case	In terms of the system	Catastrophic Critical Marginal Negligible	- Strengthen precondition - Expand postcondition - Introduce additional use cases to mitigate consequences
Executed as a part of associated use case	In terms of associated use case	See above	See above
Executed in parallel with use cases	In terms of a system executing a group of use cases: consider all possible combinations of failures	See above	See above

Table 2. An excerpt from an application of the RRAM to a use case

Move to X position									
	loss			When not required			Incorrect operation		
	consequences	Criticality	mitigation	consequences	Criticality	mitigation	consequences	Criticality	mitigation
Executed as a part of use case *"Aspirate"*	Not able to position pipette	Marginal	Retry by executing "Retry Move to X position". If not successful then shut down	In the worst case moves the head while pumping or doing Z-movement. Danger to damage the equipment or lose the experiment	Catastrophic	Strengthen precondition of "Move to X position" to prevent unexpected provision. If occurred shut down the whole system	Wrong positioning of the head. Incorrect experiment or damage of the equipment	Critical	Strengthen postcondition of "Move to X position" to insure that either positioning is correct or failure is detected. Retry if failed. Add use case "Retry Move to X position".

3.2 Example of an Application of the RRAM

The use cases at server's layer of Fillwell are independent of each other. An application of the RRAM has shown that the most serious consequences are induced by the incorrect provision of the use cases "Aspirate" and "Dispense". The loss of them deadlocks the system. Finally, an inadvertent provision of them is assumed to be unfeasible. The next iteration of the RRAM focuses of the analysis of use cases of the component's layer. An excerpt from the analysis of the use case *"Move to X position"* is presented in Table 2. The RRAM conducted at the component's layer is more com-

plex, because we need to consider various dependencies and associations. However, an application of the template presented in Table 1 significantly simplifies the analysis: it allows us to link directly failures of "Aspirate" and "Dispense" to the failures of component's layer use cases.

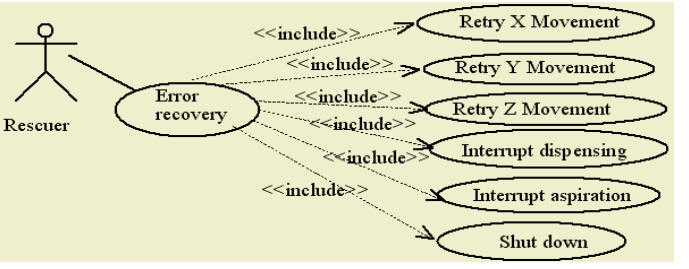


Fig. 5. Elaborated representation of "Error recovery"

3.3 Validating Proposed Approach

An application of the RRAM as proposed above allowed us to significantly improve requirements of Fillwell. For instance, as a result of preliminary hazards identification only safety requirements *SR1* and *SR2* have been stated. The additional safety requirements *SR3* and *SR4* are discovered as a result of the proposed application of the RRAM:

SR1. The operating head must stay inside allowed margins, i.e. it should not be moved outside of minimal and maximal allowed positions.
SR2. The amount of liquid in the pipette must not exceed the volume of the pipette
SR3.While the operating had is moving along Z-axes movement along X,Y-axes is prohibited
SR4. Any movement of operating head while the pump is operating is prohibited

The proposed approach encouraged us to produce an explicit description of hazards mitigation. As a result new functional requirements have been added. Namely, we elaborated on the use case *"Error recovery"* to define the physical functions to be executed for error recovery. The refined part of the use case diagram is shown in Fig. 9.
Moreover, we refined the existing description of the use cases. For instance, we added the precondition "Ensure that the request to move to position X is placed and neither Z-movement nor pumping is performed" to the specification of the use case "Move to X position" and expanded the postcondition to model possibility of failure of the use case. Furthermore, the proposed approach allowed us to establish a link between the use cases and safety requirements. This is a valuable asset since it facilitates a straightforward tracing of requirements through the whole development process.

4 Incorporating RMEA into Use Case Modeling

In the previous section we have shown how to incorporate a deductive safety technique in use case modeling. We demonstrated how an application of the RRAM

method facilitated derivation of system's requirements at functional level. However, as system development progresses, we learn about behavior of components, which are executing these functions as well as contributing to their failures. This information is usually obtained by an application of inductive safety techniques such as Risk Modes and Effect Analysis (RMEA). Hence, to incorporate safety analysis in the entire process of requirements engineering, we also need to merge use case modeling and an inductive safety technique. Next we present our approach to incorporating results of RMEA in use cases.

4.1 Extending RMEA

RMEA is an inductive analysis method, which allows us to systematically study the causes of components faults, their effects and means to cope with these faults [3]. RMEA step-by-step selects the individual components of the system, identifies possible causes of each failure mode, assesses consequences and suggests remedial actions. The results of RMEA are usually represented by tables of the following form:

Unit	Failure mode	Possible cause	Local effects	System effects	Remedial action

RMEA provides us with important information about failure modes of components and gives guidance on how to cope with these failures. Usually this information is hardware-oriented. Hence its translation into software requirements demands a good understanding of software-hardware interaction and physical processes evolving in the system. This task is challenging and therefore, needs to be assisted by software designers and safety engineers. Observe that use cases constitute an effective tool for interdisciplinary communication. Therefore, use cases can be used for establishing common communication framework between software designers and safety engineers. Next we propose a structured approach to extending a traditional representation of the results of RMEA to facilitate discovery of software requirements. The main idea behind the proposed approach is to encourage safety engineers to explicitly express how controlling software should detect errors caused by component's faults and which remedial actions software should initiate for error recovery. Unfortunately currently such a description is often confined by a phrase "modify software to detect error and initiate error recovery". It is rather unlikely that such a description will result in the traceable software requirements unless additional clarifications will be provided by a safety engineer. Therefore, while analyzing failure modes of each component we encourage safety engineer to describe

- which use cases employ the component under investigation
- what kind of information the component supplies to the use cases or which action it provides
- which use cases are involved in detection of failure of the component
- which use cases are involved in recovery from component's failure.

In Table 3 we demonstrate how the proposed extension can be incorporated in the traditional representation of the results of RMEA. Our previous experience shows that usually there are two types of most frequently used detection mechanisms: *timeout* and *a value outside of the valid scope*. For each detection mechanism we incorporate

a description, which directly allows us to generate traceable software requirements. For instance, for the detection by capturing incorrect parameter range, called *outside of valid scope,* we generate the following requirements:

Table 3. Template for representing results of extended RMEA

Component	*C*: Name of the component		
Involved in use cases	*UCName1* ... *UCNameN*: Names of use cases which use the component C in their execution		
Provides information/action	In case *C* is a sensor describe which physical value it monitors In case *C* is an actuator describe which action in performs		
Failure mode	*FM* Description of the failure mode		
Possible cause	What causes the failure		
Local effects	How the failure effects the execution of the use cases on the physical layer. Identify affected use cases UC_1, .. , UC_K		
System effect	How the failure of the use cases UC_1, ... , UC_M affects more abstract use cases and the whole system		
Detection	**Detection type: *Ouside of valid scope*** Name of parameter: *ParName* Minimum value *MinVal* Maximum value *MaxVal* Use case calculating scope *UCName1* upon event *Event_Calculate* (in typical/alternative flow of events.) Use case verifying scope *UCName2* upon event *Event_Verify* (in typical/alternative flow of events.) or **Detection type: *Timeout*** Timer *Timer* Timing constraint *Constraint* Timer activated in use case *Name_of_use_case* (where in the flow of events) Timer deactivated in use case *Name_of_use_case* (where in the flow of events) Expired deadline is detected by the use case *Name_of_use_case* (where in the flow of events)		
Remedial action	Action	Effect	Use cases involved
	Description of an action to be performed	Description of the effect which the execution of the action should have	List of the use cases involved in the execution of remedial actions

5 Conclusions and Future Work

Given the investment organizations are making in use cases and increasing use of computers to control safety-critical applications, the research on integrating use cases and safety consideration is highly demanded. In this paper we presented an approach to incorporating safety requirements into use cases. We demonstrated how the results of deductive and inductive safety techniques can be used to discover and strengthen functional and safety requirements. The proposed approach has been illustrated by a realistic case study – a liquid handling workstation Fillwell. In our previous work we demonstrated how to incorporate various safety techniques into formal approaches to system specification and development [12,14,15]. Moreover, we succeeded in formalizing UML class and state diagrams [8,10]. In our future work it would be interesting to integrate formal design techniques with use case driven approach proposed in this paper.

References

[1] John Cheesman, John Daniels, UML Components A Simple Process for Specifying Component-Based Software, Addision-Wesley, 2001

[2] J.-C. Laprie. "*Dependability: Basic Concepts and Terminology*". Vienna, Springer-Verlag, 1991.

[3] MG, OMG Unified Modeling Language Specification Version 1.5, http://www.omg.org/uml , 2003.
[4] K. Allenby, T. P. Kelly. ”Deriving Safety Requirements using Scenarios”. In *Proc. of the 5th IEEE International Symposium on Requirements Engineering (RE'01)*, Toronto, Canada, 2001, pp.228-235.
[5] Gunter Preuner, Michael Scherfl, "Integration of Web Services into Workflow through a Multi-Level Schema Architecture," Proceeding of Advanced Issues of E- Commerce and Web-Based Information Systems, Fourth IEEE International Workshop on, pp. 51-60, Jun. 2002.
[6] Ruben Prieto-Diaz, "The Common Criteria Evaluation Process," Commonwealth Information Security Center Technical Report, 2002.
[7] PMI, A guide to the Project Management Body of Knowledge(PMBOK Guide 2000 Edition), at URL: http://www.pmi.org , 2000.
[8] Troubitsyna. “Integrating Safety Analysis into Formal Specification of Dependable Systems”, in *Proc. of Annual IEEE Workshop on Fault-Tolerant Parallel and Distributed Systems.* Nice, France, April 2003.

A Component Identification Technique from Object-Oriented Model

Mi-Sook Choi[1] and Eun-Sook Cho[2]

[1] Woosuk University, 490, Hujong-ri, Samnye-up, Wanju-kun, Chonbuk , Korea
khc67_kr@hanmail.net
[2] Dongduk Women's University 23-1 Wolgok-dong, Sungbuk-gu, Seoul, Korea
escho@dongduk.ac.kr

Abstract. Today's software system environment requires rapid development and high productivity. In order to satisfy these requirements, researches have been working on the development of software reuse technology. Therefore, it is general to develop component-based software due to the advance of reuse technology nowadays. One of main issues raised in component-based development (CBD) is how to identify reusable and independent components. Existing methodologies have dealt with the problem based on only developer's heuristics, so that it is difficult to identify the components by common developers. Therefore, in this paper, we propose a new technique to identify the business components based on system component. The proposed technique applies the characteristics and degree of dependency between classes in object-oriented model. We also present a case study and experimental results to prove the practical use of our technique. We examined various examples to get the objective foundations.

1 Introduction

Software development has remained a 'craft' industry, beset with problems of delayed and cancelled projects, inadequate quality, long cycle times, and high costs. Software packages have not been an adequate alternative, often involving long, costly and difficult implementation projects, and even more difficulties in integration and upgrading [1]. Software components address these core issues in four fundamental ways. The first way is by delivering applications designed to be adaptable to business and technology change. The second way is by increasing productivity and speed of delivery without loss of quality. The third way is by providing a practical transition framework which combines legacy applications and technologies, together with modern approaches. It is strong emphasis on transition that protects existing investments and enables progressive evolution and rapid response to the changing needs of business. Finally, componentization removes the need to make binary decisions between build and buy enabling an integration approach.

In the CBD process, it is very important to identify independent components having low dependency, self-governing properties, and high reusability[2]. However, the existing methodologies have identified the components using only heuristic methods

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 778 – 787, 2005.

based on developer's intuition and experience. Furthermore, the component architecture is classified into system and business components. Most of the existing methodologies provide techniques to identify business components only. Therefore, we propose a new identification technique to refine and extend the limitations of the existing techniques. With this new technique, the system components are identified first, and then the identification of business components is next. To identify more independent components, we use structural characteristics of classes, dependency between classes according to message types, and impact scope of class according to message direction.

2 Related Works

CBD system can be separated as two types of components, system and business components. The system component is a functionally reusable component by making a group of similar functions, whereas the business components are configured to perform the functions of system component. The system component is unit of functional reuse as well as subsystem unit. However, the business component is a component being a unit of independent parts to execute the interface of system component [3].

CBD96 [4] have no system component concept to reuse the type of functions for system service. Furthermore, there are difficulties with the development tool because the business component should be identified from whole domains using developer's intuition and experiences.

Advisor methodology [3] provides a clear definition of system component to be functionally reusable in system service, but several complicated analysis is needed to complete the system component by the combination of business components because the system and business components are identified by the clear separation of the two components. Also, the identification of business component is in difficulty because it is also dependent upon the developer's intuition and experiences. Also, this method does not provide any guidelines for what classes are assigned into system components.

Rational's RUP [5] provides only activity of component identification, but it does not contain any detailed instructions and steps. Therefore, the identification of system and business components is totally dependent upon the developer's intuition and experiences.

ETRI [6] proposes a component identification method, MARMI-III, using object model and UDA table. The method of component identification using the object model is similar to the method of CBD96. The component identification method using the UDA table identifies a component using analysis of vertical relationships between use cases and classes only, but it does not consider horizontal relationships, such as method call type or direction between classes. Also, MARMI-III assigns use-cases only referring the classes into any components without defining the concept of system component. However, if it identifies the business component after identifying the system component, the problem can be solved.

3 A Method for Component Identification from Object-Oriented Model

3.1 Dependency Characteristics Between Classes

3.1.1 Dependency According to Structural Relationship Between Classes

The structural relationships between classes are classified into aggregation, generalization, and association. In details, the aggregation and generalization are vertical relationship, whereas the association is horizontal relationship between classes.

A. Vertical dependency between classes
If class A contains class B and class C and then the class A has been modified, the modification of class A should influence both class B and class C. Therefore, the classes related to aggregation or generalization are in vertical relationship as master and slave, and also they are tightly connected each other. Therefore, the classes in the vertical relationships should be included within a component because they are in very dependent relations.

B. Horizontal dependency between classes
If class A associates with class B, the relationship between class A and B is referred to as an association. It is a horizontal relationship between class A and B according to a message call. That is, when an object of class A sends a message to create, modify, delete, or refer an object of class B, the association relationship between the class A and B is occurred. In the association, the degree of change impact on an object is different according to message types rather than the change of a class does not impact on other classes as it is. Therefore, the classes in horizontal relationship can be separated into different components or they can be grouped into a component case by case.

3.1.2 Dependency According to Message Types

Message types can be classified as follows; $<A,B> \in R$, where A and B are objects, and R is relationship.

1. In case of an object A sends a message creating object B
2. In case of an object A sends a message deleting object B.
3. In case of an object A sends a message modifying object B.
4. In case of an object A sends a message referring object B.

3.1.3 Dependency According to the Message Direction

If there is an assumption that Class A depends on Class B, a change in the interface of Class B could necessitate a change in Class A. For example, A's implementation is dependent upon the facilities that made available by B [7, 8].

Therefore, if an object A sends a message creating object B, the object B depends on the object A, and then the creation of object B structurally impacts other objects manipulating the object B. In the other case, if an object B sends a message referring a state of object A, the object A depends on the object B through a message of reference; however, the object A is not impacted by the object B because the object B only refers the state of object A.

Therefore, the message call directions should be considered because the dependency relationships are different according to the message directions between the objects.

3.2 Identification of Business Component

In this section, we propose the procedures of business component identification through dependency characteristics that mentioned above in the section 3.1. The process of the component identification is shown in Fig. 1.

In the first phase, the system components are identified. To identify the system components, we identify the use cases of entire system first, and then perform the analysis of use case realization for each identified use case. At last, we group the closely related use cases into a use case, and identify the system components by considering the relations between each use case and object deduced during the uses case realization that means the robustness analysis [9].

In the second phase, an object model for the identified system components is designed. We identify business concepts from use case realization analysis and domain knowledge by applying the class identification techniques of object modeling technique (OMT) [10]. After identifying classes, attributes, associations, and operations, we construct a conceptual class diagram.

In the third phase, the classes having strong dependency with the aggregation and generalization relationships that described above in the section 3.1 are grouped into a class.

In the fourth phase, we make CRWD (Create, Read, Write, and Delete) matrix [11] to present what method is performed to each class by the system interface. The reason is that the operations for manipulating objects are classified as "create, delete, write, and read" in order to execute the functions of use case as a system interface. When we construct the matrix, we assign a value according to the relationship between the use case and class, and the relationships are defined as one of four types, "Create," "Delete," "Write," or "Read" in this paper.

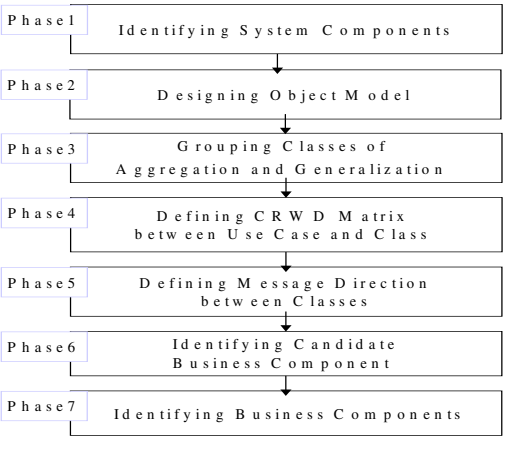


Fig. 1. Process of Component Identification

In the fifth phase, we define message call directions between classes in CRWD matrix.

In the sixth phase, we group classes related in creation and deletion into a class. The grouped class is defined as a candidate business component.

Finally, in the seventh phase, a complete business component is identified after refining the defined candidate business components.

4 Case Study and Assessment

4.1 Case Study

In the first phase, system components are identified using use case and its realization. The use case model of dental hospital system to identify the system components is shown in Fig. 2.

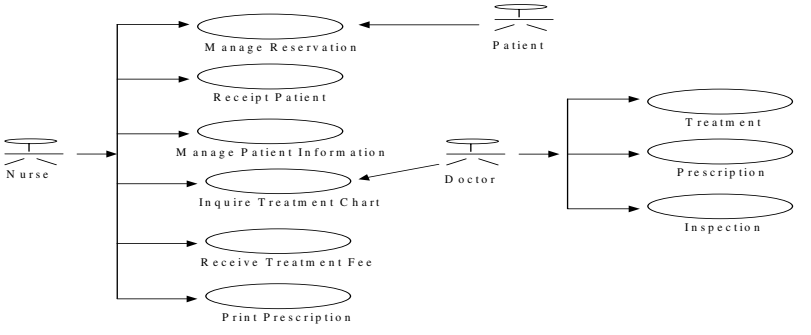


Fig. 2. Use Case Diagram

Usecase / Objects	Reservation	Receipt	Patient Inform.	Treatment Chart	Prescription	Treatment Item	Treatment Fee	Inspection	Inspection Result	Medicine	Insurance	Non Insurance
Manage Reservation	V		V									
Receipt Patient		V	V	V								
Register Patient			V	V								
Inquire Patient			V									
Update Patient Inf.			V									
Receive Treatment Fee			V	V		V	V	V				
Inquire Treatment Chart			V	V	V	V	V	V	V			
Inspect Patient			V	V				V	V			
Patient Treatment			V	V	V	V	V	V	V	V	V	V

Fig. 3. Relationships between Use Case and Class

Fig. 3 describes the table showing the relationships between use cases and objects. The system components are identified through the relationships between use cases and objects. The identified system components are “Receipt”, “Registration”, “Treatment”.

In the second phase**,** a class diagram for the identified system components is completed (Fig. 4) . Classes are contained in each system component. This diagram is a basis to identify business components. This paper identified business components for treatment system component among system components.

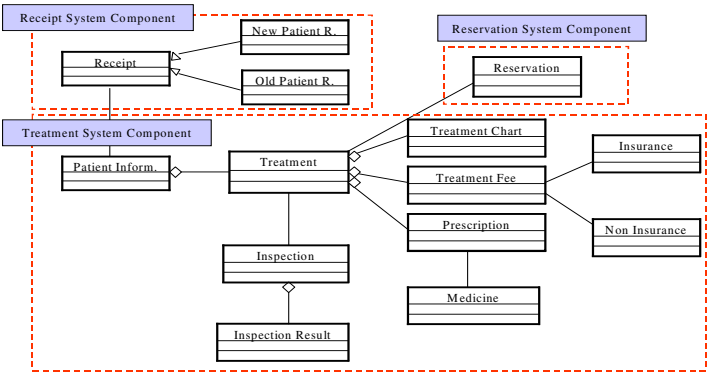


Fig. 4. Classes and System Components

In the third phase, the classes existed in aggregation or generalization relationships are grouped into a class. Fig. 5 shows the results of the grouped classes.

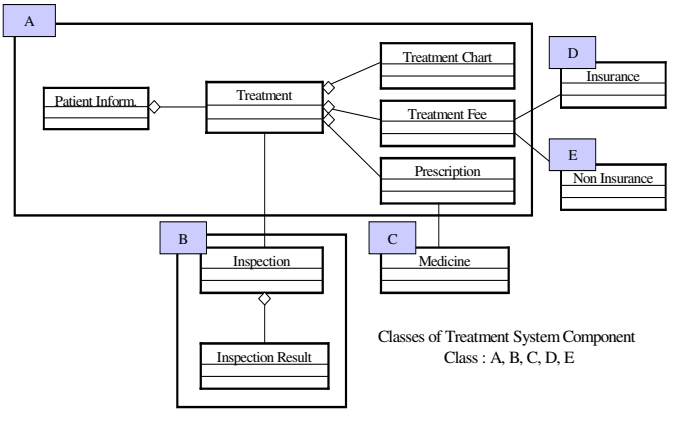


Fig. 5. Group of Classes

In the fourth phase, we make CRWD matrix table based on use cases and classes contained in the system components. As shown in Table1, there is C, R, W, or D value in each cell of CRWD matrix table.

Table 1. CRWD Matrix

Classes / Usecases	A					B		C	D	E
	Patient Inform	Treatment Chart	Prescription	Treatment Item	Treatment Fee	Inspection	Inspection Result	Medicine	Insurance	Non Insurance
Register Patient	C	C								
Inquire Patient	R									
Update Patient Inf.	W									
Receive Treatment Fee	R	R		R	R	R				
Inquire Treatment Chart	R	R	R	R	R	R	R			
Inspect Patient	R	R				C	C			
Patient Treatment	R	R	C	C	C	R	R	R	R	R

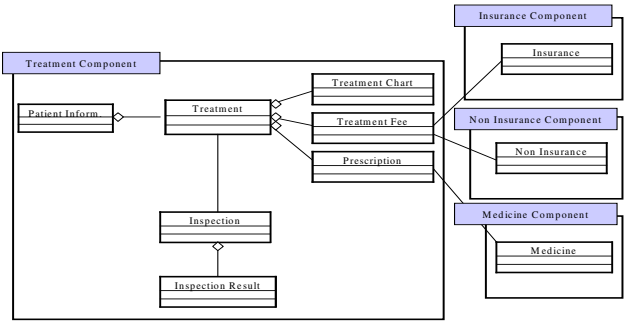


Fig. 6. Group of Classes

After assigning value, in the fifth phase, we define message call directions between classes in Table 1.

In the sixth phase, the classes related to C or D relationships in the message call directions are grouped into a class. Each defined class becomes a candidate business component as shown in Fig. 6.

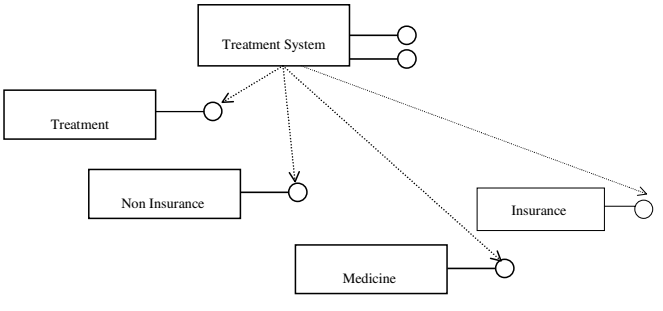


Fig. 7. Identified Business Components

Finally, we identify the complete business components and design the component architecture by refining the candidate components (Fig. 7).

Medicine, Non-Insurance, Insurance components are related in reference relationships with Treatment component, so that those components cannot be grouped; thus they are separated from the Treatment component. Therefore, we eventually identify the business components in the treatment system component of the dental hospital system, which are "Treatment", "Medicine", "Non-Insurance", and "Insurance" components. This result is identical with our intuitive partitioning.

4.2 Assessment

We measure the cohesion and coupling of the identified components to assess whether the components are well identified or not. To inspect our method for business component identification, the cohesion and coupling are computed in two different conditions, with applying CRWD factors for the component identification and without applying them. The cohesion and coupling are measured by the application of Henderson-Sellers' cohesion and coupling metrics [12]. Also, we differentiate weight values according to the order of message types, creation and deletion > modification > read, because the degree of dependency between components is different due to the message types in the association relationship. The metrics [12] are applied into Table 1 that presents the definition of message types between classes. The results are depicted in Table 2 and 3. As shown in Table 3, the results measured without applying CRWD factors present the facts that the business components are identified; thus the results are not matched with our intuition. However, the results obtained with applying CRWD factors show the fact that they can be grouped in a business component and identified even though messages are few between classes, so that the results are matched with the expected components identified by our intuition. Therefore, the proposed method is objective and more useful than other methods. The definitions of classes (A, B, C, D and E) in candidate system component are described in Fig. 5.

Table 2. Applying CRWD factors

System Component	Cohesion	Coupling	Cohesion/Coupling
S_1 (A,B,C,D,E)	0.1789	0.01481	12.0750
S_2 (AB,C,D,E)	0.1000	0.00556	18.0000
S_3 (AC,B,D,E)	0.1646	0.01790	9.1940
S_4 (AD,B,C,E)	0.1646	0.01790	9.1940
S_5 (AE,B,C,D)	0.1646	0.01790	9.1940
S_6 (ABC,D,E)	0.0895	0.00556	16.1111
S_7 (ACD,E,B)	0.1821	0.02407	7.5641
S_8 (ADE,B,C)	0.1821	0.02407	7.5641
S_9 (ABCD,E)	0.0889	0.00556	15.9999
S_{10} (ACDE,B)	0.2326	0.04259	5.4620

Table 3. Without Applying CRWD factors

System Component	Cohesion	Coupling	Cohesion/Coupling
S_1 (A,B,C,D,E)	1.0000	0.12037	8.3077
S_2 (AB,C,D,E)	0.9167	0.11111	8.2500
S_3 (AC,B,D,E)	0.8889	0.12346	7.2000
S_4 (AD,B,C,E)	0.8889	0.12346	7.2000
S_5 (AE,B,C,D)	0.8889	0.12346	7.2000
S_6 (ABC,D,E)	0.8272	0.11111	7.4444
S_7 (ACD,E,B)	0.8025	0.12963	6.1905
S_8 (ADE,B,C)	0.8025	0.12963	6.1905
S_9 (ABCD,E)	0.6944	0.11111	6.2501
S_{10} (ACDE,B)	0.6667	0.11111	6.0001

Existing methodologies are dependent upon developer's intuition and heuristics. However, we consider not only vertical relationships but also horizontal relationships between classes; thus, our method can identify correct and independent business components than other existing methods do.

Existing methodologies are dependent upon developer's intuition and heuristics. However, we consider not only vertical relationships but also horizontal relationships between classes; thus, our method can identify correct and independent business components than other existing methods do.

5 Conclusion Remarks

The proposed method refined and extended the limitations of existing component identification methods to transform object-oriented systems into component-based systems. Especially, we have developed a method by considering dependency characteristics, such as structural relationships, message call types, and message directions between classes in object oriented model. After applying these factors for the component identification, the method proposed can identify more independent components with less effort and time consumption as comparing with other existing methods do. Also, the proposed method was applied to identify target components in various systems, such as internet shopping mall system, hospital system, banking system, and so on. The cohesion and coupling of components identified by the method proposed are better than those of components identified by other methods. Therefore, we have concluded that the proposed method is objective and more useful than other existing methods. Also, the proposed method could complement the conventional developing methodologies that identify components by developer's intuition. Moreover, if an automatic tool for the identification of components using the proposed method is developed, it could identify components very effectively.

Future work is needed to formalize and automate this proposed methodology. Also, we will study other characteristics of components and apply the information to the identification method proposed.

References

1. Bulter Group, "Component-Based Development: Application Delivery and Integration Using Componentized Software", Sep. 1998.
2. Desmond Francis Dsouza, Alan Cameran Wills: Objects, Component, and Frameworks with UML: the Catalysis approach, Addison Wesley, 1999.
3. John Cheesman, John Daniels, *UML Components: A Simple Process for Specifying Component-Based Software*, Addison-Wesley, 2001.
4. John Dodd: "Identifying & Scoping CBD96 Components", Texas Instruments Inc., 1999.
5. Ivar Jacopson, Grady Booch, James Rumbaugh, *The Unified Software Development Process*, Addison Wesley, 1999.
6. ETRI, "CBD Methodology-MARI-III", *Technical Report* , 2002.
7. D. Kung, J. Gao, Pei Hsia, F. Wem, Y. Toyoshima and C. Chen, "Change Impact Identification in Object Oriented Software Maintenance", *Proceedings International Technical Conference on Ciecuit/Systems, Computers and Communications*, 1999.

8. David C. Kung, Jerry Gao and Pei Hsia, “Class Firewall, Test Order, and Regression Testing of Object-Oriented Programs”, *Journal of Object-Oriented Programming*, pp. 51-65, 1995.
9. Martin Fowler, *UML Explained*, Addison-Wesley, 2001.
10. Rumbaugh J., et.al., *Object-Oriented Modeling and Design*, Prentice-Hall, 1991.
11. Eun Sook Cho, Soo Dong Kim and Sung Yul Rhew, “A Domain Analysis and Modeling Methodology for Component Development”, *Internation Journal of Software Engineering and Knowledge Engineering*, Vol.14, No.2, April, 2004.
12. Mi Sook Choi, Eun Sook Cho, “A Technique for Identifying Business Components based on Independency”, Vol.11-D, No. 3, Korea Information Processing Society, 2004.

Retrieving and Exploring Ontology-Based Human Motion Sequences

Hyun-Sook Chung[1], Jung-Min Kim[2], Yung-Cheol Byun[3], and Sang-Yong Byun[3]

[1] Department of Computer Science, The Catholic University, Korea
hsch@catholic.ac.kr
[2] School of Computer Engineering, Seoul National University, Korea
jmkim@oopsla.snu.ac.kr
[3] Division of Communication & Computer Engineering, Cheju National University, Korea
{ycb, byunsy}@cheju.ac.kr

Abstract. A framework for semantic annotation of human motion sequences is proposed in this paper. Motion capture technology is widely used for manufacturing animation but it has a significant weakness due to the lack of an industry wide standard for archiving and retrieving motion capture data. It is difficult for animators to retrieve the desired motion sequences from motion capture files as there is no semantic annotation on already captured motion data. Our goal is to improve the reusability of motion capture data. To archive our goal first, we propose a standard format for integrating different motion capture file formats. Second, we define motion ontologies that are used to annotate and semantically organize human motion sequences. This ontology-based approach provides the means for discovering and exploiting the information and knowledge surrounding motion capture data.

1 Introduction

Motion capture technology is frequently used to solve real-time animation problems, as it enables motion data to be created easily with physically perfect motions. These days, motion capture plays an important role in the making of movies and/or games by the larger companies within the respective fields. However, motion capture technology has a significant weakness. The motion capture process is labor-intensive, expensive and time-consuming. If the existing captured data does not satisfy the user's requirements, then further modifications may be required at cost of labor, time and money. To overcome this drawback many researchers have studied and proposed motion synthesis techniques[5,7]. Though these approaches possess significant preprocessing costs due to the low flexibility of motion capture data and the difficulty in retrieving particular motion clips from a corpus of captured data [4,11,12].

To solve above problems we define a standard format for integrating motion capture data represented by different formats. Our standard format for motion capture data is a markup language that can express motion capture data based on XML , and is called MCML[6]. MCML defines a set of tags to integrate ASF (Acclaim Skeleton File)/AMC (Acclaim Motion Capture data) [1], BVH (Biovision Hierarchical data) [2] and HTR (Hierarchical Translation-Rotation) [9]. These three are the most popular motion capture data formats and have recently become supported by various types of

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 788 – 797, 2005.

motion software. MCML has an extensible structure, by means of which new capture file formats can be easily added.

MCML can be used to store motion capture files in a database, subsequently retrieving the motion clips from the database using a query expression. By having a standard format, we can eliminate the duplication of motion capture files while creating a compact sized motion database. In addition, we propose a semantic annotation scheme to retrieve motion clips relevant to annotation that describe a particular motion such as running, walking, turning and jumping. We construct motion ontologies as data models for semantic annotation. Ontologies are introduced as "explicit specification of a conceptualization"[16]. Ontologies define common vocabulary for animators who need to share motion descriptions**.**

The structure of this paper is as follows. In Section 2, we look at other related studies. In Section 3, the structure and contents of MCML are explained in detail. Section 4 describes the structure and data models of motion ontologies for semantic annotation. Finally, section 5 concludes this paper with future research directions.

2 Related Work

First, we summarize the related studies dealing with new markup language development using XML as the method for representing data in character animation, virtual reality and other fields. Morales [4] proposed a motion capture data storage method based on XML. In this method, motion capture data is stored by being converted into XML data format, so that animation staff had access to the data in a mutually cooperative environment, e.g. in the web-based environment. The system was designed with XML and ASP (Active Server Page) technologies so that motion capture data could be used easily. In contrast to our study, this approach dealt only with motion capture data stored in a simple format based on segments and did not consider a hierarchical structure. Moreover, it did not suggest the use of a standard markup language for motion capture data, such as the MCML language proposed in this paper, but only alluded to the possibility of data conversion using XML.

VRML (Virtual Reality Modeling Language) [18] is a language designed to simulate three dimensional environments on the web, and H-ANIM [19] is a standard specification established by the Web3D Consortium, which describes the structure to be used for the three dimensional modeling of an Avatar. The specification of humanoids in H-ANIM follows the standard method of representing humanoids used in VRML 2.0. The structure of the Humanoid node of H-ANIM is similar to the structure of motion capture data, because that node serves as the overall container for the Joint, Segment, Site and Viewpoint nodes, which define the skeleton, geometry and landmarks of the human figure. The particular interest of our system is in the archiving and exchanging of motion capture files in different formats. H-ANIM is a good language for representing human beings in an online virtual environment, but it is too complex to be a standard motion capture format as it has too many additional features.

Second, in the context of video data, several systems have been developed to retrieve video data based on low-level information, i.e. color, texture, format, etc, or high-level information (i.e. free text description). These systems segment and annotate video based on editing points or scene analysis [9]. Arikan et al. describe an ex-

ample-based approach to synthesizing human motion that satisfies sparse temporal annotation and pose constraints [11,12].

3 MCML DTD Specifications

In this section, we describe the tags and element structure of MCML. We define the tag names after analyzing the bone names and keywords contained in the motion capture file formats. We then define the logical structure of an MCML document.

3.1 Tags of MCML

Tags for Header Data. The motion data file has a header data area containing supplementary information, such as the file type, version, comments, etc. MCML provides a set of tags so that it can include all of the header information for these different file types.

Bone Names of Skeleton. ASF/AMC, BVH and HTR formats describe the skeleton, which is composed of a number of bones, usually in a hierarchical structure. The bone is the basic entity used when representing a skeleton. Each bone represents the smallest segment within the motion that is subject to individual translation and orientation changes during the animation. These three formats have different bone names and different hierarchical structures for the skeleton. To integrate these into one unified format, we create a more detailed hierarchical structure for the skeleton and define the bone names. Table 1 shows the MCML bone names and their corresponding names for each of these file formats.

ASF/AMC file formats use the names of the human bones while the BVH File format uses the names of marker locations to represent the joints of the body. Although HTR uses the names of human bones, it has only a few names, as it is a newly released format still in its initial stage. In order to be able to integrate these different types of files, MCML has extended power of expression so that it can contain all of these three formats.

Tags for Character Skeleton. Motion capture data contains hierarchy information about the modeled character. MCML has element and attribute sets that can represent the hierarchical structure, joint length and comparative distance from the root and each joint for the modeling of a human character.

The MCML Document of Fig. 1 shows an example of the hierarchical structure of a character. The *root* element designates the initial location of the character and the *skeleton* designates the hierarchy structure, location and size of each joint. The hierarchy structure can be expressed using a comprehension relation of the bone elements. The *name* element has the name of each joint, and the value of this name element should be one of the bone names defined in Table 1.

Tags for Motion Data. Motion data is composed of: the total number of frames, the time per frame, number of translations per frame, number of rotations per frame, etc. Motion data describes the animation of each bone over a period of time. We can examine a series of lines pertaining to a frame of animation for each of the segments defined in the skeleton in motion capture files.

```
<skeleton>
<root order="Tx Ty Tz Rz Rx Ry" axis="XYZ" position="0.00 0.00 0.00"
                                          orientation="0.00 0.00 0.00"/>
  <bone id="1">
    <name> hips </name> <offset> 0.00 0.00 0.00 </offset>
    <channels num="6"> Tx Ty Tz Rz Rx Ry </channels>
    <bone id="2">
      <name> torso_1 </name><offset> 0.00 4.21 0.00 </offset>
      <channels num="3"> Rz Rx Ry </channels>
      <bone id="3">
        <name> neck1 </name><offset> 0.00 17.08 0.00 </offset>
        <channels num="3"> Rz Rx Ry </channels>
```

Fig. 1. An example of a character skeleton described with MCML

3.2 MCML Document Structure

The *mcml* element, which is the root element of MCML, is composed of the *meta* element, which expresses the metadata, the *header* element, the *skeleton* element and the *motion* element.

Table 1. Mapping between the bone names of MCML and the bone names of ASF/AMC, BVH and HTR File formats expressing human joints (summarized)

	MCML	ASF/AMC	BVH(1)	BVH(2)	HTR
1	root	h_root	root	root	*undefined*
2	head	h_head(head)	head	head	head
3	neck1	h_neck1(upperneck)	neck	neck	*undefined*
4	neck2	h_neck2	*undefined*	*undefined*	*undefined*
5	left_shoulder	h_left_shoulder(lclavicle)	leftcollar	lshoulderjoint	*undefined*
6	left_up_arm	h_left_up_arm(lhumerus)	leftuparm	lhumerus	lupperarm
7	left_low_arm	h_left_low_arm(lradius)	leftlowarm	lradius	llowarm
8	left_wrist	(lwrist)	*undefined*	*undefined*	*undefined*
9	left_hand	h_left_hand(lhand)	lefthand	lwrist	lhand
:	:	:	:	:	:
50	right_toe_two	h_right_toe_two	*undefined*	*undefined*	*undefined*
51	right_toe_three	h_right_toe_three	*undefined*	*undefined*	*undefined*
52	right_toe_four	h_right_toe_four	*undefined*	*undefined*	*undefined*
53	right_toe_five	h_right_toe_five	*undefined*	*undefined*	*undefined*

```
<! ELEMENT meta (title, creator, subject, description, date, format, duration, category)
<! ELEMENT title (#PCDATA)>
<! ELEMENT creator (#PCDATA)>
<! ELEMENT subject (#PCDATA)>
<! ELEMENT description (#PCDATA)>
<! ELEMENT date (#PCDATA)>
<! ELEMENT format (#PCDATA)>
<! ELEMENT duration (#PCDATA)>
<! ELEMENT category (#PCDATA)>
```

Fig. 2. The structure of MCML metadata.

Meta element. MCML metadata is based on 8 elements and describes the contents of an MCML document. This is depicted in Fig. 2.

The *title* element is the name given to the MCML document by the creator. The *creator* element is the person(s) or organization(s), which created the original motion capture file. The *subject* element is the topic of the actions contained within the motion capture file (for example, ballet, dance, etc.). The *description* element is a textual description of the actions contained in the motion capture file. The *format* element is the data representation of the motion capture file, such as ASF/AMC, BVH or HTR. The *duration* element is the playing time of the frames contained in the motion capture file. The *category* is used to classify the type of motion capture data (for example, sports, locomotion, human interaction, etc.)

Header element. The *header* element is composed of 15 sub-elements, these are depicted in Fig. 3. The *header* element is used to convert an MCML document into an HTR file.

```
<! ELEMENT header (filetype, datatype?, filename, version?, skeleton_name?, units?, num_segments?,
num_frames, dataframe_rate, euler_rotation_order?, calibration_unit?, rotation_unit?,
global_axis_of_gravity?, bone_length_axis?, scale_factor?)
```

Fig. 3. The structure of MCML header element.

Skeleton element. The MCML *skeleton* element represents the hierarchical structure of a human figure. In Fig. 4, we examine the logical structure of the *skeleton* element.

```
<! ELEMENT skeleton (root, bone+)>
<! ELEMENT root EMPTY>
<! ATTLIST root order CDATA #REQUIRED    axis CDATA #REQUIRED
          position CDATA #REQUIRED      orientation CDATA #REQUIRED>
<! ELEMENT bone (name, direction?, length?, position?, axis?, order?, dof?, limits?,
bodymass?, cofmass?, offset?, channels?, bone+)>
<!ATTLIST bone id ID #IMPLIED>
```

Fig. 4. The structure of MCML skeleton element.

The *root* element describes the parent of the hierarchy. The *axis* and *order* attributes describe the order of operations for the initial offset and root node transformation. The *position* attribute describes the root translation of the skeleton and the *orientation* attribute defines the rotation.

The *skeleton* element has one or more *bone* elements. The hierarchical structure of the *bone* element may be recursive to represent the skeleton information. The PCDATA of the *name* element is the bone name according to the bone-naming rule shown in Table 1.

Motion element. The MCML *motion* element is composed of one or more *frame* elements and zero or more *motion_name* elements. This element actually describes the animation of each bone over time. The logical structure of the *motion* element is shown in Fig. 5.

```
<! ELEMENT motion (frames, frametime, frame, motion_name*)>
<! ELEMENT frames (#PCDATA)>
<! ELEMENT frametime (#PCDATA)>
<! ELEMENT frame (frame_name, frame_bone+)>
<! ATTLIST frame id ID #IMPLIED>
<! ELEMENT frame_bone EMPTY>
<!ATTLIST frame_bone name CDATA #REQUIRED
Tx CDATA #IMPLIED   Ty CDATA #IMPLIED   Tz CDATA #IMPLIED
Rx CDATA #REQUIRED  Ry CDATA #REQUIRED  Rz CDATA #REQUIRED>
```

Fig. 5. The structure of the MCML motion element.

The *frames* element is the number of frames while the *frametime* element is the playing time for each frame. The *frame* element has one or more *frame_bone* elements to represent the actions of each bone defined in the *skeleton* element. One *frame_bone* element represents one frame in the motion capture files.

4 Ontology-Based Motion Annotation

Annotation is the text data associated with particular motion clips of the motion capture data. If motion databases enable annotation over the motion clips, the animators can easily retrieve the motion clips using keywords describing certain actions such as running, walking, dancing and jumping.

There are two schemes to annotate over motion clips, segmentation and stratification. The segmentation approach is the oldest and simplest. With this approach, a motion sequence is split into independent and contiguous time segments that are annotated individually. The stratification approach creates layers of descriptions called stratum, where each stratum describes the temporal occurrences of some concept, like move, go or run. The stratification allows an overlapping of descriptions; therefore the animators can specify several levels of descriptions.

But the stratification approach does not provide methods allowing the sharing of descriptional data between different motion clips. Descriptional data may have se-

mantic relationships such as synonym, intension or extension, though this approach does not include these relationships.

We construct motion ontologies that are semantic models to specify concepts and their relationships and to connect motion clips to corresponding concepts.

4.1 The Structure of Motion Ontologies

Ontology is an explicit specification of a conceptualization and is a hierarchically structured set of controlled vocabularies for describing domain knowledge. Fig. 6 shows that motion ontology is composed of several specific human movement ontologies. The terms describe basic human movement such as; run, walk, pick or turn left, which may be collected from WordNet. Carefully handcrafted by cognitive scientists, WordNet is both a lexical database, and a fairly comprehensive common sense thesaurus. WordNet distinguishes between words as literally appearing in texts and the actual word senses, the concepts behind words. Often one word has multiple senses, each of which is briefly described in a sentence or two and also characterized by a set of synonyms, words with the same sense, called synsets in WordNet. In addition, WordNet has captured hypernym (i.e., broader sense), hyponym (i.e., more narrow sense), and holonym (i.e., part of) relationships between word senses.

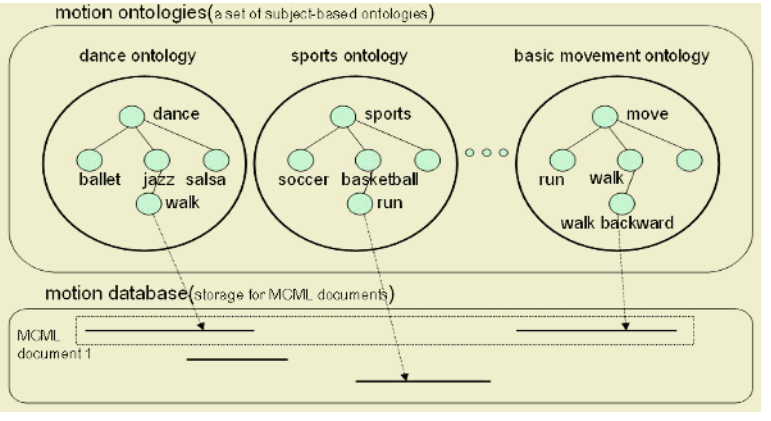


Fig. 6. The structure of motion ontology

Top-most categories of motion ontology are classified based on subject, for example, sports, dance, fighting or locomotion. Basic movement is also included in the top-most categories. Each of top-most categories is an ontology that is composed of a set of terms and a set of relationships between terms. Hierarchical relationships between terms represent 'is-a' or 'part-of' relation, for example, ballet is a dance. Each ontologies has connection to motion clips within MCML documents stored within motion databases. A motion clip is a set of continuous frames between start frame and end frame in a MCML document. Start frame and end frame are id attributes in the frame element, for example, <frame id="f000001"> for start frame and <frame id="f000214"> for end frame. Because a topic in an ontology may have several oc-

currences, the user can retrieve multiple motion clips describing a certain action, walking, from different MCML documents.

4.2 Construction and Retrieval of Motion Ontologies

RDF/RDF Schema[3] and topic maps[8] are data models that are used to represent and encode an ontology. In this paper, we use topic maps to represent motion ontologies as they allow us to specify internal descriptions for resources whereas in RDF the user must use only a URI scheme to specify resources.

Fig 7 shows a XTM fragment to define a topic which includes two occurrences and a 'is-a' relationship. XTM[8] is a XML-based formal language to specify Topic maps.

```
<topic id="walk001">
    <baseName><baseNameString>Walk backward<baseNameString></baseName>
    <occurrence>
        <instanceOf><topicRef xlink:href="#description"/></instanceOf>
        <resouceData>A person walking backward</resourceData>
    </occurrence>
    <occurrence>
        <instanceOf><topicRef xlink:href="#srart-of-motion-clips"/></instanceOf>
        <resouceRef xlink:href="/resources/dance_walk_back.mcml#f000001"/>
    </occurrence>
    <occurrence>
        <instanceOf><topicRef xlink:href="#end-of-motion-clips"/></instanceOf>
        <resouceRef xlink:href="/resources/dance_walk_back.mcml#f000214"/>
    </occurrence>
</topic>
<association>
    <instanceOf><topicRef xlink:href="superclass-subclass"/></instanceOf>
    <member><roleSpec><topicRef xlink:href="superclass"/></roleSpec>
            <topicRef xlink:href="#walk"/></member>
    <member><roleSpec><topicRef xlink:href="subclass"/></roleSpec>
            <topicRef xlink:href="#walk001"/></member>
</association>
```

Fig. 7. A XTM fragment to define a topic

We use Protégé 2000 [13] as the ontology editor. It is a popular editor tool to edit, search and generate ontologies while providing editing functionality for XTM documents with the tmtab plugins.

We developed a web-based searching and browsing interface which enables full-text search and tree structure-based browsing. The user is able to retrieve motion clips relevant to keywords entered into the full-text search field.

5 Conclusion

A standard format based on XML called MCML employed to integrate heterogeneous motion capture files and store them into a XML or a Relational database system is

introduced. We also propose a semantic annotation scheme based on construction of motion ontologies.

If the MCML documents are stored in a motion database, it is easy for the animator to obtain motions that do exactly what he or she wants by searching or browsing over motion databases through motion ontologies. This offers many advantages for motion synthesis or motion editing applications. Also, in order to provide increased security for the data and more convenient data management, commercial animation software can be used in conjunction with a database for the storage of the motion capture data. Thus, MCML and motion ontologies can improve the reusability of the motion capture data.

References

1. Acclaim, ASF/AMC File Specifications page, http://www.darwin3d.com/gamedev/acclaim.zip.
2. Biovision, BVH Specifications page, http://www.biovision.com/bvh.htm.
3. Brickley, D. and Guha, R.V, Resource Description Framework (RDF) Schema Specification. World Wide Web Consortium:http://www.w3.org/TR/PR-rdf-schema.
4. C.R. Morales, Development of an XML Web Based Motion Capture Data Warehousing and Translation System for Collaborative Animation Projects, In Proceedings of the 9-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2001.
5. E. Hsu, S. Gentry and J. Popovic, Example-Based Control of Human Motion, Eurographics/ACM SIGGRAPH Symposium on Computer Animation, 2004.
6. H.S. Chung and Y.B. Lee, MCML:Motion Capture Markup Language for Integration of Heterogeneous Motion Capture Data, Computer Standards & Interface, 26, pp.113~130, 2004.
7. L.M. Tanco and A. Hilton, Realistic Syntehsis of Novel Human Movements from a Database of Motion, In Proceedings of the IEEE Workshop on Human Motion HUMO 2000.
8. K. Ahmed, Developing a Topic Map Programming Model, In Knowledge Technologies 2001.
9. M. Hacid, C. Decleir and J. Kouloumdjian, A Database Approach for Modeling and Querying Video Data, IEEE Transanction on Knowledge and Data Engineering, 12(5), pp.729~750,2000.
10. M. Meredith and S.Maddock, Motion capture file formats explained, Department of Computer Science Technical Report CS-01-11.
11. O. Arikan and D.A. Forsyth, Interactive Motion Generation from Examples, In Proceedings of ACM SIGGRAPH 2002.
12. O. Arikan, D.A. Forsyth, and J.F. O'Brien, Motion Synthesis from Annotations, In Proceedings of ACM SIGGRAPH 2003.
13. Protégé-2000, The Protege Project. http://protege.stanford.edu.
14. S. Rosenthal, B. Bodenheimer, C. Rose, and J. Pella, The process of motion capture: Dealing with the data, In Proceedings of the 8th Eurographics Workshop on Animation and Simulation 1997.
15. S.P. Chao, C.Y. Chiu, S.N. Yang and T.G. Lin, Tai Chi synthesizer: a motion synthesis framework based on key-postures and motion instructions, Comp. Anim. Virtual Worlds 2004, vol.15, pp.259-268.

16. Uschold, M. and Gruninger, M, Ontologies: Principles, Methods and Applications.Knowledge Engineering Review 11(2), 1996.
17. W3C, Extensible Markup Language (XML) 1.0, http://www.w3c.org/TR/1998/RECxml-19980210, 1998.
18. Web3D Consortium, VRML International Standard,http://www.web3d.org/technicalinfo/specifications/ISO_IEC_14772-All/index.html.
19. Web3D Consortium, H-ANIM 2001 Specification, http://www.h-anim.org/Specifications/H-Anim2001.

An Integrated Data Mining Model for Customer Credit Evaluation

Kap Sik Kim[1] and Ha Jin Hwang[2,*]

[1] Department of Computer Information Science, Daegu Polytechnic College,
Manchon 3 Dong, Soosung Gu, Daegu, Korea
kskim@kms.tpic.ac.kr
[2] Department of Management, Catholic University of Daegu,
Kyungsan-si, Kyungbuk, Korea
hjhwang@cu.ac.kr

Abstract. Based on the customer information relating to details of financing and payment histories from a financial institution, this study derived single data mining models using MLP, MDA, and DTM. The results obtained from these single models were subsequently compared with the results from an integrated model developed using GA. This study not only verifies existing single models and but also attempts to overcome the limitations of these approaches. While our comparative analysis of single models for the purpose of identifying the best-fit model relies upon existing techniques, this study presents a new methodology to build an integrated data mining model using GA.

1 Introduction

This study was designed to perform a comparative examination of prediction accuracy of different data mining models including multilayer perception-based neural network model (MLP), multiple discriminant analysis (MDA) and decision tree model(DTM). The study was then extended to build an integrated model using the genetic algorithm (GA), in an attempt to derive a reliable credit risk prediction model, using the actual data of the behavioral scoring on customers purchasing an installment financing. Credit scoring is an important tool used by today's financial institutions to assess the creditworthiness of prospective customers and to monitor customer credit ratings. Credit scoring contributes to preventing bad debt and enables companies to offer differentiated financial products and benefits based on credit information from customers. Providing ways to more effectively manage customer relationship, credit scoring ultimately leads to increased profitability for companies. This is why many financial institutions have been trying to find out credit scoring methods offering a more accurate prediction rate.

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 798 – 805, 2005.

2 Characteristics of Data Mining Models

Most common models for credit evaluation include statistical models such as MDA, logistic regression analysis [22], and management science models.[17]. More recent studies make use of decision trees, artificial neural networks. Some of the most successful studies in this field were found in artificial neural networks [12, 21, 23].

With increasing interest in data mining techniques in recent years, artificial neural network techniques (MLP) have received extensive academic attention. The methods have wide application areas and can be used for a broad range of domains. While artificial neural networks do not require statistical assumptions, they are outstanding explanatory tools for nonlinear regression models, and their ideal aptitude for credit scoring has been widely acknowledged [3]. MDA is a statistical tool that allows values of specific characteristics observed in sets of data to be classified into two or more different groups. MDA enables to predict sample subjects' behaviors in given situations, based on various characteristics of different social phenomena [13]. Building a MDA for data mining is relatively simple, and the training period is short as well. Meanwhile, MDA require that independent variables satisfy basic statistical assumptions, present the disadvantage of necessitating validation [2].

DTM is primarily used in data mining during the classification process. The experiment consists of analyzing a collection of data gathered from the past to uncover patterns of existing elements, and organizing the characteristics of each category into combinations of attributes to form a classification tree. Compared to other classification techniques, the advantage of decision trees is that the analytic process can be more easily understood and readily explained by the researcher performing the procedures [4]. These classification models are faster and simpler to build than others and can be converted into easy and self-explanatory rules [11]. Furthermore, this technique is especially helpful when a large number of factors must be considered as the basis of decision-making, and is a highly useful tool to identify factors that must be considered from those that are irrelevant [18]. GA, belonging to the category of artificial intelligence techniques, is an effective and flexible method for searching through complex 2D or 3D spaces and detecting optimal global solutions [7]. GA is based on the evolutionary concepts of natural selection and survival of the fittest. The basic principle assumes that, when a new population is being formed, there is a high probability that singles strings of superior fitness from the old population would transmit their genes to the new one [6, 9, 10]. GA techniques have extensively evolved and achieved significant progress.

Determining which method is the best technique for data mining is an extremely difficult task. It is also common that studies on the subject are often limited due to the fact that meaningful information such as credit reports issued by credit bureaus is not easily obtainable for researchers for the comparative analysis. However, the complexity of the matter is mainly derived from distinctiveness of the studies conducted on the various topics resulting in diverse research findings for different data mining models. For this reason, the need for an integrated model and methodology for credit evaluation has been increasingly suggested recently [16].

3 Research Design

3.1 Methodology

The study proposed an integrated model to predict risks involved in customer credit evaluation based on the customer information from actual financing processes, and then tried to integrate different models using GA. In order to achieve this integrated model, we derived several classification models using MLP and several others using MDA. One classification model was derived using DTM. Then, we obtained three representative models from MLP and MDA, using GA. Next, using the same method, we integrated these representative models derived from these two types of models with DTM to obtain the final model.

$$O_i = \sum_{j=1}^{K} w_{ij} m_{ij}$$

$$\begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_N \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1K} \\ m_{21} & m_{22} & \cdots & m_{2K} \\ \vdots & \vdots & & \vdots \\ m_{M1} & m_{N2} & \cdots & m_{NK} \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{N1} \\ w_{12} & w_{22} & \cdots & w_{N2} \\ \vdots & \vdots & & \vdots \\ w_{1K} & w_{2K} & \cdots & w_{NK} \end{bmatrix} \quad (1)$$

$$E(x) = \begin{cases} S, \ if \ o_s = \max_{i \in \Lambda}(o_i) \ and \ o_s \geq \alpha \\ reject, \quad otherwise \end{cases} \quad (2)$$

$$HF(WS_q) = \begin{cases} 1 \ , if \ correctly \ matched \\ 0 \ , otherwise \end{cases} \quad (3)$$

As MLP and other models have distinctive properties, and achieve different performances depending on the context of research, it is impossible to issue a definitive judgment on which of them is superior to another. Therefore, in order to obtain the best-fit model for a given research problem, it is necessary to integrate them. The study performed a weighted integration of single models, in which the weight matrix used for the combining module for classifiers was optimized using GA [6], and the models were processed in parallel [16]. The resulting values from these classifiers were defined as in Formula 1 where N is the number of values that a dependent variable can hold (in other words, there are N classification groups); K is the number of classifiers and O_i, is the resulting value. And, as shown in Formula 2, E(x), the value a pattern can hold, is assigned the highest of the values of O_i, and a value is entered only when that value is above a certain cutoff value (α); no value is given in the contrary case. Moreover, if the value of E(x) is equal to the original value, as shown in Formula 3, 1 is entered in the fitness function of GA, and 0 if it fails to correctly match the original value.

3.2 Sample Data and Application of the Model

Sample data used in this study is customer information and financing process data, collected and stored by a Korean installment purchase financing company during a period starting from July 1997 and ending in May 2000. Out of roughly 200,000 pieces of Single customer data, 6,500 pieces of complete data with no missing values were sampled from both good and bad credit groups.

Table 1. Use of Sample Data

Use	Sample data	Balancing
Single Classifier Development (training, validation, test)	3,500	Good 1,500/Bad 1,500/Undecided 500 Training1,750/validation 875/Test 875
Single Model Prediction Accuracy Evaluation (scoring)	1,000	
Genetic algorithm-based integrated model (training)	1,000	Good 450/Bad 450/Undecided 100
Final Prediction Accuracy Evaluation (scoring)	1,000	

Of the total selected data, 3,500 pieces were used for developing Single classification models, and this set of data was again used for training (1,750 pieces), validation (875 pieces) and test (875 pieces). Moreover, for the evaluation of Single models' prediction accuracy, 1,000 other pieces of data were used in addition to the 3,500 pieces previously used. For the integrated model (training model) using genetic algorithm, 1,000 new pieces of data were used, which comprised 450 data belonging to creditworthy customers, 450 others belonging to non-creditworthy customers, and 100 pieces relating to undecided cases. 1,000 other pieces of data were used to perform the final prediction accuracy scoring on the integrated model.

3.3 Variables and Research Methods

Items are variables resulting from reprocessing of the primitive data through a normalization process so as to adapt them into acceptable input data for the prediction model. Dependent variable serves to determine whether a customer's credit is to be rated good or bad, and its value is 1 (bad) if installments are past due for a period of 4 months or more, 2 (undecided) if past due for 3 months, and 3 (good) if past due for 2 months or less during the observation period. Other variables (excluding loan number) are input variables; variables related to amounts of loan are normalized by dividing them with mean values. Among the variables used to reflect the history of the financing process during periods prior to the observation period, and take into account payment histories from periods of 3 or 6 previous months. Also, as values of these input variables vary widely in number of months and amounts, we normalized them by dividing them with the mean value of a corresponding variable to obtain real numbers between 1 and 0. The models require that variables be scaled to a range of 0.0 to

1.0 or -1.0 to 1.0. In order to satisfy this condition, variables are scaled down by dividing them by the number of months or the average amount. Through this process, values of variables such as amounts and numbers of months were normalized into values in the range between 0 and 1. This study derived several single classifiers based on actual customer information and payment transaction histories obtained from the financial market, and these models were subsequently combined to produce an integrated model using GA. Finally, this integrated model was compared with each of the single models to determine the best-fit model.

4 Experiment and Results

The weight matrix derived from the combining process using GA is shown in Table 2. Each row of the weight matrix indicates weights assigned to single models, and each column corresponds to weights assigned to different classification results. In both MDA* and MLP*, high weights were assigned to value 2 (undecided value) of the first of the three models.

Table 2. Optimal Weight Matrix for Models Created through Preliminary Integration

Models Created through Preliminary Integration	Weight Matrix (W)
MDA*	$W = \begin{bmatrix} 0.1 & 0.69 & 0.1 \\ 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$
MLP*	$W = \begin{bmatrix} 0.1 & 0.45 & 0.1 \\ 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$

The weight matrix for the final integration is shown in Table 3. The weight matrix elements reveal that weights were assigned to value 1 (bad credit) of the first MDA and the third DTM. To enhance the accuracy rate of the final integrated model, weights were assigned to values corresponding to classification groups on which Single prediction models performed most poorly. Hence, the value of the 'good' group on which the MDA showed the lowest prediction accuracy is assigned 0.87. Also, the value of the 'undecided' from the MDA is assigned a weight of 0.65.

Table 3. Weight Matrix for Model Created through Final Integration

Integrated Model	Weight Matrix (W)
NN*	$\begin{bmatrix} 0.57 & 0.65 & 0.87 \\ 0.1 & 0.1 & 0.1 \\ 0.09 & 0.1 & 0.1 \end{bmatrix}$

The prediction accuracy rate of the final model as shown in Table 4. was obtained using Palisade Evolver 4.0. The result of the comparative analysis of several different prediction single models(SM) for consumer credit indicated that the combined model integrating SM through GA was superior in performance to single classifiers.

Table 4. Comparison of Prediction Accuracy among Models–Final Integration

Model Type		Total Prediction Accuracy (%)	Group		
			Bad	Undecided	Good
		81.92	81.02	2.02	97.37
	MDA	82.27	82.87	3.03	94.01
		81.81	81.48	6.06	95.04
SM		83.75	78.24	13.13	93.72
	MLP	84.44	78.70	22.22	89.34
		84.53	77.31	28.28	88.76
	DTM	82.20	80.56	53.54	86.86
	MDA	83.80	81.48	2.02	96.35
PI	MLP	82.40	78.24	23.23	89.64
	DTM	82.20	80.56	53.54	86.86
FI		84.40	81.02	2.02	97.37

* SM=Single Model, PI=Preliminary Integration, FI=Final Integration

The credit scoring produced by SM during the initial experiment showed substantially high prediction accuracy in the range of 81.81% ~ 84.53%. During the preliminary integration (PI) of MDA and MLP using GA, weights were assigned to the values of the 'undecided' group. The comparison of prediction accuracy following PI, revealed that the MDA was the best-performing model, recording an accuracy of 83.80%. However, the same model's prediction accuracy on the 'undecided' group was extremely poor, measured at 2.02%. In contrast, the DTM and MLP models showed more consistent performance across the three result groups - bad credit, undecided and good credit. During the final integration (FI), weight was assigned to the 'undecided' value of the MDA and to the 'good' value of the DTM. The integrated model created through assigning weights in this way exhibited a prediction accuracy of 84.40%, appreciably higher than the rates achieved by the models produced from the preliminary integration. By classification group, the integrated model achieved an accuracy rate of 97.37% on the good credit group; 1% higher than 96.35% recorded by the MDA, which was the best-performing model among those produced from the preliminary integration. The accuracy on the undecided group was measured at a comparatively low rate of 2.02%. However, this may not be a real impediment for the practical application of the model, as the industry of necessity must strive to reduce this classification group for the sake of better customer credit rating monitoring.

Meantime, the integrated model failed to bring about improvement in prediction accuracy on the bad credit group, recording 81.02%, which is inferior to 81.48% achieved by the MDA model.

5 Conclusion

This study was designed to apply data mining models for the prediction of customer credit risks. The research findings indicated that the integrated model combining single models using the genetic algorithm turned out to be the most accurate tool of prediction.

Beyond verifying various models, this study attempted to overcome the limitations in the existing methodologies [8, 1, 19, 24, 5], by generating an integrated model using GA, which imposes a new methodology for data mining model integration. The integrated model developed in this study showed a prediction accuracy rate of 84.40%, substantially higher than the 60 to 70% recorded by other models mentioned in previous studies. Therefore, the genetic algorithm-based integration method is highly effective and useful as a technique for predicting consumer credit risks. Furthermore, the prediction model suggested by this study can substantially contribute to building the financial risk prediction tools. An assessment system built based on this model can significantly enhance the accuracy in identifying valuable customers.

The study was conducted based on the data set consisting of 200,000 total pieces, far exceeding the amount of data used in previous studies ranging usually from several dozens to several hundreds. Therefore, the relevance of our approach is supported by outstanding data reliability and validity. For this reason, the results obtained from this study are highly realistic and this prediction model can be easily applied to similar situations without major modifications.

References

[1] Boyle, M., Crook, J. N., Hamilton, R., Thomas, L. C.: Methods for Credit Scoring Applied to Slow Payers. In Thomas, L. C., Crook, J. N., Edelman, D. B. (eds.): Credit Scoring and Credit Control. Oxford University Press, Oxford (1992) 75-90

[2] Chae, S.: Social Science Research Methodology. 2nd edn. Hakhyunsa , Seoul (1999)

[3] Cheng, B., Titterington, D. M.: Neural Networks-A Review from a Statistical Perspective-.Statistical Science, Vol. 9. (1994) 2-30

[4] Choi Jong-hu, Han Sang-tae.: Analysis of a Decision Making Tree for Data Mining Using AnswerTree. SPSS Academy, Seoul (2000).

[5] Desai, V. S., Convay, D. G., Crook, J. N., Overstreet, G. A.: Credit Scoring Models in the Credit Union Environment Using Neural Networks and Genetic Algorithms. IMA Journal of Mathematics Applied in Business and Industry, Vol.8. (1997) 323-346

[6] Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, MA. (1989)

[7] Gupta, Y. P., Gupta, M. C., Kumar, A. K., Sundram, C.: Minimizing Total Intercell and Intracell Moves in Cellular Manufacturing: A Genetic Algorithm Approach. INT. J. of Computer Integrated Manufacturing, Vol.8. No.2. (1995) 92-101

[8] Henley, W. E.: Statistical Aspects of Credit Scoring. PhD Thesis, Open University (1995)
[9] Holland, J. H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI (1975)
[10] Hon, K. K. B., Chi, H.: A New Approach of Group Technology Part Families Optimization. Annals of the CIRP, Vol. 43. No.1. (1994)
[11] Imielinski, T., Mannila, H..: A Database Perspective on Knowledge Discovery. Communications of the ACM, Vol. 39. No. 11. (1996) 214-225
[12] Jain, B. A., Nag, B. N.: Performance Evaluation of Neural Network Decision Models. Journal of Management Information Systems, Vol. 14. No. 2. (1997) 201-216
[13] Jeong, C., Choi I.: A Statistical Analysis using SPSSWIN. Muyok Publishing, Seoul (1998)
[14] Kim G.: Integrated Data Mining Model for Prediction of Customer Credit Risk in Installment Purchase Financing. Ph.D Dissertation, Catholic University of Daegu (2003)
[15] Kim H.: Prediction Model for Credit Risk Assessment in Installment Purchase Financing Integrating Several Classifiers through Genetic Algorithm. MA Thesis, Daegu University (2001)
[16] Kim, E., Kim, W., Lee, Y.: Purchase Propensity Prediction of EC Customer by Combining Multiple Classifiers Base on GA. Proceedings of International Conference on Electronic Commerce, (2000) 274-280
[17] Mangasarian, O. L.: Linear and Nonlinear Separation of Patterns by Linear Programming. Operations Research, Vol.13. (1965) 444-452
[18] Mehta, D.: The Formulation of Credit Policy Models. Management Science. Vol.15. (1968) 30-50
[19] Srinivasan, V., Kim, Y. H.: The Bierman-Hausman Credit Granting Model-A Note. Management Science, Vol.33. (1987) 1361-1362
[20] Thomas, L. C.: A Survey of Credit and Behavioral Scoring: Forecasting Financial Risk of Lending to Consumers. International Journal of Forecasting, Vol. 16. (2000) 149-172
[21] West, D.: Neural Network Credit Scoring Models. Computers & Operations Research, Vol. 27. (2000) 1131-1152
[22] Wiginton, J. C.: A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior. Journal of Financial and Quantitative Analysis, Vol. 15. (1980) 757-770
[23] Wong, B. K., Bodnovich, T. A., Selvi, Y.: Neural Network Applications in Business: A Review and Analysis of the Literature(1988-95). Decision Support Systems, Vol. 19. (1997) 301-320
[24] Yobas, M. B., Crook, J. N., Ross, P.: Credit Scoring Using Neural and Evolutionary Techniques. Credit Research Centre, University of Edinburgh, Working Paper (1997)

A Study on the Component Based Architecture for Workflow Rule Engine and Tool

Ho-Jun Shin, Kwang-Ki Kim, and Bo-Yeon Shim

201, HyunSan B/D, 108-7, YangJae-Dong, SeoCho-Ku, Seoul, 137-891, Republic of Korea
{hjshin, kkkim, byshim}@cspi.co.kr

Abstract. The component based development and architecture technology have the potential to be more powerful than traditional. In this paper, we propose 4 views for architecture development, which are use case view, logical view, component view and implementation view. We present component based architecture through ABCD pattern in a component viewpoint. In addition, we apply 4 viewpoints for workflow rule and tool development and propose each design. We present the user interface of actuality developed workflow engine and tool in implementation view. It actually allowed the various stakeholders to find what they want to know about the software architecture. This architecture viewpoints separation accurately will accurately reflects the practice in development of real software systems. Also, we expect to decrease the complexity of implementation, improve reuse and reconfiguration and productivity in workflow system and similar domain.

Keywords: Component Based Architecture, Component Based Development, Architecture View, Workflow Rule Engine & Tool.

1 Introduction

Software architectures describe how a system is decomposed into components, how these components are interconnected, and how they communicate and interact with each other. It is easy to be convinced that building applications from software components is an effective development approach. But how do you plan the best set of components for an enterprise, that can be assembled to meet today's application needs, and which can accommodate the ever-changing demands of the enterprise? The problem can be tackled in part, by having a systematic process for identifying the major "business components", based on application requirements, or the requirements of multiple applications for some area of the business. Also, it needs preparing initial component architecture for software development[1].

We propose to organize the description of a software architecture by using several views, each one addressing one specific set of concerns. The views are called as use case view, logical view, component view and implement view. The use case view is used to capture an user requirement. The logical view describes a whole system to 5 layers. Also, We define the structural items to be located in each layer. The component view defines the component to organize by using ABCD pattern. The ABCD

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 806 – 816, 2005.

pattern can be applied to the problem of architecting a target system in terms of software components. The implementation view presents context diagram and implementation specification based on contents that is defined from logical view point. We apply 4 views to develop an workflow engine and tool. It actually allowed the various stakeholders to find what they want to know about the software architecture. We wish to present variety of development through architecture viewpoints separation.

2 Related Works

2.1 Concept of Software Architecture

Software architecture is served as a framework for understanding system components and its interrelationships. This understanding is necessary for the analysis of current systems and the synthesis of future systems. In support of analysis, software architecture captures domain knowledge and community consensus and ease simulation and prototyping. In support of synthesis, software architecture provides a basis for establishing Product Line Systems(PLS) and using domain knowledge to construct and maintain modules, subsystems, and systems in a predictable manner[2]. Especially, the goal of the PLS is to enable widespread product line system through architecture-based development. PLS focuses on engineering and reengineering software systems from a product line perspective. Software architecture is concerned with capturing the structures of a system and the relationships among the elements both within and between structures. The structures we found fell into severer categories[3].

- The **conceptual architecture** describes the system in terms of its major design elements and the relationships among them.
- The **module interconnection architecture** encompasses two orthogonal structures.
- The **execution architecture** describes the dynamic structure of a system.
- The **code architecture** describes how the source code, binaries, and libraries are organized in the development environment.

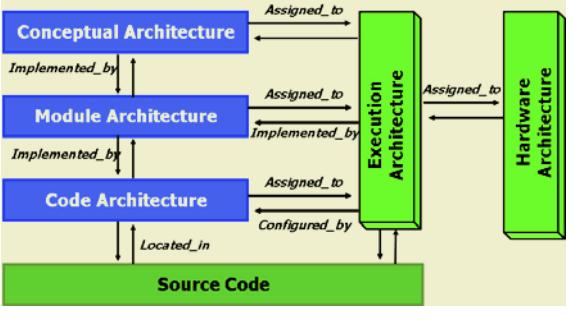


Fig. 1. Relationship among the Software Architecture

These categories address different engineering concerns. The separation of such concerns, combined with specialized implementation techniques, decreased the complexity of implementation, and improved reuse and reconfiguration. These relationships are summarized in Figure 1.

2.2 4+1 View Architecture Model

Software architecture deals with the design and implementation of the high-level structure of the software. It is the result of assembling a certain number of architectural elements in some well-chosen forms to satisfy the major functionality and performance requirements of the system, as well as some other, non-functional requirements such as reliability, scalability, portability, and availability. Perry and Wolfe put it very nicely in this formula2, modified by Boehm: Software architecture = {Elements, Forms, Rationale/Constraints}

Also, Software architecture deals with abstraction, with decomposition and composition, with style and esthetics. To describe software architecture, we use a model composed of multiple *views* or perspectives. In order to eventually address large and challenging architectures, the model we propose is made up of five main views (cf. fig. 2)[4]:

- The *logical* view, which is the object model of the design (when an object-oriented design method is used),
- The *process* view, which captures the concurrency and synchronization aspects of the design,
- The *physical* view, which describes the mapping(s) of the software onto the hardware and reflects its distributed aspect,
- The *development* view, which describes the static organization of the software in its development environment.

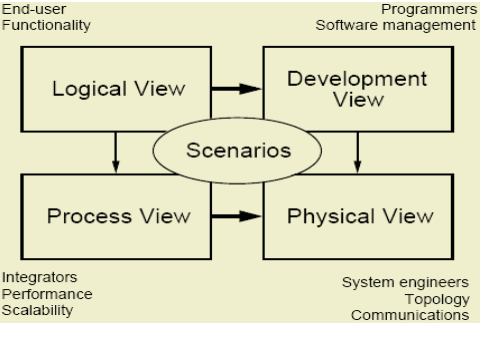


Fig. 2. 4+1 View Model

The description of an architecture—the decisions made—can be organized around these four views, and then illustrated by a few selected *use cases*, or *scenarios,* which become a fifth view. The architecture is in fact partially evolved from these scenarios as we will see later.

3 Component Based Architecture for Workflow Rule and Tool

3.1 Use Case View

Use Case View uses to reflect user's requirement in architecture. The workflow rule engine and tool design and execute the scheduling rules or dispatching rules that need in a field. The workflow rule engine and tool design and run scheduling rules or dispatching rules that need in field. Specially, there are very important role for productivity elevation in product line as support system of MES(Manufacturing Execution System). Figure 3 shows the use case diagram for the workflow rule design and execution in rule designer side.

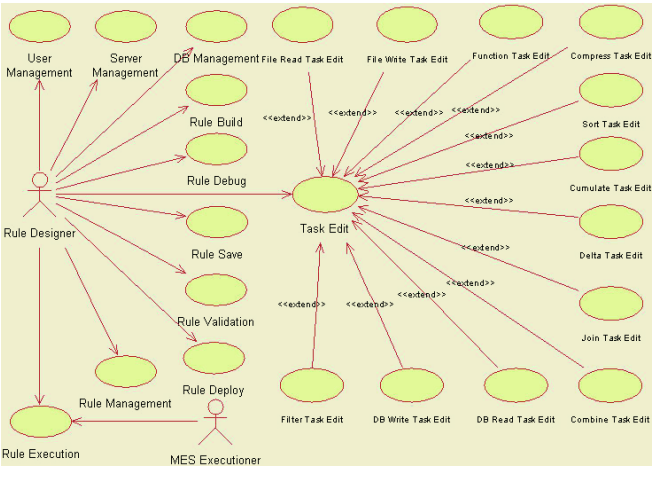


Fig. 3. Use case diagram for Workflow Rule Engine and Tool

3.2 Logical View

We propose architecture that consists of 5 layers with figure 4. These contents are as following.

Client layer takes charge the interaction of user and system. Client layer receives user's request and delivers to the system. As the result, client responses and handles in system to user.

- User Interface: Take charge the interaction between user and system.
- Client-side Control: Offer various functions that can not offer with User Interface. There is program or library in User Interface outside as ActiveX, Applet, Plug-in.

Presentation layer connects with client layer and business layer. It verifies user request in client layer changes to correct form and delivers from business layer. Also, there communicate response that handle in Business layer to client layer.

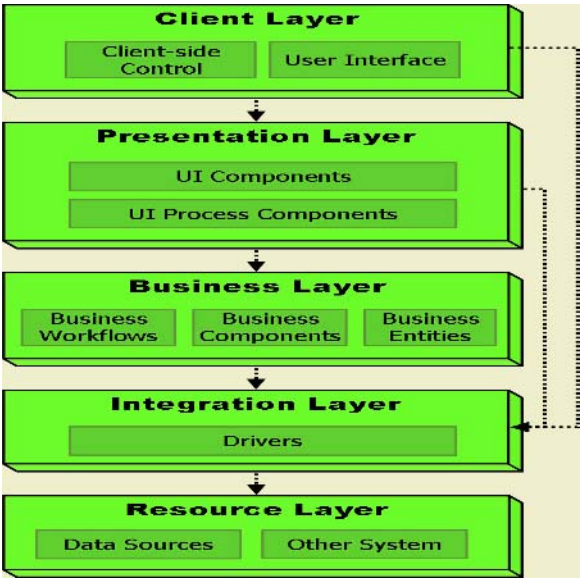


Fig. 4. Basic Architecture for Workflow Rule Engine and Tool

• UI Components: User offers so that it does interact with application. This can be implemented by taking advantage of .Net technologies of window form, web form etc.
• UI Process Components: User process component is used to control processing procedure that is accompanied with user interaction as the same time. It may not do coding directly to user interface item as using component that flows of process and state administration logic. It can reuse equally user interaction engine in several user interfaces.

Business: Business layer handles business according to business rule. Business layer verifies data in presentation layer. It makes data as new result by applying suitable business rule data and existing data, storing or delivering to outside system.

• Business Workflows: We collect data to need Business Workflows through a : user process. We can use this data at a business process processing. A business workflow defines a business process of many steps and accomplishes the facility to combine. We use a business process officer tool and can implement.
• Business Components: The business component implements business logic. It operates a business task which the application needs.
• Business Entities: We deliver data between components and express a business entity. We design the entities to the foundation and implementation classes to be deduced in an object modeling course.

Integration layer connects with integration a system and various outside resource. Integration layer can be connect to database of resources layer as well as that connects

with system that know physically with outside resources and do abstract to work that happen in interaction such as data transmission and conversion .

• Drivers: We connect a system to outside system and convert data and deliver.

A resource layer is an outside resource which the system uses. Outside resources defines by everything (system, library, module, control and so on) in system outside.

• Data Sources: It means data pool of various forms including database system.
• Other System: There are outside system which connects with the system. There is abstract legacy system service or web service etc. as services that exist on system outside.

The layers has the dependency with a downside layer but there is following exception according to each situation. Client layer knows that if must have dependence, but connect mainly to outside system such as database system or legacy system being lower part presentation mainly be integration layer and have relativity. Presentation

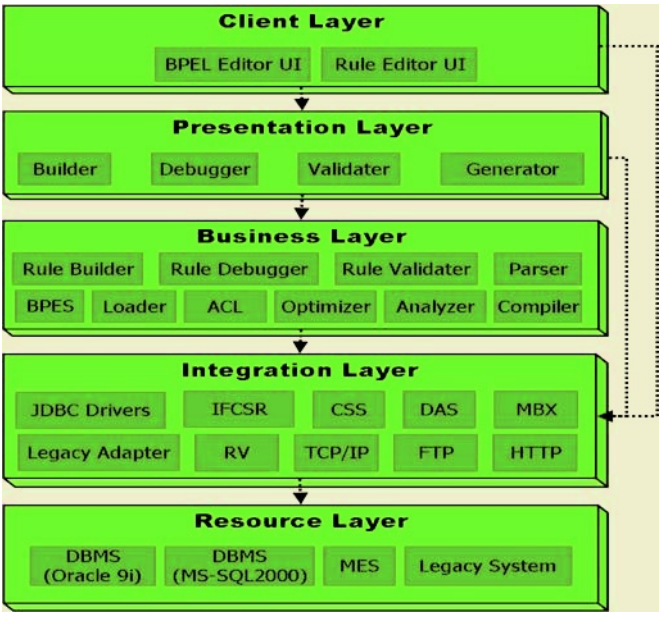


Fig. 5. Physical item of workflow rule engine and tool

layer be know that if must have relativity, but connect justly to outside system such as database system or legacy system being lower part business layer mainly justly be Integration layer and have relativity. Existing business workflow is implemented in program logic inside. The flexible about the change of the system gets low consequently. We propose the system of a framework foundation to solve such a problem. Also, we implement a business rule at the server with a figure 5.We present the architected to execute by the workflow which the user defines such rules.

3.3 Component View

We analyze a component inside action of the system which tries to develop in a component view. We present a component architected to the foundation the content. Figure 6 stores the workflow rule to database and shows the procedure to build in a rule

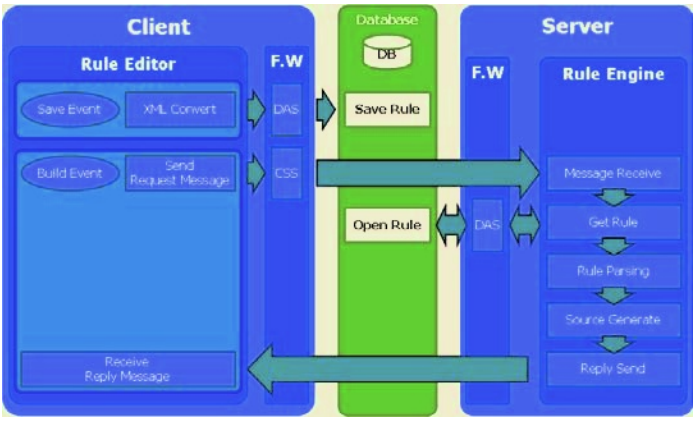


Fig. 6. Rule Save and build operation

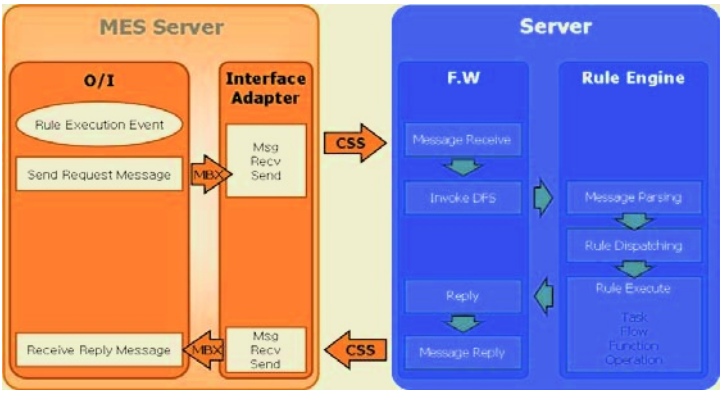


Fig. 7. Execute operation

engine. We convert to XML format and store the rule at workflow rule database server. They accomplish the build in a workflow rule engine according to the scenario. Figure 7 presents the action to execute an workflow rule. We execute the task to request dispatching in an outside MES in the server.

We propose ABCD architected pattern in this paper with figure 8.

Domain Component: Component in demand to specific business domain.

• The structure of the business component in demand to application domain practiced particular task and default business logic.
• Application module applied straightly.
• Basis building-block for domain application.

Core Business Component: The definition of the functional and technical component among the general components.

• Core business logic support what user wants.
• The detailed definition of the component is to serve the general service in the business domain.

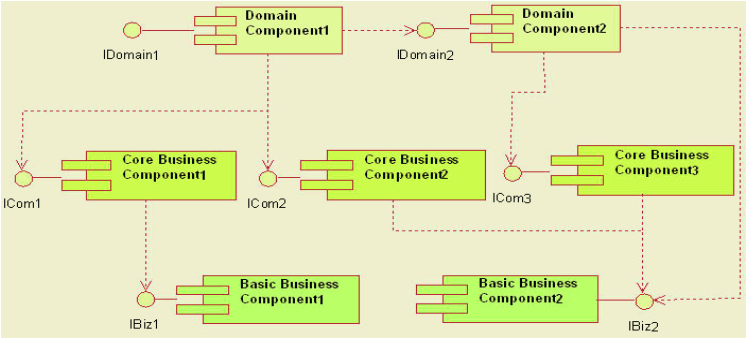


Fig. 8. Component based architecture-ABCD pattern

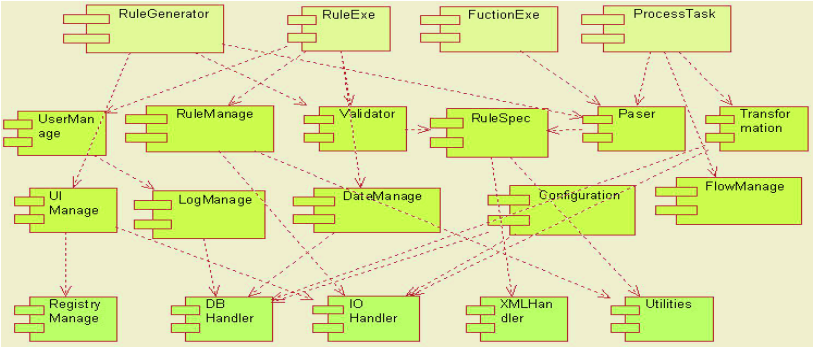


Fig. 9. Component based architecture of workflow rule engine and tool

Basic Business Component: The definition achieves an important role in the whole business.

• Functionality component for the common business achievement.
• Appearance of the basis solution set for business operation
• Infrastructure and service demanded to develop the commercial application dealt with distributed object technology.

Architecture Component(Platform) : Component depended on platform.

• Appling various platforms for running and constructing component.

Figure 9 presents a component based architected for workflow engine and tool. An architecture component did not consider in this paper.

3.4 Implementation View

A Client layer and presentation layer implement with the in house library using Microsoft's .Net framework. A business layer is developed in Java Runtime Environment. An Integration layer implements with the driver to offer from an outside system. A resource layer is MES, DBMS systems. Figure 10 presents implementation

Fig. 10. Context Diagram of Workflow Rule and Tool

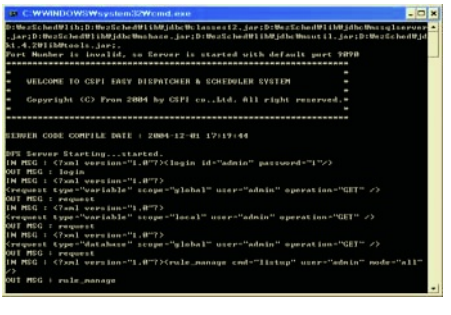


Fig. 11. Workflow Rule Engine

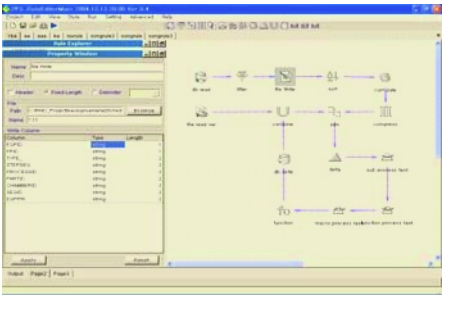


Fig. 12. Workflow Rule Tool

view which expresses in a context diagram. There are DAS(Data Access Server) and CSS(Communication Standard Service) for data access and communication between client and server. A workflow rule engine is located in the server. Figure 11 and 12 present that workflow engine executed and user interface designed for workflow rule.

4 Conclusion and Future Works

In this paper, we propose the architecture of 4 views and apply views to workflow engine and tool. The views are use case view, logical view, component view and implementation view. Especially, we discussed on logical viewpoint and component viewpoint for development of component. The logical view of Workflow domain defined 5 hierarchies laying stress on function. The hierarchies are client layer, presentation layer, business layer, integration layer and resource layer. A component architecture defined use ABCD pattern proposed in this paper. The ABCD pattern organized architecture component, basic business component, core business component, and domain component. The component based architecture can be specified definitely by applying ABCD pattern in each hierarchy of logical architecture.

As a result, implementation architecture presented whole system structure and user interface that is developed. The component base architecture that applies workflow engine and tool, presents various viewpoint of development. This offers an opportunity that can implement easily to developer, and act role as guide. Also, we expect to decrease the complexity of implementation, improved reuse and reconfiguration and productivity in work-flow system and similar domain. In the future work, there area the great needs more study about component integration based architecture for workflow system. We are also going to study on the various case study and evaluation for architecture efficiency.

References

1. Frank Svoboda and Dick Creps, "Software Architecture Technology Guide"
2. Paul Kogut, "The Software Architecture Renaissance", STSC's Crosstalk newsletter, 1994.
3. Dilip Soni, Robert L. Nord, and Christine Hofmeister, "Software Architecture in Industrial Application", Proceedings of the 17th international conference on Software engineering, pp. 196-206, 1995.
4. Philippe Kruchten, "Architectural Blueprints-The "4+1" view Model of Software Architecture", IEEE Software, Volume 12, Issue 6, pp.42-50, 1995.
5. George T. Heineman and William T. Councill, “Component-Based Software Engineering”, Addison-Wesley, 2001.
6. Claudio Riva, Petri Selonen, Tarja Systä and Jianli Xu, "UML-based Reverse Engineering and Model Analysis Approaches for Software Architecture Maintenance", Proceedings of the 20th IEEE International Conference on Software Maintenance, pp.50-59, 2004.
7. Hugo Vidal Teixeira, Regina M. Braga and Cláudia M. L. Werner, "Model-Based Generation of Business Component Architectures", Proceedings of 30th EUROMICRO Conference, pp. 176-183, 2004.
8. Guoqing Xu, Zongyuan Yang and Haitao Huang, "Article abstracts with full text online: A basic model for components implementation of software architecture", ACM SIGSOFT Software Engineering Notes, Volume 29, pp. 1- 11, Issue 5, 2004.
9. Pearl Brereton and David Budgen, "Component-Based Systems: A Classification of Issues", IEEE Computer, Vol 33, No 11, pp. 54-62, 2000.
10. Jia Yu and Rajkumar Buyya, "A Novel Architecture for Realizing Grid Workflow using Tuple Spaces", Proceedings of Fifth IEEE/ACM International Workshop on Grid Computing, pp. 119-128, 2004.

11. Lixin Wang and Loo Hay Lee, "Semiconductor manufacturing: Scheduling and dispatching: a simulation study on release, synchronization, and dispatching in MEMS fabrication", Proceedings of the 34th conference on Winter simulation: exploring new frontiers, pp. 1392 - 1400, 2002.
12. H.K. Kim, E.J.Han, H.J. Shin and C.H. Kim, "Component Classification for CBD Repository Construction", Proceedings of SNPD'00, pp. 483-493, 2000.

A Fragment-Driven Process Modeling Methodology

Kwang-Hoon Kim[1], Jae-Kang Won[1], and Chang-Min Kim[2]

[1] Collaboration Technology Research Lab.,
Department of Computer Science, Kyonggi University,
San 94-6 Yiuidong Youngtonggu Suwonsi Kyonggido, 442-760, South Korea
{jkwon, kwang}@kyonggi.ac.kr
[2] Division of Computer Science, Sungkyul University,
147-2 Anyang8dong, Manangu, Anyangsi, Kyonggido, 430-742, South Korea
kimcm@sungkyul.ac.kr

Abstract. In this paper[1], we propose an advanced modeling approach, which is called a fragment-driven process modeling methodology that enables several real actors/workers to cooperatively define a process model. We also suggest a feasible design that can cooperatively realize the methodology with maximizing efficiency and involvement of real workers in the process modeling work. In the traditional approaches, the modeling work is done by a single designer who has to know all about the detailed and complex information needed for a process model, such as relevant data, organizational data, roles, activities, related software and programs, scripts, etc. However, according to that processes have become more complicated and large-scaled in recent, it is hard to say that the approach is reasonable. Therefore, we propose a more realistic approach that enables several real actors/workers to cooperatively define a process model with disclosing just required information through completely distributed environment. In this approach, the actors need to define only their own activities, not the whole one, then the system gathers these partial sub-models (which is called process fragments), and finally compose the complete process model. We strongly believe that the methodology should be very applicable and valuable for cooperatively modeling not only intra-organizational processes but also cross-organizational e-business processes, such as SCM, e-Commerce, e-Logistics, and so on.

Keywords: Fragment-driven Process Modeling Methodology, Cooperative Modeling Approach, Information Control Net, Global Process Model.

1 Introduction

There exist a lot of technologies that make the world closer together physically and conceptually. We would categorize those technologies into two groups

[1] This work was supported by Korea Research Foundation Grant. (KRF-2002-003-D00247).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 817–826, 2005.

according to the subjects of the world that they are aiming at - people and organization. That is, one is the group of technologies targeting on bringing people closer together, the other is the group of technologies making organizations much closer together. Conceptually speaking, the former group is focusing on enhancing efficiency and productivity of intra-organizational group activities. And, the latter group is to promote inter-organizational collaborative activities. As a matter of fact, we would proclaim that the focus of information technological trend be shifted from supporting collaborative work for group of people to supporting collaborative work for group of organizations.

In the business process management literature, these technological trends are coming true, too. We can see two evidences - Merging realtime groupware into process issue and Cross-organizational process issue. The first issue is related with the complexity problem of process. A process has been becoming gradually complex more and more in terms of the structural aspect of the process as well as the behavioral aspect of the process. The structural complexity is concerning about how to efficiently model and define a large-scale process procedure consisting of a massively parallel and large number of activities. Therefore, by using the realtime groupware technology, a group of people (process designers) is able to cooperatively model and define the large-scale process procedure. The behavioral complexity is related with the run-time components of process enactment. If a certain activity's implementation (application program, for example) has to be collaboratively done by a group of actors, then in order to support the situation, the realtime groupware technology should be merged into the process enactment components. In this paper, we are looking for a method for merging groupware into the process modeling work.

Conclusively, in this paper, we would like to seek a feasible solution for resolving the previous two methods at once. It is called cooperative fragment-driven process methodology that is used for constructing a cross-organizational process procedure in a way of not only that a group of actors can be engaged in the modeling work at anywhere and anytime, and but also that it avoids the dilemma. We concretize the methodology by designing a cooperative process modeling concept with respect to ICN (Information Control Net) and by embedding the realtime groupware functionality, as well. Therefore, a group of designers or actors are able to open a session, join to the session, and cooperatively construct a cross-organizational process procedure by the methodology and the system. We describe the methodology in the section 3 and 4, after introducing related works and backgrounds in the next section. Finally, we explain about the use and extension of the methodology.

2 Backgrounds and Related Works

In the world-wide information technology market including South Korea's, process and its related technological fields, such as WS, SCM, CALM, EAI, B2B e-Commerce, etc., begin attracting great attention from the society of information science and database management in aspects of not only research topics but

also industrial one such as information application fields. It does start catching public attentions. There are several projects ongoing research and development of process systems issued by universities and research institutes. Of course, these technologies are issued and fairly settled down in the worldwide information technology arena, too.

In terms of the usability of the process and e-commerce technologies, we can see the complete disproportion, which means that the number of people and organizations doing anything more than talking about them seems to be small, and that they have not been widely installed and practiced, yet. That is precisely why this methodology is needed and why we have received so much encouragement to write them from practitioners in the field. At the same time, that's why many professionals from both business and IT disciplines ought to be engaged in the design and implementation of the inter-organizational process and e-commerce systems.

In terms of, however, the cross-organizational process modeling aspect, there are several related works in the literature. According to the degree of complexity in collaboration among organizations (or inter-organizations), there might be two types of collaborative organizations - loosely coupled collaborative organizations and tightly coupled collaborative organizations. The Inter-process project [4] conducted at the Kanagawa Institute of Technology, Japan, is one of typical frontiers pioneering inter-organizational process modeling methodology and system for the loosely coupled collaborative organizations. They focus on the definition of a global-process model (which they call inter-process) for an cross-organizational business process. This inter-process model defines the basic interaction between the associated parties, and then it is transferred into the process management systems of the parties. Within the system, the (local) processes are modified to be suitable for the needs of the individual enterprizes. On this basis, the inter-process definition tool is used to define the inter-process process. After this definition work, the translators automatically convert the inter-process process definition data into the process engines used in each organization. While on the other, as a methodology supporting the tightly coupled collaborative organizations, we propose the fragment-driven process modeling methodology and design a cooperative process modeling system, which is based upon the methodology, in this paper. The detailed idea and concept of the methodology is described in the next section. Also we explain what are differences between the [4]'s approach and ours based upon the issues, too.

3 Fragment-Driven Process Modeling Methodology

As stated in the previous sections, we would propose a methodology for modeling a process procedure in a completely different way from the traditional modeling approaches. The methodology reflects the basic philosophies of the previous two groups of technologies as much as possible. That is, we look for a way not only that makes people closer and more cooperative but also that makes inter-related organizations more collaborative in modeling a global process. According for the

process to be more complicated and to be engaged with many organizations, our methodology will be more worthy and more effective. It comes from a simple idea, in which it should be not reasonable for only a single designer to be able to define a process. Then, how can we make as many designers as possible to be engaged in modeling a process? In this section, we conceptually and operationally illustrate the idea of the fragment-driven approach, as an answer for the question, and formally define the approach by information control net.

3.1 The Conceptual Idea of the Fragment-Driven Approach

The fragment-driven process model and methodology are focusing on the definition of a collaborative business process across industries and companies particularly targeting on the process-driven e-Commerce domains. And, in terms of the cross-organizational (Note that we would use the term, global, instead of, cross-organizational, from now, for the sake of simplicity.) process model for defining the collaborative business process, it is very important for the dilemma described in [4] to be resolved. That is, the more independence of organizations is guaranteed, the more each organization has the ability or the right to decide things on its own.

However, in contrast to this, for the construction of a global process, the more interactions among the process fragments associated with the collaborative business process are precisely specified, the more the methodology for describing the global process in one place has advantages. In order to avoid the dilemma, we take a different approach from the [4]'s approach, in which they proposed the hierarchical process description methodology that can be classified into the top-down approach. As a conceptually opposite approach to the top-down approach, we propose the fragment-driven approach, as presented in the Figure 1.

In the top-down approach, a single global process designer defines a global process procedure and disseminates the global process modeling information

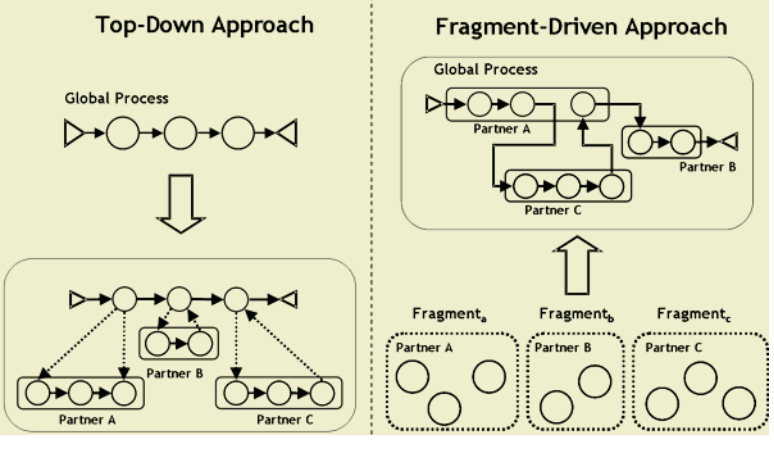


Fig. 1. Top-down Approach vs. Fragment-driven Approach

into the partners, then each partner's modeler should refine its own detailed internal sub-process model. So, the approach requires a set of translators because each partner has a different modeling tool each other. This is why it needs the hierarchical process description methodology, in which the high layer (global process) is done by a global process modeling tool, and the low layer (internal process) is done by a group of internal process modeling tools. Each partner does not need to open its internal things to the others. Conclusively, this approach is able to accomplish a secured modeling work, but it's too complicated and strict to realize the efficiency and effectiveness of the global process modeling work, itself.

In the fragment-driven approach, we conceive a logically (not physically) hierarchical process description methodology, in which all process fragments are modeled by a single global process modeling tool, but they share a global working space under the control of strict access rules. And the global process modeling tool supports the realtime groupware functionality so as for a group of designers to cooperatively model a global process procedure at the same time and from the different locations. Each organization needs to open not its whole internal things but only its partial things (that are just related with only its process fragment) to the others. Also, we do not need the translators, either. The global process model collaboratively defined by a group of cooperative designers is going to be stored on the registry/repository that can be accessible from the member organizations. Conclusively, this approach is able to accomplish not only a secured global process modeling work but also a fragment-driven/cooperative global process modeling work. Furthermore, we can expect the efficiency and effectiveness on constructing a global process model.

3.2 The Collaborative Global Process Model

The collaborative global processes graphically represented in ICN are just the e-Business processes that we are eventually targeting on for the cooperative fragment-driven process modeling system. Figure 2 is to represent an e-Logistics global process by the ICN notation, as a typical example of the process-driven e-Commerce process in real environment. The collaborative global process's fragments are scattered over five organizations (Org_a, $\sim$ Org_e). That is, the organizations cooperate each other in the fashion of tightly coupled relationship. Assume that, in this model, five process engines, which are distributed over the five organizations respectively, coordinate to perform the collaborative global process. We can imagine that the e-Logistics global process of a logistic company cooperates with three foreign service partners (Org_a, Org_b, and Org_c) and two other logistics companies (Org_d, and Org_e), in fashions of nested and synchronized pattern. This paper is targeting on how to support for a group of people to describe and specify just like the complicated e-Logistic processes by the cooperative fragment-driven process modeling system.

We would shortly introduce ICN [6] and extend it to accommodate the collaborative global process models. The Information Control Net [6] was developed for describing and analyzing information flow within an office. It has been used within actual as well as hypothetical automated offices to give a comprehensive

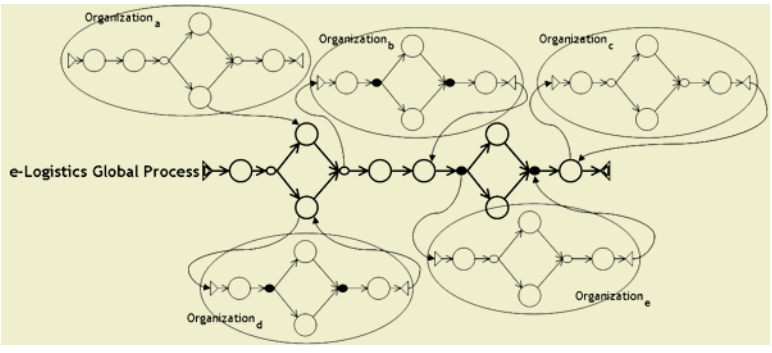


Fig. 2. A Complicated Global ICN Model with e-Logistics Process Fragments in Collaboration

description of activities, to test the underlying office description for certain flaws and inconsistencies, to quantify certain aspects of office information flow, and to suggest possible office restructuring permutations. In order to support global process modeling with multiple organizations, we extend the basic ICN model by incorporating the notion of organization including global-actors, global-groups, global-organizations. The following is the formal definition of the extended ICN for global processes:

A global ICN is 6-tuple $\Gamma = (\delta, \gamma, \pi, \kappa, I, O)$ over a set of A activities (including a set of activities), a set of R of repositories and a set C of organizations, where

- I is a finite set of initial input repositories, assumed to be loaded with information by some external process before execution of the ICN;
- O is a finite set of final output repositories, perhaps containing information used by some external process after execution of the ICN;
- $\delta = \delta_i \cup \delta_o$
  where, $\delta_o : A \longrightarrow \wp(A)$ is a multi-valued mapping of an activity to its sets of (immediate) successors, and $\delta_i : A \longrightarrow \wp(A)$ is a multi-valued mapping of an activity to its sets of (immediate) predecessors; (For any given set S, $\wp(S)$ denotes the power set of S.)
- $\gamma = \gamma_i \cup \gamma_o$
  where $\gamma_o : A \longrightarrow \wp(R)$ is a multi-valued mapping (function) of an activity to its set of output repositories, and $\gamma_i : A \longrightarrow \wp(R)$ is a multi-valued mapping (function) of an activity to its set of input repositories;
- $\pi = \pi_a \cup \pi_c$
  where, $\pi_a : C \longrightarrow \wp(A)$ is a single-valued mapping of an activity to its organization, and $\pi_c : A \longrightarrow \wp(C)$ is a multi-valued mapping of an organization to its sets of associated activities;
- $\kappa = \kappa_i \cup \kappa_o$
  where κ_i : sets of control-transition conditions, T, on each arc, $(\delta_i(\alpha), \alpha), \alpha \in A$; and κ_o : sets of control-transition conditions, T, on each arc, $(\alpha, \delta_o(\alpha)), \alpha \in A$; where the set $T = \{default, or(conditions), and(conditions)\}$.

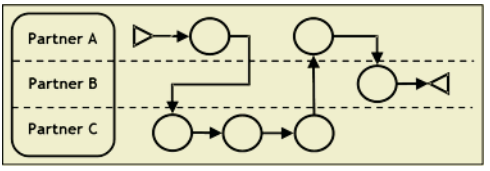


Fig. 3. The Conceptual User Interface (Swimlane and ICN) of the Modeler

In mapping ICN diagrams into formal definitions, solid arrows into an activity node correspond to the function δ_i, and solid arrows out of a node correspond to δ_o. Similarly, dashed arrows into an activity node correspond to the function γ_i, and dashed arrows out of a node correspond to γ_o. Additionally, solid lines between an activity and an organization correspond to the function π_a, and solid lines between an organization and activities correspond to π_c. Finally, the control-transition conditions, which consist of three kinds, *default*, *or*(*conditions*), and *and*(*conditions*), on the arrows into an activity correspond to the κ_i function, and the control-transition conditions on the arrows out of an activity correspond to κ_o.

3.3 The Conceptual Idea of Fragment-Driven Modeling Methodology

The modeler to be implemented based on the methodology provides a set of functions and capabilities for the process modeling work cooperatively performed by a group of users (or designers/builders). As a front-end component of the system, it has three major components - graphical user interface manager, cooperative fragment client engine, and local modeling data manager - that provide the group workflow modeling functionalities, such as user management, session control, event handling and synchronization, floor control, space sharing, access control, message handling, modeling data management, model verification, and so on.

The graphical user interface manager enables a group of cooperative designers to define, edit, verify, and simulate a collaborative global process on a shared space as shown in Figure 3. The shared space is divided into the number of designers logged on the current session, each of whom is assigned into one swimlane (three swimlanes in Figure 3). Then, the owner of each swimlane takes the responsibility for editing his/her workflow fragment. Of course, each of designers can't edit the others' workflow fragments, but they can only see and read the others' modeling works. As a result, the swimlane means, in this modeler, the access control boundary (read/write permission for the owner, read-only for the others).

3.4 The Operational Idea of the Fragment-Driven Modeling Methodology

From the point of operational view, we would name the [4]'s approach "black box approach," in which the internal process in each organization is treated as

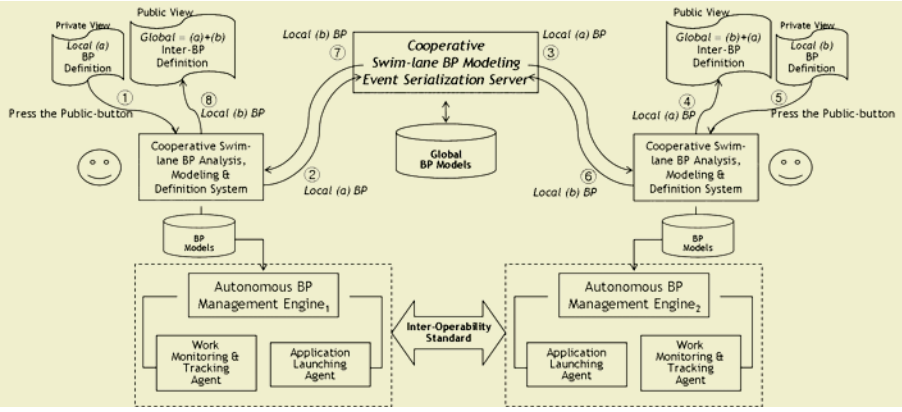


Fig. 4. The Operational Scenario of the Fragment-driven Approach

a black box, because it makes only the linkage with others visible. While on the other, our approach would be named "white box approach," because all of the internal processes (process fragments) associated with a global process procedure can be precisely described in one place where the modeler is located, and broadcasted into the other coordinators. As we described in the previous, our methodology is pursuing the closed e-business framework, not the open e-business framework. That is, while the black box approach might provide a reasonable solution for constructing the open e-business framework, the white box approach, our approach, should be appropriate for constructing the closed e-business framework. In this section, we explain how the white box approach, our modeling methodology, defines a global process and generates its collaborative process fragments, each of which belongs to an organization.

Based upon the concept of the white box approach and the conceptual idea described in the previous section, we imagine how the fragment-driven approach is basically manipulating the events through an operational scenario as shown in Figure 4. In order for the approach to be operable, the following architectural components will be implemented in the next section. The numbers (1 8) on the figure represents the sequence of events that are broadcasted to the other coordinator in realtime during the fragment-driven process modeling work. Without any further explanation, you may be easily understood the meanings of the events and the numbers.

- Cooperative process modeling and event serialization server
- Global Registry and Repository for storing the defined global process models
- Cooperative process analysis, modeling and definition system
- Local/internal repository for storing the process fragments
- Private View and Public View on the windows of a local system.

4 Implications

In summary, we have introduced our fragment-driven process modeling methodology, so far. The methodology might fit very well into the system that provides tightly coupled interactions among a group of people as well as a group of organizations in collaboration. So, we have a basis for implementing a process modeling system that is operable based upon the methodology. We are doing an on-going development of a cooperative swimlane process modeling system. We have shortly introduced the system's operational idea with a little more detailed description in the previous section, in order to show evidence that how feasible the methodology is. The system might be able to avoid the dilemma [4] and to provide a great simplicity and efficiency in process procedure modeling work that should be perfectly applicable for not only intra-organizational processes but also cross-organizational processes. Also, it can be embedded into the closed or process-centric e-business framework. The methodology's advantages can be summarized as the followings:

- It is a white box approach, because each activity can be clearly assigned into the corresponding organization that has the responsibility of execution. So, the construction unit of a global process model becomes the activity, which is exactly same with the conventional process modeling methodology.
- A group of cooperative designers are always aware of collaborative linkages, in real time, existing among internal processes (or process fragments) of a cross-organizational process model, because the cooperative fragment-driven modeling tool provides WYSIWIS functionality to the designers logged on a session, and it is able to automatically generate a cross-organizational process model by combining and arranging the process fragments.
- It needs not to develop a set of translators converting the description language of a global process model into the language of each process fragment model, because the description language of the global process model collaboratively defined by a group of designers is WPDL/XPDL, and the corresponding process fragments are stored in the registry/repository component that provides interfaces to the runtime components of each organization.

5 Conclusion

So far, we have proposed the fragment-driven process modeling methodology and proved that the methodology is a feasible and realistic by designing the cooperative swimlane process modeling system. We strongly believe that the modeling system fortifies the modeling methodology, and vice versa. Also we assure that the modeling methodology fits very well into the tightly-coupled framework for the process-driven e-Commerce domains. Especially, the methodology might have a strong advantage in terms of the high degree of modeling ability for not only complex and maximally parallelized process procedures but also collaborative global process procedures. At the same time, it should be very applicable to

the type of collaborative global process models that is defined and characterized, by us, with tightly coupled global processes in collaborative organizations.

Not only in South Korea but also in world-wide research arena, process and its related technological fields, including process-centric e-Business, software development process and e-Logistics, are catching great attentions from the society of information science and database management fields. So, there are a lot of challenges to develop and commercialize business management solutions. This paper should be one of those active attempts for pioneering global process modeling methodologies toward supporting cross-organizational processes in collaboration, too.

Acknowledgement. This work was supported by Korea Research Foundation Grant (KRF-2002-003-D00247).

References

1. Process Management Coalition Specification Document, "Process Coalition Interface 1: Process Definition Interchange Process Model", Document Number: WFMC TC-1016-P, August 5, 1998.
2. UN/CEFACT and OASIS, "ebXML Documents v1.0", May 2001
3. Edward A. Stohr, J. Leon Zhao, "Process Automation: Overview and Research Issues", Information Systems Frontiers, Volume 3, Issue 3, 2001
4. Haruo Hayami, Masashi Katsumata, Ken-ichi Okada, "Interprocess: A Challenge for Business-to-Business Electronic Commerce", Process Handbook 2001, WfMC, October 2000
5. Dong-Keun Oh, Kwang-Hoon Kim, "An EJB-Based Database Agent for Process Definition", Journal of Korean Society for Internet Information, Vol.4, No.5, December 2001
6. Clarence A. Ellis, "Formal and Informal Models of Office Activity", Proceedings of the 1983 Would Computer Congress, Paris, France, April 1983

A FCA-Based Ontology Construction for the Design of Class Hierarchy

Suk-Hyung Hwang[1], Hong-Gee Kim[2], and Hae-Sool Yang[3]

[1] Division of Computer and Information Science, SunMoon University,
100 Kal-San-Ri, Tang-Jeong-Myon, A-San, Chung-Nam, 336-840 Korea
shwang@sunmoon.ac.kr
[2] College of Dentistry, Seoul National University,
28-22 Yeonkun-Dong, Chongno-Ku, Seoul 110-749, Korea
hgkim@snu.ac.kr
[3] Graduate School of Venture, Hoseo University,
29-1 Se-Chul-Ri, Bae-Bang-Myon, A-San, Chung-Nam, 336-795 Korea
hsyang@office.hoseo.ac.kr

Abstract. One of the most tasks of object-oriented software designers is the design of the class hierarchy and the relationships among classes. Since there are many conceptual similarities with the design of an ontology, and an ontology is semantically richer than a UML class model, it makes sense to put the emphasis on ontology design. That is, an object-oriented software designer can design an ontology by organizing object classes in a class hierarchy and creating relationships among classes. UML models can then be generated from the ontology.

In this paper, we introduce the Formal Concept Analysis(FCA) as the basis for a practical and well founded methodological approach to the construction of ontology. We show a semi-automatic, graphic and interactive tool to support this approach. The purpose of this work is to provide a semi-automatic methods for the ontology developers. We describe here the basic ideas of the work and its current state.

1 Introduction

An ontology defines the terms and concepts(meaning) used to describe and represent an area of knowledge, as well as relations among them[1]. An ontology is based on a taxonomy which represents a class hierarchy in the object-oriented world. Software design artifacts, such as UML class models, and even source code can be generated directly from ontologies, thus effectively speeding up the software development process.

There are many guidelines and methodologies for ontology construction proposed [2, 3, 4, 5, 6]. But none of them can be used generally. The process of ontology build is craft rather than an engineering activity. Manual construction and description of domain-specific ontology is a complex and time-consuming process. It is extremely difficult for human experts to discover ontology from given data or text. As verification of the correctness of ontology has to be done

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 827–835, 2005.

by human experts with respect to their knowledge about domain, it can lead to inconsistency of the created ontology. One solution to the problem is to build ontology automatically or at least, use semi-automatic methods. Formal concept analysis(FCA)[7] can help structure and build ontologies. This is because FCA can express ontology in a lattice that is easy for people to understand and can serve as a guideline for ontology building.

This paper presents a formal approach to construct ontologies using Formal Concept Analysis. That is, we propose here to adopt the Formal Concept Analysis as a framework for the identifying concepts and the construction of concept hierarchies("concept lattices") because it is a natural structure that systematically factors out commonalties while preserving specialization relationships between concepts. The rest of the paper is organized as follows: In a first part, Section 2 presents an introduction to the formal concept analysis: *formal context*, *formal concept*, and *concept lattice* etc. are introduced. Section 3 gives an overview of construction of ontology based on the concept lattice; Lastly, Section 4 concludes the paper, giving the prospects of our ongoing works to study and prototype the semi-automatic tool.

2 Formal Concept Analysis

The basic structure of formal concept analysis is the context. A context is comprised of a set of objects, a set of attributes and a relation describing which objects possess which attributes. In the formal definition, the set of objects is denoted by $\mathcal{O}$, and the set of attributes is denoted by $\mathcal{A}$.

To introduce the method FCA, we first have to define the term *context* or *formal context*[†].

Definition 1. *A formal context is a triple* $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ *that consists of two sets* $\mathcal{O}$ *and* $\mathcal{A}$ *and a relation* $\mathcal{R}$ *between* $\mathcal{O}$ *and* $\mathcal{A}$*. The element of* $\mathcal{O}$ *are called the* **objects** *and elements of* $\mathcal{A}$ *are called the* **attributes** *of the context. In order to express that an object* o *is in a relation with an attribute* a*, we write* $o\mathcal{R}a$ *or* $(o, a) \in \mathcal{R}$ *and read it as "the object* o *has the attribute* a*".*

Contexts can be represented as cross tables, whose rows are headed by the objects, whose columns are headed by the attributes and whose cells are marked iff the incidence relation holds for the corresponding pair of object and attribute.

For instance, consider a context where this context is based on the set of objects $\mathcal{O}$ and the set of their attributes $\mathcal{A}$ as follows: $\mathcal{O} = \{o_1, o_2, o_3, o_4\}, \mathcal{A} = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ and the incidence relation $\mathcal{R}$ is given by the cross table(Table 1). The table should be read in the following way: Each $\times$ marks a pair being an element of the incidence relation $\mathcal{R}$, e.g. (o_1, a_2) is marked because

[†] FCA may be applied to *many-valued context* in which objects are interpreted as having attributes with values. However, in this paper, we are concentrating on the *(single-valued) context.*

Table 1. An example of formal context with four objects and seven attributes

	a_1	a_2	a_3	a_4	a_5	a_6	a_7
o_1		×				×	×
o_2			×	×	×		
o_3			×		×	×	×
o_4	×		×				

the object o_1 has a_2 as its attribute, whereas (o_1, a_3) is not marked because the object o_1 does not have attribute a_3.

The central notion of FCA is the **formal concept**. A concept (O, A) is defined as a pair of objects $O \subseteq \mathcal{O}$ and attributes $A \subseteq \mathcal{A}$ which fulfill certain conditions. To define the necessary and sufficient conditions for a formal context, we define two derivation functions as follows:

Definition 2. *Let* $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ *be a context,* $O \subseteq \mathcal{O}$ *and* $A \subseteq \mathcal{A}$*. The function* intent *maps a set of objects into the set of attributes common to the objects in* O*(*$\mathsf{intent} : 2^O \to 2^A$*), whereas* extent *is the dual for attributes sets(*$\mathsf{extent} : 2^A \to 2^O$*):*

$$\mathsf{intent}(O) \triangleq \{a \in \mathcal{A} | \forall o \in O : o\mathcal{R}a\},$$
$$\mathsf{extent}(A) \triangleq \{o \in \mathcal{O} | \forall a \in A : o\mathcal{R}a\}.$$

For $O \subseteq \mathcal{O}$, $\mathsf{intent}(O)$ is the set of attributes owned by all classes of O. With $O = \{o_1, o_3\}$, $\mathsf{intent}(O)$ is the set of attributes shared by both o_1 and o_3, and more exactly $\{a_6, a_7\}$. Symmetrically, for $A \subseteq \mathcal{A}$, $\mathsf{extent}(A)$ is the set of objects that own a_3 and a_5, i.e., $\mathsf{extent}(A) = \{o_2, o_3\}$. Together, these two functions form a *Galois connection* between the objects and attributes of the context.

Objects from a context share a set of common attributes and vice versa. Concepts are pairs of objects and attributes which are synonymous and thus characterize each other. Concepts can be imagined as maximal rectangles(modulo permutation of rows and columns) in the context table. When looking at the cross table this property can be seen if rectangles totally covered with crosses can be identified, e.g. the four cells associated with o_1, o_3, a_6, a_7 constitute such a rectangle. If we ignore the sequence of the rows and columns we can identify even more concepts.

Definition 3. *Let* $(\mathcal{O}, \mathcal{A}, \mathcal{R})$ *be a context. A* Formal Concept *is a pair* (O, A) *with* $O \subseteq \mathcal{O}$ *is called* extension*,* $A \subseteq \mathcal{A}$ *is called* intension*, and*

$$(A = \mathsf{intent}(O)) \wedge (O = \mathsf{extent}(A)).$$

In other words a concept is a pair consisting of a set of objects and a set of attributes which are mapped into each other by the Galois connection. The set of all concepts of the context $\mathcal{C} = (\mathcal{O}, \mathcal{A}, \mathcal{R})$ is denoted by $B(\mathcal{C})$ or $B(\mathcal{O}, \mathcal{A}, \mathcal{R})$, i.e.,

$$B(\mathcal{C}) = \{(O, A) \in 2^{\mathcal{O}} \times 2^{\mathcal{A}} | \mathsf{intent}(O) = A \wedge \mathsf{extent}(A) = O\}.$$

For example, $B(\mathcal{C})$ contains the following nine Formal Concepts:
$(\{o_1, o_2, o_3, o_4\}, \emptyset), (\{o_2, o_3, o_4\}, \{a_3\}), (\{o_1, o_3\}, \{a_6, a_7\}), (\{o_2, o_3\}, \{a_3, a_5\})$, $(\{o_1\}, \{a_2, a_6, a_7\}), (\{o_3\}, \{a_3, a_5, a_6, a_7\}), (\{o_2\}, \{a_3, a_4, a_5\}), (\{o_4\}, \{a_1, a_3\})$, $(\emptyset, \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\})$.

Concepts are partially ordered by inclusion of extents(and intents) such that a concept's extent includes the extent of all its subconcepts(and its intent includes the intent of all of its superconcepts). That is, the set of formal concepts is organized the partial ordering relation $\leq$ -to be read as "is a subconcept of"- as follows[‡]:

Definition 4. *For a Formal Context $\mathcal{C} = (\mathcal{O}, \mathcal{A}, \mathcal{R})$ and two Concepts $c_1 = (O_1, A_1)$, $c_2 = (O_2, A_2) \in B(\mathcal{C})$ the* **Subconcept-Superconcept** *order relation is given by:*

$$(O_1, A_1) \leq (O_2, A_2) \Leftrightarrow O_1 \subseteq O_2 (\Leftrightarrow A_1 \supseteq A_2).$$

This relationship shows that the dualism exists between attributes and objects of concepts. A concept $c_1 = (O_1, A_1)$ is a subconcept of concept $c_2 = (O_2, A_2)$ iff the set of its objects is a subset of the objects of c_2. Or an equivalent expression is iff the set of its attributes is a superset of the attributes of c_2. That is, a subconcept contains fewer objects and more attributes than its superconcept.

The subconcept-superconcept relation is transitive, which means that a concept is subconcept of any concept which can be reached by traveling upwards from it. If a formal concept has a formal attribute then its attributes are inherited by all its subconcepts. This corresponds to the notion of "inheritance" used in the class hierarchies of object-oriented modelling. For this reason, FCA is suitable for modelling and analyzing object-oriented class hierarchies.

These definitions now lead us to the following basic theorem of formal concept analysis which says that the set of concepts and the subconcept/superconcept relation defined above forms a complete lattice.

Theorem 1. *([7]) Let $\mathcal{C} = (\mathcal{O}, \mathcal{A}, \mathcal{R})$ be a context. Then $(B(\mathcal{C}), \leq)$ is a complete lattice, the concept lattice of $\mathcal{C}$. Its infimum and supremum(also known as the meet and join, respectively) operation(for any set $I \subset B(\mathcal{C})$ of concepts) are given by*

$$\bigwedge_{i \in I} (O_i, A_i) = \left(\bigcap_{i \in I} O_i,\ \mathsf{intent}(\mathsf{extent}(\bigcup_{i \in I} A_i)) \right),$$

$$\bigvee_{i \in I} (O_i, A_i) = \left(\mathsf{extent}(\mathsf{intent}(\bigcup_{i \in I} O_i)),\ \bigcap_{i \in I} A_i \right).$$

This theorem says that the set of concepts produced by a context is a lattice and the meet and join of a group of concepts can be calculated by the above

[‡] The subconcept-superconcept relation "$\leq$" on the set $B(\mathcal{C})$ of all concepts of a context $\mathcal{C} = (\mathcal{O}, \mathcal{A}, \mathcal{R})$ satisfies *reflexivity*, *transitivity*, and *antisymmetry*.

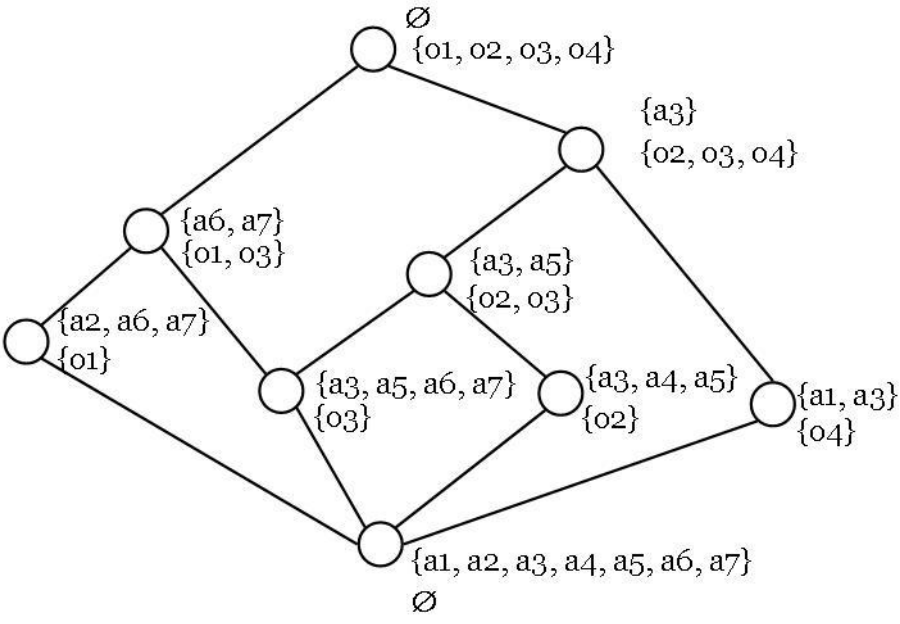


Fig. 1. Concept lattice for the context of Table 1

formula involving the intersection of attributes and objects. As result, we have a lattice denoted by:

$$\mathcal{L} := (B(\mathcal{C}), \leq, \wedge, \vee, (\mathsf{extent}(\mathcal{A}), \mathcal{A}), (\mathcal{O}, \mathsf{intent}(\mathcal{O}))).$$

The lattice $\mathcal{L}$ is called *Galois lattice* or *Formal concept lattice*.

3 Concept Lattice Construction

A concept lattice (a concept hierarchy) can be represented graphically using line (or Hasse) diagrams. These structures are composed of nodes and links. Each node represents a concept with its associated intensional description. The links connecting nodes represent the subconcept/superconcept relation among them. This relation indicates that the parent's extension is a superset of each child's intension. A node covers all of the instances covered by the union of its descendants, making the concept hierarchy a subgraph of the partial ordering by generality. More abstract or general nodes occur higher in the hierarchy, whereas more specific ones occur at lower levels. Figure 1 shows the line diagram for the context of Table 1.

The line diagram consists of nodes that represent the concepts and edges connections these nodes. Two nodes c_1 and c_2 are connected iff $c_1 \leq c_2$ and there is no concept c_3 with $c_1 \leq c_3 \leq c_2$. Attributes and objects propagate

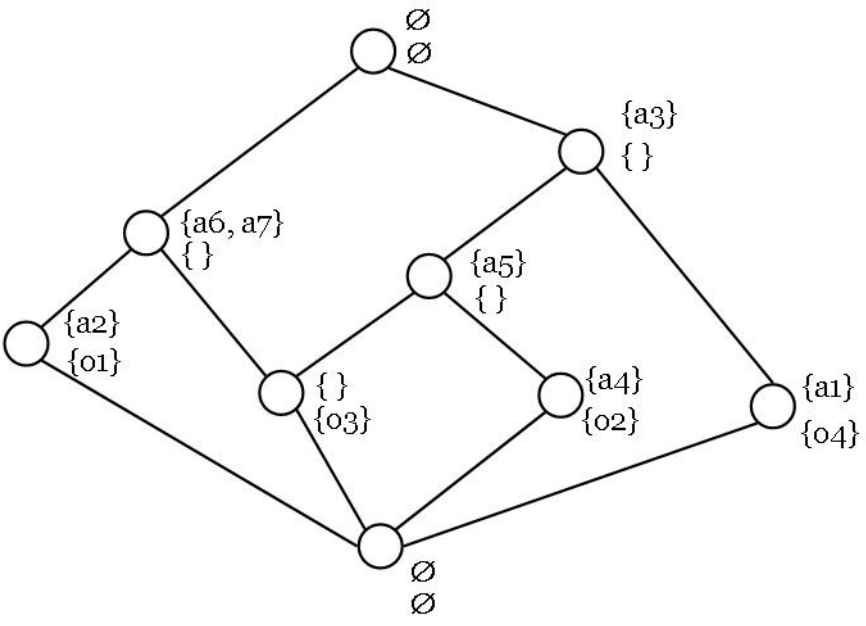


Fig. 2. Inheritance Concept Lattice for Figure 1

along the edges, as a kind of inheritance. Attributes propagate along the edges to the bottom of the diagram and dually objects propagate to the top of the diagram. Thus the top element of a line diagram(the supremum of the context) is actually marked by $(\mathcal{O}, \emptyset)$ if $\mathcal{O}$ is the set of objects. The bottom element(the context's infimum) is marked by $(\emptyset, \mathcal{A})$ if $\mathcal{A}$ is the set of attributes. Now, we can summarize the above considerations as an algorithm to construct concept lattice in **Algorithm 1:** BuildConceptLattice.

Algorithm 1 BuildConceptLattice

```
edges ← ∅;
S ← B(O, A, R);
for all c1 ∈ S do
  for all c2 ∈ S − {c1} do
    if (c1 ≤ c2) ∧ (∄c3 ∈ S − {c1, c2}[(c1 ≤ c3) ∧ (c3 ≤ c2)]) then
      edges ← edges ∪ {(c1, c2)};
    end if
  end for
end for
```

For a pair $c = (A, B)$, the elements of A will be present in every ancestor of c and symmetrically, the elements of B will appear in every descendant.

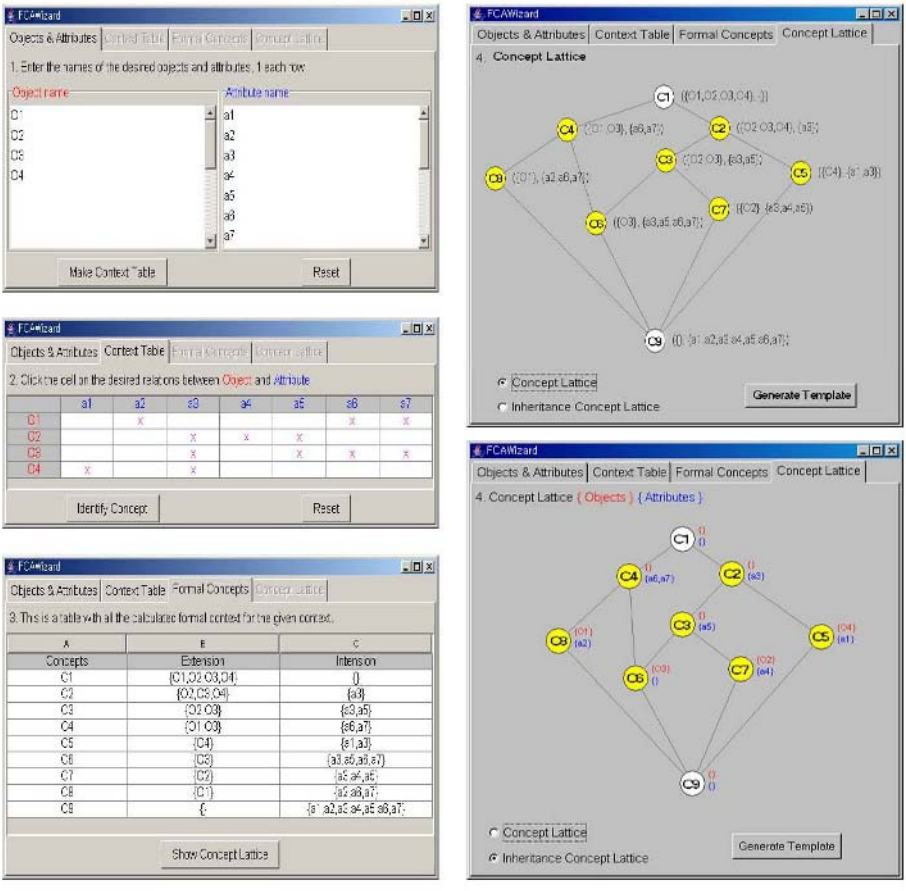


Fig. 3. Screenshots of FCAwizard

Now, let $N(A)$ be the non-redundant elements in A, and $N(B)$ the non-redundant elements in B. An object g will appear in $N(A)$ if the corresponding concept c is the greatest lower bound of all concepts containing g. Symmetrically, a property m will appear in $N(B)$ if the corresponding concept c is the least upper bound of all concepts containing m. The lattice property guarantees the existence of the greatest lower bound and the least upper bound.

Definition 5. *A inheritance concept lattice* ICL *is isomorphic to a concept lattice* $\mathcal{L}$*. For a concept* $c = (X, Y) \in \mathcal{L}$*, let* $N(X)$ *be the non-redundant elements in* X*, and* $N(Y)$ *the non-redundant elements in* Y*:*

$$N(X) = \{x \in \mathcal{O} | \mathsf{intent}(\{x\}) = Y\}$$
$$N(Y) = \{y \in \mathcal{A} | \mathsf{extent}(\{y\}) = X\}$$

A inheritance concept lattice is defined as the set of concepts $(N(X), N(Y))$.

The inheritance concept lattice(ICL) is obtained from the concept lattice $\mathcal{L}$ by replacing each concept (A, B) with the pair $(N(A), N(B))$. We call this lattice an *inheritance concept lattice*.

Figure 2 is the inheritance concept lattice for the example of figure 1. The Concept Lattice and Inheritance Concept Lattice are isomorphic latices. The inheritance concept lattice also shows where each property should be declared. One important property of the inheritance concept lattice is that each property appears exactly in one place.

Based on the definitions and algorithm, we implement a semi-automatic tool(called FCAwizard) to support the ontology construction. FCAwizard creates concept lattices in a number of steps(see figure 3): first it builds lists of the names of objects and attributes, and their relations; from this it creates a list of formal concepts. There are then put into the algorithm to build the concept lattice, and displayed to a line diagram(Hasse diagram) format.

4 Conclusions

In this article, we have shown that Formal Concept Analysis(FCA) can be used for semi-automatic construction of ontologies from a given domain-specific context. We have presented our research-in-progress concerning the semi-automatic tool for ontology construction. FCA can help structure and build ontologies. This is because FCA can express ontology in a lattice. The lattice is easy for people to understand and can serve as a guideline for ontology building.

There are many applications of formal concept analysis in the research area of software engineering[8]. A concept lattice can be used to discover cohesive units and suggest a decomposition of the program into modules[9]. Several works[10, 11] apply formal concept analysis to the exploration and search of software component repositories, and consider the stored software components to be objects with certain sets of properties. Works dealing with reverse engineering exploit the classification of information into concepts to recover different elements or formal models from the implementation or architecture of the program[12, 13].

The whole of our work is based on the concept lattice of the Formal Concept Analysis which allows to construct a "well defined" ontology with maximally factorized properties. Since there are many conceptual similarities with the design of an ontology, and an ontology is semantically richer than a UML class model, it makes sense to put the emphasis on ontology design. That is, an object-oriented software designer can design an ontology by organizing object classes in a class hierarchy and creating relationships among classes. UML class model can then be generated from the ontology. Software design artifacts, such as UML class models, and even source code can be generated directly from ontologies, thus effectively speeding up the software development process.

Acknowledgements

This research was supported both by the MIC(Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center)

support program supervised by the IITA(Institute of Information Technology Assessment), and by a grant of the Korea Health 21 R&D Project(0412-M102-0404-0002), Ministry of Health & Welfare, Republic of Korea.

References

1. T. R. Gruber. Toward principles for the design of ontologies used for knowledge sharing. Presented at the Padua workshop on Formal Ontology, March 1993.
2. M. Gruninger and M. S. Fox. Methodology for the design and evaluation of ontologies. In Proceedings of IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.
3. Uschold, M. King, M. Towards a Methodology for Building Ontologies. Workshop on Basic Ontological Issues in Knowledge Sharing. 1995.
4. Fernandez, M.; Gomez-Perez, A.; Juristo, N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Symposium on Ontological Engineering of AAAI. Stanford (California). March 1997.
5. Barbara A. Kitchenham, Guilherme H. Travassos, Anneliese von Mayrhauser, Frank Niessink, Norman F. Schneidewind, Janice Singer, Shingo Takada, Risto Vehvilainen, Hongji Yang, "Towards an ontology of software maintenance", Journal of Software Maintenance: Research and Practice Volume 11, Issue 6, pp.365-389, 1999.
6. M. Annamalai and L. Sterling. Guidelines for constructing reusable domain ontologies. In AAMAS03 Workshop on ontologies in agent systems, Melbourne, Australia, pages 71–74, 2003.
7. B. Ganter and R. Wille, "Formal Concept Analysis, Mathematical Foundations," Springer-Verlag, 1999.
8. T. Tilley, R. Cole, P. Becker, and P. Eklund. "A Survey of Formal Concept Analysis Support for Software Engineering Activities," In G. Stumme, editor, Proceedings of the First International Conference on Formal Concept Analysis-ICFCA'03, Springer-Verlag, February 2003.
9. A. van Deursen and T. Kuipers. Identifying objects using cluster and concept analysis. In Proceedings of the 21st International Conference on Software Engineering, pages 246–255. ICSE99, IEEE Computer Society Press, 1999.
10. C. Lindig. Concept-based component retrieval. In Working Notes of the IJCAI-95 Workshop: Formal Approaches to the Reuse of Plans, Proofs, and Programs, pages 21–25, Montreal, Aug. 1995.
11. Y. Shen and Y. Park. Concept-based retrieval of classes using access behavior of methods. In Proc. of the Int. Conf. on Inf. Reuse and Integration, pages 109–114, 1999.
12. G. Antoniol, G. Casazza, M. D. Penta, and E. Merlo. A method to reorganize legacy systems via concept analysis. In Proceedings of the 9th International Workshop on Program Comprehension, pages 281–291, Toronto, Canada, May 2001.
13. R. Waters, S. Rugaber, and G. D. Abowd. Architectural element matching using concept analysis. In Proceedings of the 14th IEEE Conference on Automated Software Engineering, pages 291–294, Oct. 1999.

Component Contract-Based Formal Specification Technique*

Ji-Hyun Lee, Hye-Min Noh, Cheol-Jung Yoo, and Ok-Bae Chang

Department of Computer Science,
Chonbuk National University, Korea
{puduli, cjyoo, hmino, okchang}@chonbuk.ac.kr

Abstract. When we analyze the business domain, we have to decide what business concepts are to be encapsulated into a component and find what business concepts are to be built by using a reuse component. Also, as a component is reused in the form of a black-box, the reuser must have detailed information about the component, such as the functional and non-functional performance which is necessary to reuse or integrate. So, we will propose a formal approach to design a robust component. First, we analyze a business domain by using Z and category theory. Second, we extract the components and the interfaces from previous analysis results. Lastly, we add component contracts(functional and non-functional performances) to the result. We specify business concept based on DbC which is used broadly to specify the behavior of an interface in an object-oriented area. Also, we will define rules for extraction of components and component contracts from specification. Specially, we will use category theory to analyze the relations between components.

1 Introduction

Component technology is all about interoperation of mutually alien components. Because of this, well-designed component and well-understanded the meaning of component are becoming important issues. Actually, [1], [2], [3], and [4] have formalized syntactic or behavioral aspects of components by using formal languages such as *Z*, *LTL formulae*, and *Wright*. They provide a specification not about how to use and design them but about what they provide. That is to say, they are specifying documents written in natural language into formal specification by using formal language. But, components are integrated in a system rather than used independently. And most of all, as the components are reused in the form of black-box, a reuser must have detailed information about the component. To support this, [5], [6], [7] have described component's interfaces using contracts.

To make this issue work, in our formal specification technique, first, we analyze the business domain by using Z and then we classify components and their

* This work was supported by grant No. R05-2003-000-12253-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 836–845, 2005.

interfaces. Specially, we specify the business concepts based on DbC(Design by Contract) which is used broadly to specify the behaviors of interfaces in an object-oriented area. And we define rules for extracting components and their interfaces from the specification. We also add contracts to every interfaces to provide information which is necessary to reuse. To reuse correctly, the interfaces of a component should state more than type signatures[8].

The main purposes of this paper are as follows.

- The notion of component contract.
- The basic steps and the specification rules for specifying component contracts using Z.
- Analysis of the relation between components using category theory.
- Application by example.
- Discussions about merits/demerits of presented formal specification technique.

We have chosen the specification language Z, because it has gained considerable popularity in the industry and comes equipped not only with methodology [9] but also with some tool support, such as type checking [10] and theorem proving [11]. Z is designed to specify state-based systems which is in proper accordance with the reality of safety-critical systems. An undeniable deficiency of Z is the fact that neither time nor complex control structures can be specified.

2 Component Contract Features

When we apply contracts to components, we found that such contracts can be divided into four levels of increasingly negotiable properties[12].

- Basic contracts(IDL, type systems)
- Behavioral contracts(invariants, pre/post conditions)
- Synchronization contracts
- Quality-of-Service contracts

In this paper, we formalize component contracts as contracts with basic contracts, behavioral contracts and quality-of-service contracts point of view. Previous works on contracts suggest that programmers annotate the implementation of their methods with contracts, in particular, pre/post conditions. Unfortunately, such contracts cannot capture enough information of a component. To fully specify a component, it is necessary to specify components in terms of performance measurement results such as time/space complexity and resource requirements besides the operations they provide and the operations they require.

3 Component Contract Design Approach

As previously stated, CBSE focuses on design rather than coding. Well-designed component will be reused more frequently and a precisely specified component

will be reused more safely. In this section, we will propose steps for designing component based system(CBS) and explain some rules for extracting components and their interfaces from requirements.

3.1 Steps for Specification

We present the formal steps for CBS design. The formal design steps is composed of five steps, which are *Requirements Anlaysis using Z*, *Components and Interfaces Identification*, *Add Contracts*, *Relation Analysis between Components*, and *Add Non-functional Contracts*[13]. This design approach is developed for extraction of high cohesive components and their interfaces from requirements, addition of contracts, and analysis of relation between components. For more detail and exact analysis of requirements and extraction of more cohesive components, we use Z as formal specification language and category theory for analysis of relationships.

3.2 Component Specification Rules

Well-defined components and well designed interfaces are more important than any thing else in the CBSE. Several definitions for specification of software components using contracts are given [7]. We have to identify all of the services needed from the requirements, and from these we have to define the components and interfaces. In this section, we will define specification rules for this.

Rule1(Types are components): We manipulate types as components because types are encapsulating maximal sets of a single business concept.

Rule2(Each of modularized problems are components): We attempt to put into practice the notion of modularization because the task has the interdependencies between different aspects of the problem. It is clear that each of the modularized problems will take the form of subsystems from the more complex overall system.

Rule3(Schemas for the specification of system's operation are components): To identify two adjacent states of system, Z uses Δ, Ξ notation. Schemas used for the specification of system's operation are apt to perform a common business logic.

Rule4(Each operation of modularized problems is interface): Modularized problems have several operations, and these act as interfaces of a component. In Z, each schema consist of triple, namely, schema name, signature part, and predicate part. In specification of a component contract, the basic schema structure of Z represents as follows:

- **Schema name:** Naming a schema introduces syntactic equality of a component or operation. According to the Rule5, we give schema name.
- **Signature part:** This part introduces interfaces and invariants in case of a component, and declares variables and their types using mathematical data structures such as sets, relations in case of an operation . According to Rule6 and Rule7, we give variable and instance name.

– **Predicate part:** We describe constraints on the values that the identifiers in the signature part may take.

Rule5(Naming Schema): We decided to express that the schema of a component be $Schema\_name_{comp}$ and the schema of an operation can be $Operation\_name_{comp\_name}$ which *comp_name* is the component name providing a given operation.

Rule6(Naming Instance): If an interface is an *indirect-use interface*[7] or an *indirect-use operation*[7], we describe the name of instance in the form of $InsName_{indir}$: *ComponentName* or $InsName_{indir}$: $InterfaceName_{comp\_name}$.

Rule7(Naming Variable): The input variables which correspond to precondition are named in the form of $Variable\_name_{pre}$ while the output variables which correspond to postcondition are named in the form of $Variable\_name_{post}$. Component composition or assembly is the combination of two or more software components that yields a new component behavior. A component composition standard supports the creation of a larger structure by connecting components and the insertion or substitution of components within an existing structure. We specify these relations between components by using binary relation.

4 Application by Example

System Analysis. There is a system whose description is as follows:

Name. J's Computer Parts
Version. 1.0
Brief Description. J wants to lower the cost of doing business by selling her computer parts directly to the end customer, through an e-commerce web-based sales model.

1. User authentication: Registered users would first log into the web site, and only registered users should be able to browse and purchase.
2. An online catalog: User should be able to browse her complete product line on the web and view details of each product.
3. Online quote generation: While browsing the catalog, a user should be able to view the current shopping basket.
4. Specialized pricing functionality: Users who items in bulk should get a percentage discount.
5. Order generation: Once the user select his/hers and has committed to ordering the products, a permanent order should be generated.
6. Billing functionality: Once the user has placed order, we should bill it to him/her. If the user does not have required balance, the order should be cancelled.

Step1: Analyze Requirements Using Z. First, let us attempt to put into practice the notion of abstraction. The above paragraph discusses what is necessary. Now, in order to describe the system's behavior, we introduce certain

non-empty sets as types. The type *Customer* stands for all possible users. Such users should be able to browse her complete product line. So the type *Product* denotes the set of all possible products. *Quote* and *Order* denote products which are in the shopping cart and is a committed order. Any set of objects whose internal structure has been ignored for one reason or another is referred to as a basic or generic type.

$$[Customer, Product, Order]$$

The second dimension of complexity in our task is the interdependencies between different aspects of the problem. Thus it is clear that we may treat this problem separately that we enumerated before(brief description 1-6). We adopt them as a schema for modularization of our specification. Modules for an example system are *Product*, *Customer*, *Quote*, *Pricer*, *Order*, and *Billing*. In this example, we center on the online quote generation(module *Quote*).

Quote represents a shopping cart that a customer is working on and is a set of products that the customer is interested in but has not committed to buying. Namely, the customer surfs our e-commerce web site, quote stores all the products and quantities he or she is interested in. A quote consists of a series of line items. Every line item represents one particular product and a quantity of a single product which the customer wants. In this respect, the two attributes:

- *lineitem* (temporary product and its quantity what the customer wants)
- *purchase* (Transforming line items into permanent order)

should be sufficient. The *lineitem* has the mathematical structure between *Product* and quantity, and quantity must be a positive natural number. So, we introduce auxiliary functions for *lineitem*.

TempPurchase

$$lineitems : Product \rightarrow \mathbb{N}$$

When we convert this temporary line item into permanent order, we need to generate unique ID for the identification of the others. This requires additional type and schema as follows:

$$[UniqueID]$$

PurchaseOrderID

$$id : UniqueID \rightarrowtail Purchase$$

And then, we add product and quantity as long as the quantity of an item is not zero. Then, we compute the subtotal price. Finally, we compute tax and total price. We note that *subtotal*, *totalprice*, and *tax* must be a positive natural number, that is

$$subtotal, totalprice, tax : \mathbb{N}$$

ComputeSubtotal

$$
\begin{array}{l}
Purchase \\
p : Product \\
subtotal : \mathbb{N} \\
\hline
subtotal = p.price * \operatorname{ran} lineitem
\end{array}
$$

ComputeTax

$$
\begin{array}{l}
Purchase \\
s : ComputeSubtotal \\
tax : \mathbb{N} \\
\hline
tax = s * (1/10)
\end{array}
$$

ComputeTotal

$$
\begin{array}{l}
Purchase \\
s : ComputeSubtotal \\
t : ComputeTax \\
total, total' : \mathbb{N} \\
\hline
total' = total + s - t
\end{array}
$$

For a large number of purchases, an appropriate mathematical representation would be following schema:

Purchase

$$
\begin{array}{l}
TempPurchase \\
ComputeTotal \\
ordered : Product \times \mathbb{N} \\
issued : Customer \pfun Order \\
customers : \mathbb{F}\, Customer \\
orderitem, orderitem' : \mathbb{F}(Product \times \mathbb{N}) \\
\hline
\operatorname{ran} lineitem > 0 \\
\operatorname{dom} issued \subseteq customers \\
\#ordered = 0 \\
orderitem' = orderitem \cup odered \\
lineitem' = lineitem \setminus odered \\
lineitem' = \emptyset
\end{array}
$$

Step2: Extract Component and Interface. Now, we extract components and interfaces for this example by using specification rules. First, the three generic types are components by specification **Rule1**. And then, we select *TempPurchase* and *Purchase* as a component according to **Rule2**. According to

Rule4, *ComputeSubtotal*, *ComputeTax*, *ComputeTotal* is a common procedure for computing price. So, we select this as a component named *Pricer* and each of operations is regarded as an interface. Now, each component and operation is renamed by using **Rule5** and **Rule7**.

Step3: Analyze Relation between Component. In this paper we will use categories to represent possible abstractions of domains. Given a particular domain, we construct a category where objects represent the concepts, and arrows represent axioms. We do not consider other aspects of the domain. But, there is no need to reason about a concrete representation for concepts or axioms at this process.

Now, let's adjust category theory to the component design. The *Customer* component has four states:

- *LogIn*: The customer log in to Jasmine's Web site for the first time.
- *BrowseProduct*: To start adding products to the Quote, the customer can browse the list of products.
- *LineItem*: The customer is able to view and modify the quote at any time.
- *Purchase*: When the customer has selected products all s/he wants, s/he can convert quote into an order
- *Billing*: When the customer has placed order, s/he has to pay for products.

The initial state for the *Customer* component is *LogIn* and the Figure 1 is a diagram which shows the process of the order generation. The defining attribute values in the state diagram at which the operation arrow begins are the preconditions for the operation. If the operation appears as an arrow between more than one pair of states, the preconditions for the operation are the distributed intersection of the definitions of all of the states at which the arrow starts. Likewise, the postconditions are defined by the distributed intersection of the definitions for the states at which the operation arrow terminates.

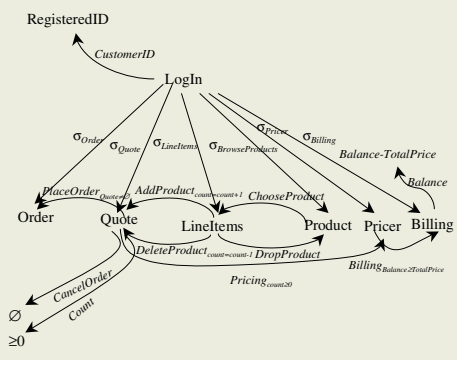


Fig. 1. Customer State Diagram for J's Computer Parts

From this figure, we can drive functional relationships as follows.

$PlaceOrder : LineItem \rightarrow Purchase$
$AddProduct : BrowseProduct \rightarrow LineItem$
$Billing : Purchase \rightarrow Billing$
$CancelOrder : Billing \rightarrow Purchase$
$DropProduct : LineItem \rightarrow BrowseProduct$

From the Figure 1, we can see that *LineItem* maps to *TempPurchase* and *BrowseProduct* maps to *Product*. Namely, component *Product* and *TempPurchase* have one-to-one relationship. Because *Quote* consists of *TempPurchase*, *Quote* and *TempPurchase* have many-to-one relationship. In this way, we can get overall relationship as follows:

$[JsCompouterParts]$

$Customer$
$Product$
$Order$
$Quote$
$TempPurchase$
...
$1 \leftrightarrow 1 : \mathbb{P}(Customer \rightarrowtail Quote)$
$1 \leftrightarrow * : \mathbb{P}(Quote \leftrightarrow TempPurchase)$
$1 \leftrightarrow 1;\ \mathbb{P}(TempPurchase \rightarrowtail Product)$
...

Step4,5: Add Contracts, Add Non-Functional Contracts We add behavioral contracts to which the components and interfaces we extracted in step 2. From the category state diagram, we can get pre/post conditions.

$Quote$

$ejbCreate_{required}$
$ejbRemove_{required}$
$makeUniqueID$
$getNumberOfItem$
$getLineItem$
...
$addProduct$
$getProduct_{indir} : TempPurchase_{comp}$
$getQuantity_{indir} : TempPurchase_{comp}$
...
$resouce : \mathbb{P}\, RESOURCE$
$t\_complexity : \mathbb{T}$
$s\_complexity : \mathbb{S}$

$resouce = DataBase$
$t\_complexity = min - sec(0, 1 \operatorname{div} 1000)$
$s\_complexity = KB - Byte(2, 37)$

5 Discussions

The concise advantages of component contract-based specification technique proposed in this paper are as follows.

- It is possible to design consistent component by standardization of component design process.
- Component designer can provide more detailed information to implementor using component contract.
- The category theory used for understanding relation between objects can be helpful to understand accurate relations between components.
- Open design approach can improve the reusability and the essential objective of component based software engineering.
- Formal specification language Z has gained considerable popularity in industry and comes equipped not only with methodology but also with some tool support.

However, this process has the following disadvantages.

- Formal specification language Z is complicated and takes much man-power, time and cost.
- Categorical approach imposes some formal restrictions. First, arrows can only express binary relationships. Second, a meaningful composition operator must be defined for arrows. Third, the composition operator must be associative[14].

6 Conclusions

We're described so far what component contract is, and why we need to specify component contract. And we proposed an design steps for this. In the formal steps, we described how we can apply the category theory and DbC in CBS. Finally, we applied this approach to a simple ordering system.

In this paper, we saw interfaces as contracts and proposed a formal specification technique to specify components by using contracts. Sometimes, we have to redo large portions of code because we need to change an interface to get at some piece of data. This can be avoided by 10 minutes of effort during design time. So, we tried to formalize the design steps by applying a notion of component contract. As we noted previously, this process won't work in every instance. This process takes a lot of man-power, time and cost, but it is helpful for making a cohesive and reusable component.

References

1. Sullivan, K., Marchukov, M., and Socha, J., Analysis of a Conflict Between Aggregation and Interface Negotiation in Microsoft's Component Object Model, *IEEE trans. SE*, Vol. 4, pages 584-599, 1999.

2. Sousa, J. and Garlan, D., Formal Modeling of the Enterprise JavaBeans ^{TM} Component Integration Framework, Proc. FM'99, pages 1281-1300, 1999.
3. Kreuz, D., Formal Spcecification of CORBA Services using Object-Z, The University of Hamburg-Harburg Telematics Group 21073 Hamburg, Germany, 1999.
4. Nakajima, S. and Tamai, T., Behavioural Analysis of the Enterprise JavaBeans ^{TM} Component Architecture, Proc. of the 8th international SPIN workshop on Medel checking of Software, pages 163-182, 2001.
5. Han, J. , Temporal Logic Based Specification of Component Interaction Protocols, Proceedings of the 2nd Workshop pf Object Interoperability at ECOOP, 2000.
6. Reussner, R. H., Enhanced Component Interface to Support Dynamic Adaption and Extension, Proceedings of the 34th Hawaiin International Conference on System Science, *IEEE*, 2001.
7. Lee, J. H. , Yoo, C. J., Chang, O. B., Component Contract-Based Interface Specification Technique Using Z, IJSEKE(International Journal of Software Engineering and Knowledge Engineering) Vol. 12 No 4, August 2002.
8. Szyperski, C., *Component Software: Beyond Object-Oriented Programming*, Addison-Wesley, 2002.
9. Potter, B., Sinclair, J., and Till, D., *An Introduction to Formal Specification and Z*. Prentice Hall, 2nd edition, 1992.
10. Spivey, J. M., The fuzz manual, Computng Science Consultancy, Oxford, 1992.
11. Bowen, J. and Gordon, M., Z and HOL. In *Z User Workshop*, Workshops in Computing, pages 141-167, Springer-Verlag, 1994.
12. Beugnard, A., Jexequel, J.-M., Plouzeau, N. and Watkins, D., Making components contract aware, *IEEE Software*, Pages 38-45, June 1999.
13. Lee, J. H., Yoo, C. J., Chang, O. B., Component Contract-Based Component Contract-Based Process for High Level Design, SNPD03, Lubeck, Germany, pp. 6 11, 2003.
14. Wendorff, P., A Formal Approach to the Assessment and Improvement of Terminological Models Used in Information Systems Engineering, ACM SIGSOFT Software Engineering Notes, Proceedings of the 8th European software engineering conference held jointly with 9th ACM SIGSOFT symposium on Foundations of software engineering, Volumn 26 Issue 5, Sep. 2001.

A Business Component Approach for Supporting the Variability of the Business Strategies and Rules

Jeong Ah Kim[1], YoungTaek Jin[2], and SunMyung Hwang[3]

[1] Kwandong University
clara@kd.ac.kr
[2] Hanbat National University
ytjin@hanbat.ac.kr
[3] Daejeon University
sunhwang@dju.ac.kr

Abstract. In this paper we present how to use rule for manipulating the variability of a business. We can classify the business component as business domain components and business process components. The most important thing to design the component, designer should identify the variability among the same business components. If we put the variability into the components, we should modify those components whenever we reuse that. It is not easy to modify the component since the user of that component might not be same with the writer. We extend component property to business rules. We define the component property to customize the component execution environment. Like this, we can define the difference business strategy in rules. With our rule engine, the component can use the rule during the component execution time so that the user of component can redefine or rewrite the rules according to their business context. Based on this approach, the reusability of component is improved. In this paper we describe new component architecture including rule component, our rule engine and case study.

1 Introduction

A component based development is a promising way that software products can be made by obtaining and assembling pre-developed components, and then software productivity and quality can be improved. This component-based approach presents different views from the traditional software development in sense that the approach needs much effort in requirements, testing and integration rather than designing new code [1].

In the business point of view, business organizations not only have to cope with the continuously changing requirements, environment and business processes, but also need the reuse of the business process to adapt to new business requirements as quickly as possible[2]. Business applications also must be designed to be adapted to the modification of the existing process and the rearrangement of the process to create new one. Then, the component technology can be combined with the business application development to support both the business and the software [3].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 846 – 857, 2005.

A business component may be defined as a software implementation to the autonomous business concept or the business process while keeping the basic characteristics of components. The business components consist of business domain components which represent the static aspects, business process components which describe the dynamic and variable aspects of the business [4,5,6]. In the process-centered application development using business components, it is important to consider how components can be built and how business processes and business rules can be automated and represented within the components. Many researches presented the modeling and the building guidelines of the components comprising business objects [7]. The research on the business process automation used the workflow and rule-based technologies [8,9]. However, there are only a few case studies on their adaptability and reusability when the business strategies and rules are changed [10]. That is to say, the way of handling variability inherent to the business may be differed according to how business components are designed.

In this paper, we present a way of building the business application using rule centered component based development. This approach tries not to bury variable business aspects in one component but to separate them into another component in designing the business components. Then, it has to be able to meet the changing requirements by replacing the separated variable parts not by modifying the codes during component running. For this, we componentized the business process and associated rules specified by a rule specification into a rule component and developed a rule interpreter executing the rules. As a case study, we developed rule-based business components for the insurance products, the subscription and the insurance planning in an insurance sales domain.

Section 2 of this paper describes our model and concept of business component approach. Section 3 presents general idea of rule-based component development approach to handle the variability of the business rules. In Section 4, we describe the observations on the representation, construction and its usage when applying rule based component for insurance sales system. Also, this section shows the execution results when we used the rule components in the case study and describes the conclusion with how this approach fits well into the business application such as the case study. Section 5 concludes this research.

2 Component-Based Business Development Approach

A software component is defined as the independently developed and deployed software unit having well-defined interfaces to perform specific functions. It may also be assembled with other components to produce the entire application [11]. The construction steps of the component follow domain analysis and modeling, component identification, component design and implementation process. The components are usually delivered in black box style whose source code is unavailable, and component specification or the associated information is required to understand the component. It is also necessary for user to adapt components to the new context whose requirement can be changed even if components are well-designed and built for general usage in the domain. The business components consisting of the aggregate of business objects are the autonomous, independent and large granularity unit which

implements the real world business concepts. It may be categorized into three points of view: granularity of component, characteristics of component, implementation of components [12].

The process of building the business application through the component based development approach may roughly divided into two steps. One step is to make the business components and other step is to assemble the pre-developed components to produce the target business applications. The most important factor in the construction of business components is to make the components which can be reused in the same domain or among similar domains. The architecture layers organizing the business components, which may be defined differently depending on the approach, consist of UI component layer, enterprise layer which can further be decomposed into business domain and business process component layer, and resource layer[4,6]. This paper considers only the business domain and business process layer.

2.1 Business Domain Component

It corresponds to the business domain model which represents entities defining the characteristics of the business domain and defines the states necessary for the business process. It represents the static view of the business and corresponds to the components consisting of objects in the technical view. Fig.1 shows the Adaptive Component Model which provides the business concept to build the business domain component [13].

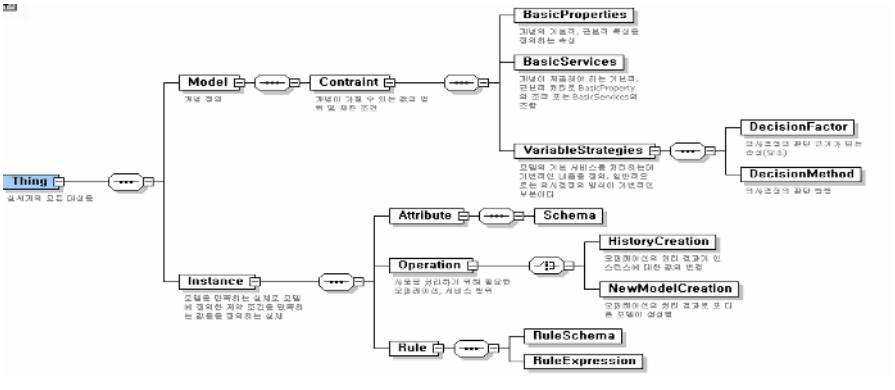


Fig. 1. Adaptive Component Model

The model defines constraints (interpretation range and method) to explain the thing. The constraints are divided into essential parts and variable parts. The essential parts mean the attributes to explain the thing and includes the basic services which interpret and manipulate the attributes. The variable ones mean the continuously changing properties according to the situations. It can not be predicted what the variability is or what variability happened. However, the variable parts also have to define necessary elements to decide which are variable one and its interpretation methods.

For example, customer may become the business domain component which represents the customer subscribing various insurance. The customers essentially have the information such as name, address and social security number. This information is the necessary to describe and explain the customers. Those customers perform basic operations such as enrollment, payment and order operations. Each company has different strategies to manage customer class. The factor for deciding the customer class may be sex, age and the buying power. Because this factor can be changed, we can define this as the variability. Also this variable factor has to be defined and interpreted separately.

2.2 Business Process Component

Business process components are the componentization of the variable and dynamic process occurred in the business system. These process components represent the workflow of tasks and act on the business domain components. Through this componentization of the process, not only the many parts of the existing process can be reused without affecting the basic system design, but also the modification or adaptability requirements can be limited to either the single components or the assembly of the corresponding components.

The business process model defines business workflow and working conditions. The business process is the main element of this model and refers to business states. The business rule is the rule for controlling the one or more business process behavior and stored on the rule base. The business process component decides its progress according to the value of the state variable defined in the business domain component. The business domain component is the entity for defining not only the domain properties but also the necessary information for business processing. We define the business rule as both business process rule and business domain rule. The business domain rule defines the variable method which interprets the value of the thing.

For example, to compute the legal age or insurance age, we can use the social security number (because it consists of birth date plus additional number in Korea). The business process rule is the rule for defining the variable parts of the type, sequence and processing conditions of the business process necessary to perform business task. For example, the insurance planning has the sequence of the entry condition check followed by insurance payment followed by pension amount computation.

3 Rule-Based CBD Approach for the Variability

It is very important and difficult to decide about how we can set up the granularity of the components when working with the component based development. The reusability of component would be decreased if it contains the general concept while the reusable extent of it may small if it implements the specific and concrete concept. It is realistically hard thing to reuse the components as it is. The reusability through the modification of components requires the understanding of their internal implementation, so that it becomes same difficult as the maintenance of the general module.

The business requirements are continuously changed during the system development or after the completion of the development. If the variable aspects such as business strategy and principles are included in the business components, it needs to modify code of the component whenever the modifications are occurred. It can bring the improvement of the component reusability if we can draw the clear line between the variable parts and the common parts and can redefine the variable ones separately when developing components. There may be difficult cases to decide which parts are variable. Generally, the variable parts in the software mean the modification of the models and model interpretation method. In the concrete sense, this may be the addition of attributes, modification or addition of operation names and the modification of implementation method of the operations.

However, it is very difficult to predict all the things the why and which attributes of which components can be added and which operations can be renamed. As the generalization of those cases, this can be considered as the modification of the models and methods of model interpretation. It means that the models and the methods of model interpretation can be changed whatever components. Therefore, it needs to separate and extract those changed parts from the business domain and process components so that it can adapt flexibly to the variability of the business strategies and rules. Then, we make a rule component to capture those requirements. Figure 2 shows the construction of the rule components.

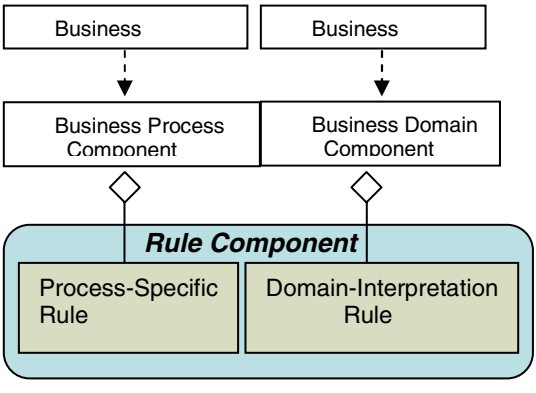


Fig. 2. The construction of rule component

The rule controls one or more behaviors of business. It is not only the constituents of rule components but also contains the business process and domain rules. The business domain rule defines the method of interpretation about the value of the entity in the domain. For example, obtaining the insurance age or legal age can be acquired by using the security number. The business process rule is the one which describes the kind of tasks, sequence and processing conditions necessary to perform a task.

The business process represented by those rules creates the business domain or accesses the existing one. Within this business process which represents the sequence of the basic business activity, may be hierarchically decomposed into activities and

sub activities. For example, 'Insurance Planning' can be specified in the sequence of 'entrance condition check', 'premium calculation'. The tasks are not progressed to the next steps until the entrance condition check is completed. In this business processing, the modification of the processes or business rules are inevitable but many parts not modified can be reused.

Then, there are needs for representing business logic into the pre-defined rules or rule sets forming the rule components to adapt new business requirements quickly. Those rules or rule sets have to be able to be not only modifiable without affecting the rest of the system but also testable independently. Through this, highly adaptable business services can be provided.

3.1 Rule-Based Component Architecture

To increase the reusability and extendibility, it is rather separate the business variation from the business components than encode it into the component inside. We defined these variable parts as the rules, and then rule components. Figure 3 describes a rule based component layer. We define the role of rule components added in this layer. It ensures that the extendibility and flexibility may be achieved by the separation of business rules and strategies with the variation from the common services of components.

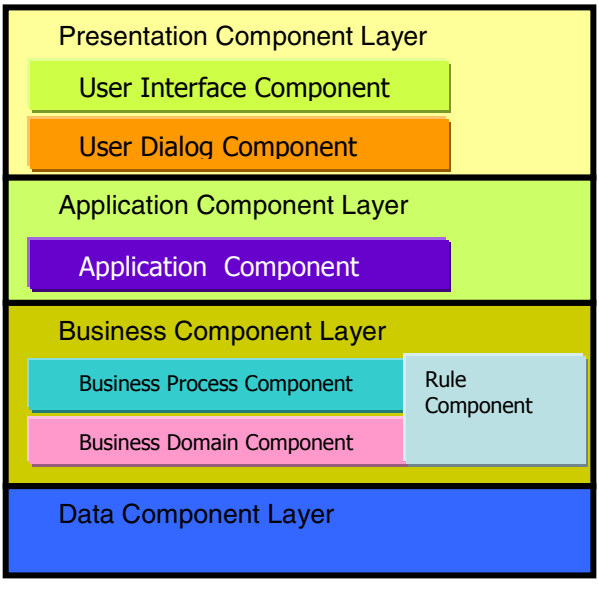


Fig. 3. Rule based component layer

The role of rule component is to enhance the extendibility and flexibility of the components through the separation of basic component services into variable one which incorporates the business rules and strategies. The relationship among components are described in Figure 4.

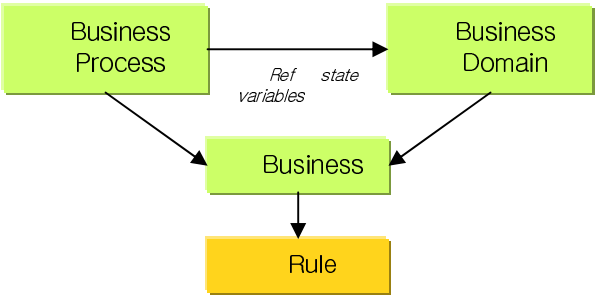


Fig. 4. The relationship among the components

3.2 The Construction of Rule Components

The product rules to which the domain components refer are the basic product information represented by PRODUCT_XML and include product name, interest, insurance type, entry limit and age at entry etc. The process rules to be included in the business process components are presented according to the rule syntax in Figure 5 and they define the process themselves and the processing method.

```
ReturnType Rule Heading {
        IF ( condition)
        THEN
        ELSE
        ADDITIONALINFO
```

Fig. 5. Syntax for Rule

For example, there are the rules such as 'age at entry' check rule and 'entry limit' check rule. In the case of defining the rules, the common rules and dependent rules of insurance products are defined separately. We also introduce the rule execution sets when many rules can be grouped to form a sequential business workflow. Figure 6 represents the construction of the rule execution sets.

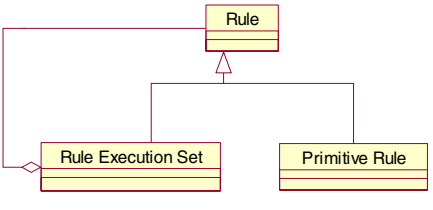


Fig. 6. The construction of rule execution set

A business process can be completed through the execution of these sets. The business components use these sets as the call units of the rules. All the things in the

same business processing must have same number and name of the rule execution sets. These rule sets may be considered as one rule but these sets have to include the lists of rules to be called for business processing. Figure 7 describes an application of the rule execution set checking the entry conditions.

```
Boolean MainInsuranceEntryCheckRule() {
    IF( FIRE("InsuranceProduct_MainInsurance_MaximumEntryAmountCheckRule") a
nd
    FIRE("InsuranceProduct_MainInsurance_MinimumEntryAmountCheckRule") and
        FIRE("InsuranceProduct_MainInsurance_MinimumEntryAgeCheckRule") and
        FIRE("APlanInsurance_MainInsurance_MaximumEntryAgeCheckRule") )
    THEN Collection(Boolean true)
    ELSE Collection(Boolean false)
    ADDITIONALINFO ELSE "Not Pass EntryCheckRule" }
```

Fig. 7. Example of rule execution set

Using the rule editor and rule interpreter, the result screens of an application of the rules or rule sets is presented. Figure 8 describes the fact that the modified business rule is immediately reflected without changing the domain components when the minimum entry age is changed to legal age at entry.

4 Case Study: The Development of an Insurance Sales System

This application has the following properties. First, the same type (life or accident) of the insurance products has similar characteristics and many parts can be shared when the insurance product manager plans a new product. Second, the insurance processes or rules are similar but these are diverse and variable. This application consists of the enrollment of the insurance products, the insurance planning, the result processing of the insurance planning and the issue of proposal form. This system compares one insurance product with another product and then provides various information to the customers who are interested in the specific insurance product. It also compares each application conditions within single insurance product.

The insurance product means the sales unit which can be sold to the customers and consists of the package of the insurance. To understand the product, it is necessary to understand the business documents describing the products. These documents contain not only the basic descriptions of the insurance products but also specify the necessary criteria and processing method about each insurance process. Using these documents, the business domain components consisting of the business objects can be identified.

The business process such as the enrollment of the insurance product means the enrollment of the rules necessary to deal with the product. The current product rules contain the basic data of those products such as name, interest and code etc. As an example of the business process, the insurance planning inputs the insurance product which a client wants to make an entry and various conditions and then shows the results such as the insurance premium and surrender value after simulation. There always exist variable factors in the business processes and these factors must be able

to be defined and interpreted separately from the business domain. Then, the changing business requirements can immediately be adjusted without concerning the modification and re-distribution of the components.

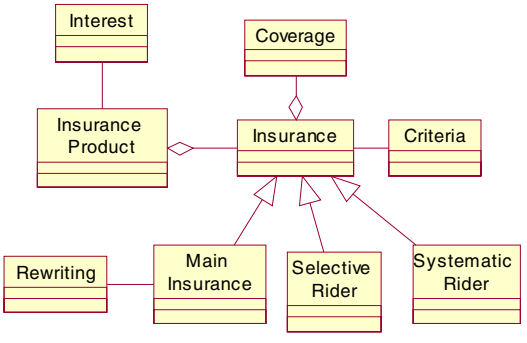


Fig. 8. Construction of the insurance product domain component

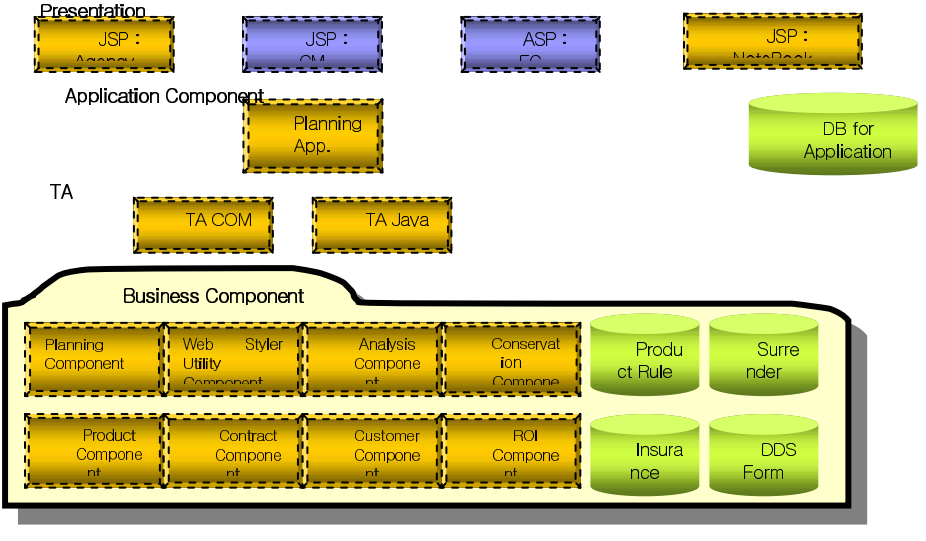


Fig. 9. Component layer of the insurance sales supporting system

Figure 9 shows the sample product line applying rule component in insurance application. In business component folder, we present the core business asset and rule component to customize that business component according to business context. With those business core assets, we could product different products that they have different presentation technologies and different planning strategy. For different strategy, we apply rule component. For different presentation, we apply adapter component.

Using the rule editor and rule interpreter, the result screens of an application of the rules or rule sets is presented. Figure 10 describes the fact that the modified business rule is immediately reflected without changing the domain components when the minimum entry age is changed to legal age at entry.

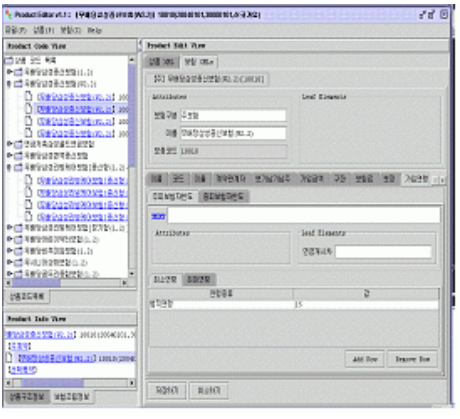


Fig. 10. Screen shot of editing rule data

Figure 11 shows the rule for defining the insurance entry amount as the common rule of the entry design according to the rule syntax. It checks if there is an insurance which clients want to join, it computes the entry amount suitable. Otherwise, it returns false value. It is used for the entry design of all insurance products.

Figure 12 represents the rule execution set when many rules can be grouped to form a sequential business workflow. The modification of business process or business rules mean the modification of the rule set and the modified rules can be applied without changing the process component.

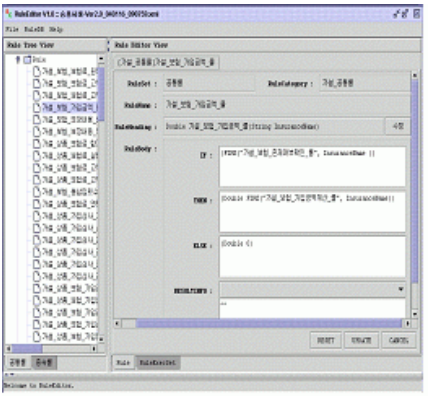


Fig. 11. Screen shot for editing rules

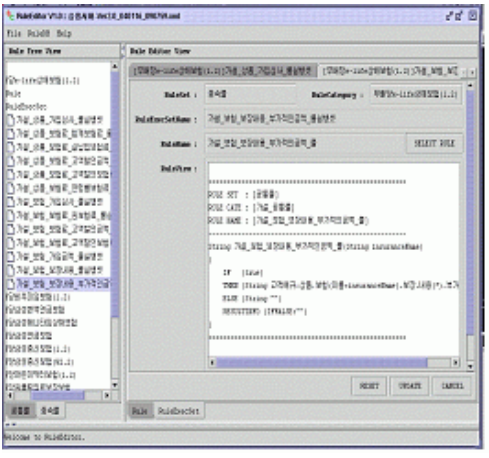


Fig. 12. Screen shot for editing rule execution set

5 Conclusion

The most important factor in the construction of business components is to make the components which can be reused in the same domain or across domains. In the business point of view, business organizations not only cope with the continuously changing requirements, environment and business processes, but also need the reuse of the business process to adapt new business requirements as quickly as possible.

Business applications also must be designed to be adapted to the modification of the existing process and the rearrangement of the process to create new one. The insurance sales system as a case study has the properties that the insurance processes or rules are similar but diverse and variable. For the time-to-market, component technology can be introduced but it can not entirely resolve the variability and reusability problem due to the dynamic characteristics of this system. For this, we componentized the business policy and associated rules specified by a rule specification and developed a rule interpreter executing the rules. We built rule-based business components for the insurance products, the subscription and the insurance planning in the insurance sales domain.

By using this rule component approach, we show that the modification of the business strategies and rules can be reflected and reusability can be increased.

References

[1] M. Morisio and C.B. Seaman et al, Investigating and improving a COTS-based software development process, ICSE 2000, pp.31-40,2000.

[2] Herbert Weber, Asuman Sunbul and Julia Padberg, Evolutionary Development of Business Process Centered Architectures Using Component Technologies, Technical University Berlin.

[3] Greg Baster, Prabhudev Konana, and Judy E.Scott, Business Components: A case study of Bankers Trust Australia Limited,CACM,vol.44,no.5,pp.92-98,2001.
[4] Peter Herzum and Oliver Sims, The Business Component Approach, OOPSLA'98 Business Object Workshop IV, 1998.
[5] Peter Herzum and Oliver Sims, Business Component Factory, OMG Press,2000.
[6] Hans Albrecht Scmid, Business Entity and Process Components, Business Object Design and Implementation III, OOPSLA'99 Workshop Proceedings, D. Patel, J. Sutherland and J. Miller(Eds).pp.131-145, 1999.
[7] Sandy Tyndale-Biscoe and Oliver Sims et al, Business Modeling for Component Systems with UML,EDOC'00, 2000.
[8] Stefan Schreyjak, Using Components in Workflow Activities. Business Object Design and Implementation II, OOPSLA'96, OOPSLA'97 and OOPSLA'98 Workshop Proceedings, D. Patel, J. Sutherland and J. Miller(Eds),pp. 126-135, 1998.
[9] Pamela M. Rostal, Agent-Oriented Workflow: An Experience Report, Business Object Design and Implementation II, OOPSLA'96, OOPSLA'97 and OOPSLA'98 Workshop Proceedings, D. Patel, J. Sutherland and J. Miller(Eds) pp.62-75,1998.
[10] Hans Albrecht Scmid and Fernando Simonazzi, Business Process are not represented adquately in Business Applications and Frameworks!, Business Object Design and Implementation III, OOPSLA'99 Workshop Proceedings, D. Patel, J. Sutherland and J. Miller (Eds). pp.76-85, 1999.
[11] Clemens Szyperski, Component Software -Beyond Object-Oriented Programming, Addision Wesley,1998.
[12] Jun Ginbayashi, Rieko Yamamoto and Keiji Hasimoto, Business Component Framework and Modeling method for Component-based Application Architecture, EDOC'00, pp.184-193, 2000.
[13] JeongAh Kim, Adaptive Component Model, Technical Report, Kwandong University, 2004.

A CBD Application Integration Framework for High Productivity and Maintainability

Yonghwan Lee[1], Eunmi Choi[2,**], and Dugki Min[1,*]

[1] School of Computer Science and Engineering, Konkuk University,
Hwayang-dong, Kwangjin-gu, Seoul, 133-701, Korea
(yhlee,dkmin)@konkuk.ac.kr
[2] School of Business IT, Kookmin University
Chongung-dong, Songbuk-gu, Seoul, 136-702, Korea
emchoi@kookmin.ac.kr

Abstract. Under a rapidly evolving e-business environment with frequent employee turnover and limited resources, a large volume of web-based software development projects require the CBD method for highly reusability and flexibility. This paper presents the architecture of our CBD application integration framework that has been used for developing large-scale e-business applications to achieve high development productivity and maintainability. This framework is flexible and extensible dynamically so as to be customized to various applications. It provides a development and maintenance toolkit as well as common business logics that can be reused with a few adaptation steps. This framework copes with a number of software architectural qualities in various aspects.

1 Introduction

One of the major goals of e-business projects is to develop an application fast and effectively that not only satisfies the given functional requirements, but also handles frequent changes of its requirements. For this purpose, many e-business development projects employ very flexible and extensible application frameworks that produce high development productivity with high software quality [5]. An application framework is a semi-application [1, 9] of which some parts may be changed or reused. The framework architecture models the interaction among the major components to satisfy non-functional qualities and functional requirements. The framework provides skeletons or template codes to extend the functionality of its business logics. In addition, it provides the common services, such as security, log, naming and caching services. An advanced framework supports a tool-kit to plug a business logic component into the framework.

* Corresponding author: Dugki Min (dkmin@konkuk.ac.kr).
** This work was supported by the Korea Science and Engineering Foundation (KOSEF) under Grant No. R04-2003-000-10213-0. This work was also supported by research program 2005 of Kookmin University and Kookmin research center UICRC in Korea.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 858 – 867, 2005.

In order to cope with frequent changes of the functional and quality requirements, the semi-application framework could solve the following design forces. The first force is that the architecture of the framework should be flexible so that it can easily afford to change the interaction among its major components [2]. Second, since the policies, rules, and environment configurations of e-business applications vary according to the applied business domains, a framework should provide a dynamic reconfiguration mechanism that can accept the changes easily without affecting the framework architecture [12]. Third, an application framework should provide development and operational environment for dynamic configuration. The last force is to provide effective mechanism for monitoring and managing various contents and resources in both application and middleware service levels [7].

In this paper, we present the architecture of our CBD application integration framework that has been used for developing large-scale e-business applications to achieve high development productivity and maintainability. This framework is dynamically flexible and extensible so as to be customized to various applications. It provides a development and maintenance toolkit as well as common business logics that can be reused with only a few adaptation steps. This framework copes with a number of software architectural qualities in various aspects. To achieve the flexibility, the framework architecture is designed to solve the following problems: flexible control flow, flexible maintenance, decoupling between UIs and business logics, development separation of each architecture tier, and flexible changes of business logic interfaces. Also our framework achieves the extensibility design forces, such as extensible request processes, functional extensible in multiple-instance environments of Web application servers, and extensible login component.

The remainder of this paper is organized as follows. In section 2, as the related work, we compare three web application frameworks; Struts, Spring, and Hibernate. Section 3 presents the overall architecture of our application framework. Section 4 explains the flexibility aspects of the framework. Section 5 shows the extensibility aspects of our framework. Section 6 draws a conclusion.

2 Related Works

We can classify frameworks into a system infrastructure framework, a middleware integration framework, and an enterprise application framework [9]. The system infrastructure framework is used to effectively develop system infrastructure, such as operation system. The middleware integration framework is required to integrate distributed applications and components. The enterprise application framework is used to develop business application. The framework proposed in this paper is in the category of enterprise application framework.

In the enterprise application framework, there are three popular web application frameworks: Struts [13], Spring [14], and Hibernate [15]. The Struts is the framework for presentation tier, which contains the MVC architecture pattern [3]. It is an extensible web application framework based on the design patterns approved [11]. However, it does not cover the business logic tier like the EJB. The Spring is mainly re-

sponsible for managing objects in business logic tier. It uses a layered architecture pattern and is good for test-based projects. The Spring provides infrastructure services required for application development. The Hibernate is the framework for mapping between object and its relation table. It provides an association, a composition, an inheritance, and a polymorphism relationship. In addition, it provides a powerful query with Hibernate query language.

Those three kinds of frameworks cover their specific target tiers, but do not act as systematic framework through all of the three tiers. They do not have dynamic configurable mechanism in a business area for a software architecture, policy, rule, and environment configuration. Also, there is no systematic integration with development tool for high productivity, and easy maintenance through the three tiers.

3 Framework Architecture

This section presents the overall framework architecture that we propose in this paper as in Figure 1. The framework architecture is composed of three major subsystems: Presentation Tier, Business Tier, and Admin Console. The Presentation Tier accepts requests through client browsers, processes session management, security, and data translation, and transfers the requests to EJB-module which contains business logics. The Business Tier contains the business logics to process the requests from the Presentation Tier with the help of EJB-module. The Admin Console has the development tool and management tool.

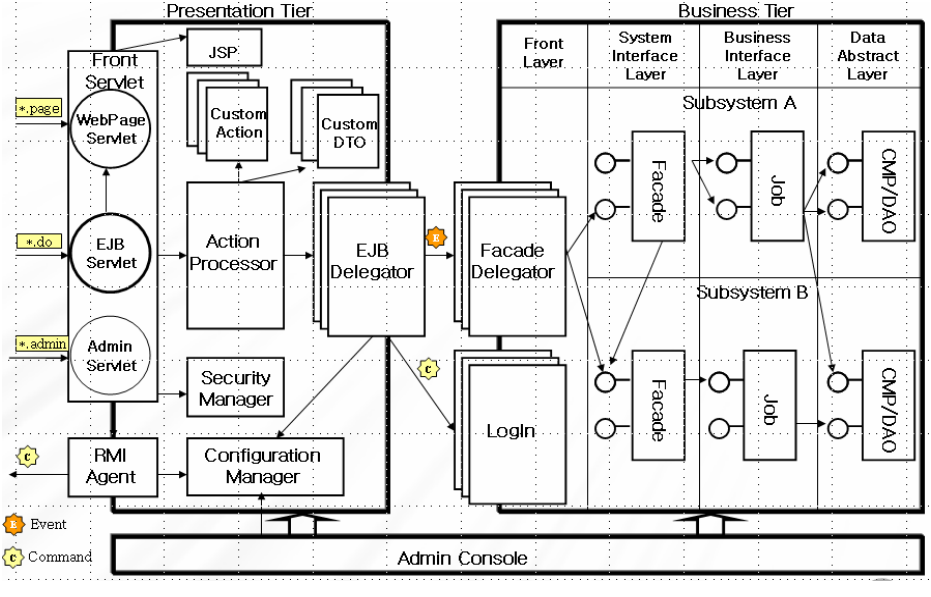


Fig. 1. The Overall Framework Architecture

3.1 Presentation Tier Architecture

In the framework architecture of Figure 1, the Presentation Tier which locates at the front part processes client request. The *FrontServlet* keeps a number of servlets, and proceeds client requests according to the requesting URL pattern, such as **.page, *.do, *.admin, *.login*, etc. As for a **.do* URL pattern, the FrontServlet makes the EJB Servlet process a business logic. As for a **.page* URL pattern, it makes the Page-Servlet process a UI task according to the page construction information, which is organized by the Admin Console. The *Action Processor* is in charge of processing action objects plugged into the framework. The action object is generated from the framework after a developer develops an action class according to the hook method defined by the framework for processing a business logic. The *EJBDelegator* communicates with EJBs for clients' sake. The transmitted communication information and invocation information are set in the Admin Console, because the Admin Console contains all the information and parameters classified. The *Security Manager* performs the tasks of authentication and access control. The *Configuration Manager* processes the system configuration information, which is stored in a XML repository. The *RMI Agent* performs the synchronization task of source code and configuration information among the multiple instances of the same object. It also works for Single-Sign-On.

3.2 Business Logic Tier Architecture

Figure 1 also contains the Business Tier that includes EJB business logics. The *FaçadeDelegator* is the entry point for the requests from the presentation tier. Its major responsibility is to call the façade bean for invoking an EJB component containing an appropriate business logic. The *FaçadeDelegator* also performs general-purpose tasks, such as exception handling and logging, which is independent of a specific subsystem. The *LogIn* is in charge of processing the authentication task for each user. The *FaçadeBean* provides interface that can be used outside of the Business Tier to invoke the *JobBean* that has the actual business logic. The CMP/DAO processes the relational database tasks.

There are two kinds of scheme for a single *FaçadeDelegator* to call multiple façade beans. The first scheme is that the *FaçadeDelegator* uses the EJBMetaData Object and a java reflection. The second scheme uses a callback method. The *FaçadeDelegator* refers the object that includes the callback method and calls the callback method.

3.3 Admin Console Architecture

The Admin Console is a subsystem that is in charge of managing source codes and various parameters according to the concerned task unit. When a client request arrives into the Presentation Tier, the *FrontSevlet* in the Presentation Tier parses the client request and makes the Admin Console process the client request according to the client URL patterns. The client URL has a category information classified by the Admin Console and a URL pattern. For example, when a client requests http://localhost/aaa.bbb.do , which is a **.do* URL pattern, the Admin Console develops the *bbb* node under the *aaa* node in a tree form and provides the related source code to perform the task in that node. In other word, the Admin Console categories

tasks in a tree form, provides the related source code (DTO, Action, EJB code, etc.) to the task, edits them, compiles them, packages them, distributes them, and tests them via a tool. Also the Admin Console manages setting information which is necessary for each task.

The structure of the Admin Console is as follows. The *admin console* includes a policy admin, a development tool, and a configuration and monitoring tool. The *policy admin* manages and specifies policies and rules. The *development tool* helps to implement business logic, compile a source code, package, and deploy. Systematic integration processed by the framework and tools together takes the overall system development process rapidly. The *configuration* manages various environment configuration parameters. The *monitoring* manages all resources related to the framework.

4 Achieving the Flexibility

Among a number of software architectural qualities, our framework provides a good flexibility in many aspects. In this section, we introduce the flexibility aspect of our framework.

4.1 Flexible Control

The overall software architecture follows the MVC pattern in all of the Business Tier, the Presentation Tier, and the Admin Console. Figure 2 shows the control flow among model, view, and controller by using the Admin Console. The *EJB Bean* acts as the

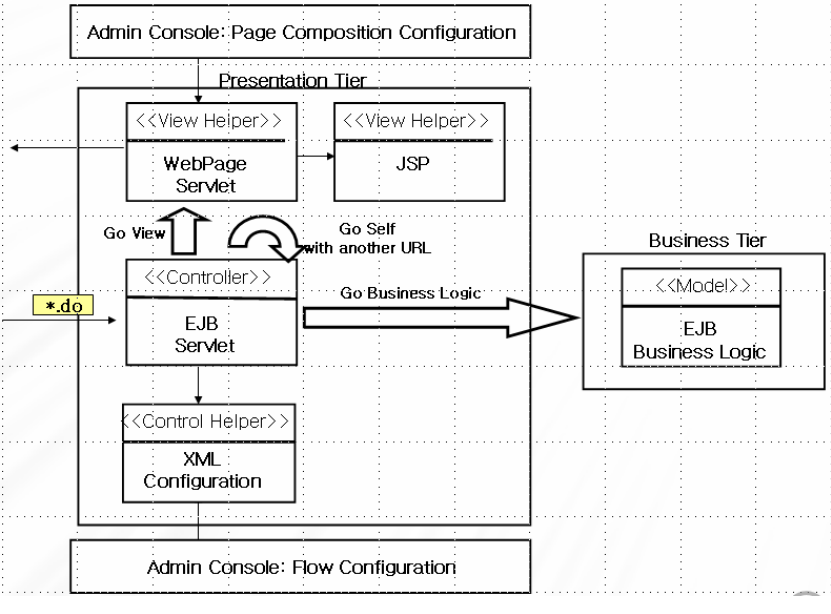


Fig. 2. Flexible Control by Using the Admin Console

model role by containing a business logic. The *FrontServlet* is in charge of flow control for PageServlet to assemble JSP pages according to the Composite Pattern. The Admin Console has all the setting values for this flow control and page decision in the hierarchical form of task nodes and the next URL properties. If the next URL pattern is in **.page* form, the control moves to WebPage Servlet after processing the business logic. If the next URL pattern is in **.do* form, the control moves to the EJB Servlet with a different URL by means of redirection. Since the Admin Console has the setting values and parameters to control the flow, it gives a flexible control flow via a model and a view.

4.2 Flexible Maintenance

In order to maintain the framework, it is necessary to update, compile, package, distribute, and test source codes and setting parameters without stopping the server. Figure 3 shows the structure of flexible maintenance by choosing the related components in the framework. The user may update the setting parameters and source codes through the *Admin Console*. If there is any change in setting of the *Admin Console*, the information in the *XML Repository* is changed accordingly by the *Configuration Manager* and the information in memory is changed by the XML Controller. These changes affect the processing dynamically because the *Front Servlet, Action Processor,* and *EJB Delegator* retrieve the related information from the *Configuration Manager*. Thus, without stopping the server, we are able to test the source codes that should be inserted by developer during operation, which are codes for action, DTO, EJBs, etc. This kind of codes are compiled and inserted into the *Object Pool* with the help of *Pool Manager*.

4.3 Decoupling Between UIs and Business Logics

There are always updates in the UI level. By decoupling between the UIs in the Presentation Tier and business logics in the Business Tier, changes in the UI level do not

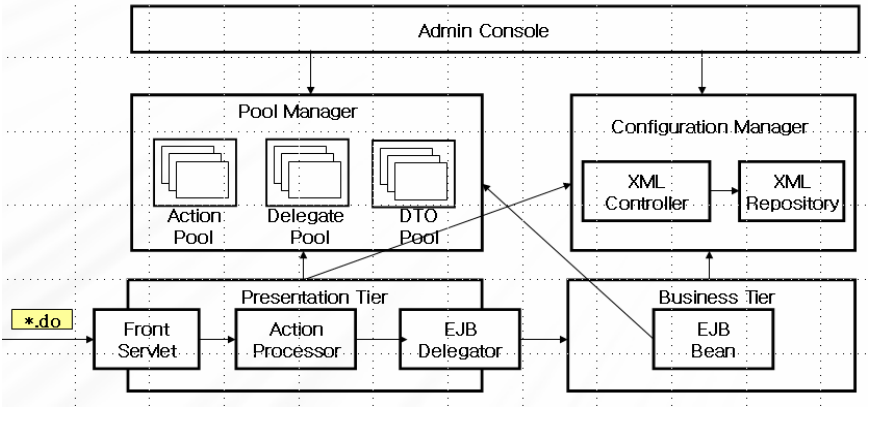


Fig. 3. Flexible Maintenance of Source Codes and Setting Updates

affect the EJBs of business logics. We separate the *Custom DTO*, which contains data to transfer to the Presentation Tier, and the *Domain DTO*, which has the EJB business logic. Any UI changes minimize the dependency on the business logic, yielding the high flexibility.

4.4 Separate Development of the Presentation and the Business Logics

Our framework is designed to have coupling between the Presentation Tier and the Business Tier be minimized so that the development of business logics can be fully separated from the presentation parts. As shown in Figure 4, the Presentation Tier is dependent on the Business Tier only by message invocations from the *EJBDelegator* to the *Façade Delegator*. By hiding all EJBs and providing only the interfaces of *Façade Delegator*, those two tiers are minimally related.

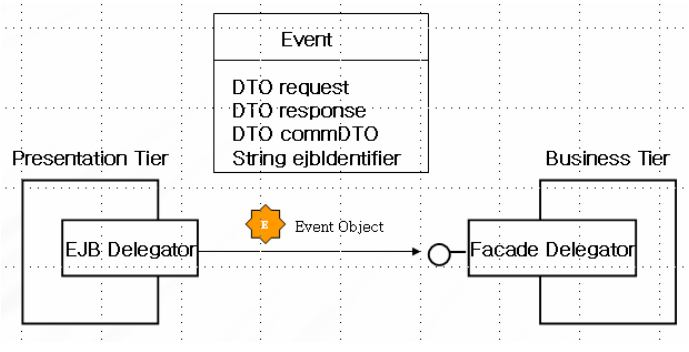


Fig. 4. Minimal Dependency between the Presentation and the Business Tiers

When transmitting data to the *FaçadeDelegator*, the Event Object is used to carry data. In the Event Object, there are a number of contents. The *request DTO* is data to process a business logic in EJB. The *response DTO* is returned after processing in EJB. The *common DTO* has data that is repeatedly needed in every request. The *ejb Identifier* has information showing the description about all methods in the *Façade bean* to be used to invoke through the *Façade Delegator*. By using this *Event Object*, the Presentation Tier that is associated with the EJB client is separately developed and tested. Also when we use a commercial product of Presentation part, it is possible to use it only by adding a Delegater to invoke EJB. It is one of the strong points of our framework.

4.5 Flexible Changes of Business Logic Interfaces

The interface of business logic in the *Job Bean* is hidden to outside of the Business Tier, so that the effect of any change can be minimized. We use the Façade Pattern, and have the interface of the *Façade Bean* be the multi-grained interface as shown in

Figure 1. Thus, any changes in the interface of *Job Bean* do not affect client levels that invoke.

5 Achieving the Extensibility

As the second software architecture quality we achieve, the functional extensibility is presented in this section.

5.1 Request Process Extensibility

Since all Web requests are transmitted to the *FrontServlet,* we extend request functionality by easily adding more processes in the *FrontServlet*. Also, all EJB-related requests are transmitted to the *Façade Delegator*. It is extensible to add more functionality by inserting more processes into the *Façade Delegator*.

5.2 Functional Extensibility in Multiple-Instance Environment of Web Application Servers

Figure 5 shows the multiple-instance environment of Web application servers. Since there are a number of Web application servers that have to have the same image, we need to perform synchronization. In this environment, there are a number of considerable issues, such as single-sign-on, source code synchronization, setting parameter synchronization, node resource monitoring, version management, etc. To extend the functionality of this kind of tasks, we need to add commands whenever there are changes or upgrades. In order to support this functional extensibility, we use Command Pattern in our framework.

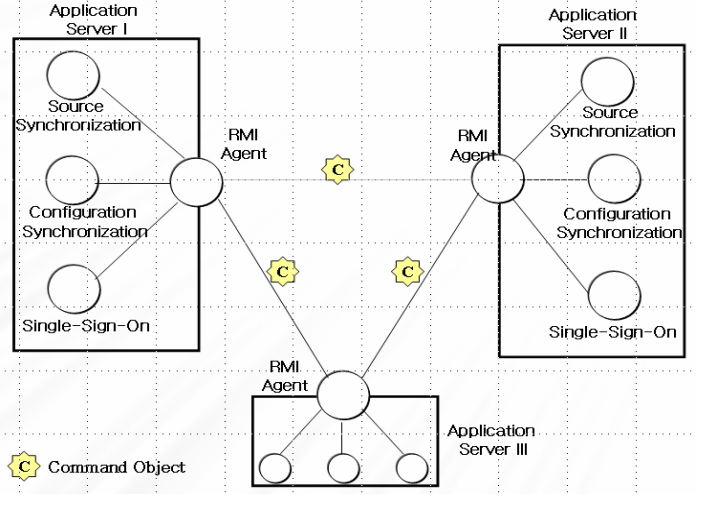


Fig. 5. Functional Extensibility using RMI Agent and Command Pattern

5.3 Extensibility of Login Component

Since the authentication, session information, and data transmitting mechanism are different for every project, it is necessary to extend the login functionality. For this purpose, we employ the Command Pattern to the EJB Login component.

6 Conclusion

The major goal of system integration project is to achieve the software quality and functional satisfaction required by customers fast and effectively in a software development process. To achieve this goal, recent software development projects in business domains employs application frameworks that are partially developed in advance. In this paper, we presented an advanced framework that supports the dynamic re-configuration of software architecture, environment parameter setting, and source codes. We explained how to achieve the flexibility and the functional extensibility in our framework which leads to the high productivity and maintainability. The software architecture of our framework is systemically integrated with its supported development environment so as to achieve high productivity in application development.

References

[1] D.C. Luckham, J. Vera and S. Meldal, Key Concepts in Architecture Definition Language, in Foundations of Component-Based System, Ed. Gary T, Leavens and Murali Sitaraman, Cambridge University Press 2000

[2] J. Kramer and J. Magee, Analyzing dynamic change in distributed software architectures, IEE Proceedings-Software, 145(5), Oct. 1998.

[3] D. Le Metayer, Describing Software Architecture Styles Using Graph Grammars, IEEE Transactions on Software Engineering, 24(7):521-533, July 1998

[4] Sun MVC Information at Sun Micro System Web Site, http://java.sun.com/blueprints/guidelines/designing_enterprise_applications_2e/web-tier/web-tier5.html

[5] Web Application Framework Information at Apache Struts Framework, http://jakarta.apache.org/struts/

[6] SUN, JavaServer Pages. http://java.sun.com/products/jsp/, 2004

[7] Fumihiko Kitayama, Shin-ichi Hirose, Goh Kondoh, Design of a Framework for dynamic Content Adaptation to Web-Enabled Terminals and Enterprise Applications, IEEE, 1999

[8] D. Schwabe and G. Rossi, “An Object-Oriented Approach to Web-Based Application Design,” Theory and Practice of Object Systems (TAPOS), special issue on the Internet, vol. 4, no. 4, 1998, pp. 207-225

[9] M. Fayad, D, Schemidt, and R. Johnson, eds., Building Application Framework, Wiley & Sons, New York, 1999

[10] Sencan Sengul, James W, Gish, James F. Tremlett, Building a Service Provisioning System Using The Enterprise Java Bean Framework, IEEE, 2000

[11] E. Gamma, R. Helm, R Johnson, J. Vlissides, “Design Patterns: Elements of Reusable Object-Oriented Software”, Addison-Wesley, 1995

[12] D.M. Heimbigner and A.L. Wolf. Post-Deployment Configuration Management. In Proceedings of the Sixth International Workshop on Software Configuration Management, number 1167 in Lecture Notes in Computer Science, pages 272–276. Springer-Verlag, 1996.
[13] Apache, Struts Framework, http://jakarta.apache.org/struts/index.html
[14] Spring Framework, http://www.springframework.org/
[15] HiberNate Framework, http://www.hibernate.org

Integrated Meta-model Approach for Reengineering from Legacy into CBD

Eun Sook Cho

Div. Of Computer and Information Science,
Dongduk Women's University,
23-1 Wolgok-dong, Sungbuk-gu, Seoul 136-714 Korea.
escho@dongduk.ac.kr

Abstract. There is an increasing interest in migration legacy systems to new hardware platforms and to new software development paradigms. The reason is that high maintenance costs and lack of documentation. In order to migrate or transform legacy system, various approaches such as screen scrapping, wrapping, semi-development, and re-development, tools, and methodologies are introduced until now. However, architecture or requirements level's transformation is not suggested because most of those approaches focus on code-level transformation or a few model-level transform. In this paper, we suggest a meta-model driven approach applying 3D space concept, which can be applied into architecture and requirement phase. Proposed integrated model drives seamless migration or co-evolution from code to architecture of reverse engineering and from architecture to code of forward engineering.

1 Introduction

Reengineering approaches for transforming legacy system into component-based system are wrapping or screen scrapping-centered techniques [1,2,3]. Those methods mostly transform simply user interface parts of legacy system. Therefore, business logics or data parts of system are still remained in forms of legacy system. These strategies have many limitations in respect to transforming legacy system into component-based or web-based system. The reason is that any transformation of business logics or data, core parts of system, is not realized. Also, most domestic case studies are one-to-one mapping legacy codes into cods of target system. Simple code level transformation cannot reflect new requirements into legacy system as well as cannot improve quality or maintainability of system because architecture or design elements are not transformed. Although new system is constructed by using those approaches, ultimately modernized system holds still the complexity or maintenance problems of legacy system.

Recognizing those problems of existing approaches, in this paper, we propose a new reengineering technique; model-based or model-driven reengineering technique in order to support modernizing legacy system into more systematic and extensible component-based system.

This paper is organized as follows. Section 2 introduces existing reengineering techniques and presents limitations of those techniques. Section 3 defines meta-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 868 – 877, 2005.

models based on 3D space, presents reengineering process based on meta-model, and defines relationships among meta-models. Section 4 describes case studies reengineering based on proposed meta-models into legacy system. We discuss proposed approaches and existing approaches in Section 5. Finally Section 6 presents concluding remarks and future works.

2 Limitations of Existing Researches

Many researches of modernization techniques legacy system have been introduced. One of them is to redevelop whole parts of legacy system completely from scratch. This method has a merit that we can start anew requirement analysis, design, and implementation based on good experience and technologies. On the other hand, it needs enormous costs and time, high risk, much of knowledge and effort required in understanding legacy system [4]. Therefore, many researches recommend incremental transformation approaches as well as apply these in industry. Incremental transformation method first recovery design or architecture models by applying reverse engineering techniques into existing legacy codes. And then, it applies forward engineering based on identified information into legacy system and transforms legacy system into target system. When applying forward engineering into legacy system, we can use two methods selectively. One is to develop new target system by adding new requirements into legacy system. The other is to transform legacy system by reusing components of legacy system [5]. However, existing approaches still have some limitations.

The first, existing reengineering methods produce many models and apply them during reengineering, but connections between models are defined or established. As a result, reengineers are under the burden of producing excessive documents, and there are a number of case studies that produced models are not reflected on system transformation process. This situation brings about low reusability of resources of legacy system.

The second, models are represented differently nesting the same semantics or unnecessary information is identified in each methodology. This situation is the cause that representation level or abstraction degree is not reflected adequately.

In this paper, we apply 3D space concept[6] to reengineering meta-model, establish and present hierarchy of meta-models according to phase or description level in order to overcome these limitations.

The third, there are cases that identified or applied models focus on only one side or aspect of system. This brings about many difficulties in construction of architecture or specifying other aspects of system. In this paper, we present information of various models with meta-models.

3 Meta-model for CBD Reengineering

In this paper, we first define a general reengineering process model based on existing reengineering processes, and suggest meta-models for models identified in each phase

of defined reengineering process. The complex nature of software can be represented as a 3D software product space. We use 3D software product space concept [6].

In the D1 pyramid, meta-models are described. 2D and 3D pyramid deals whole hierarchy. In the D2, information for models applied to reverse engineering and forward engineering phase is depicted as meta-model. As shown in Figure 1, we divide phases into 3 phases; requirement analysis, architecture design, and detailed design and describes meta-models according to each phase. Left side of Figure 1 describes models applied to reverse engineering phase. Right side of Figure 1means models applied to forward engineering or CBD phase. Horseshoe shape of Figure 1is referenced with CORUM model proposed by CMU's SEI [7].

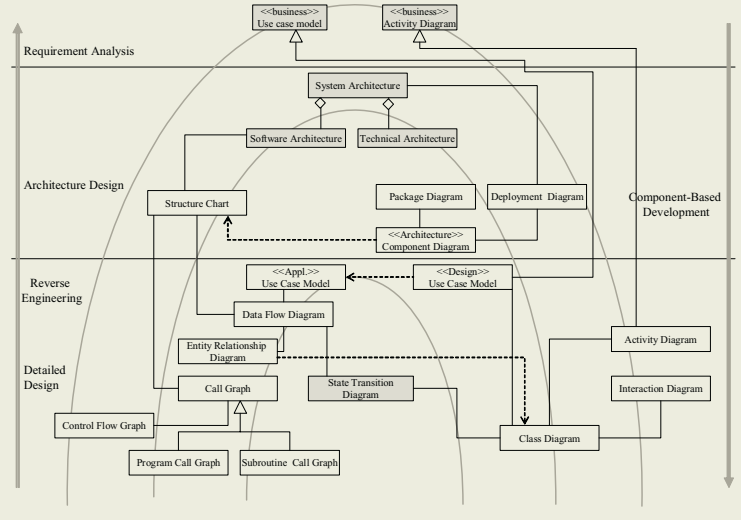


Fig. 1. Architecture of Reengineering Meta-model

3.1 Meta-model of Reverse Engineering

As described in Figure 2, tasks of reverse engineering are divided into 3 layers; detailed design, architecture design, and requirement analysis, and models to describe in each layer and information of models are represented as UML's class diagram notation. For example, a use case model exists in both detailed design layer and requirement analysis layer, but the degree of representation information is different. However, because requirement analysis layer is more abstract than detailed design layer, the granularity of use case of this layer becomes a business process or a business task. In order to distinguish this difference, a use case is represented with <<business>> or <<appl.>> by using stereotype of UML.

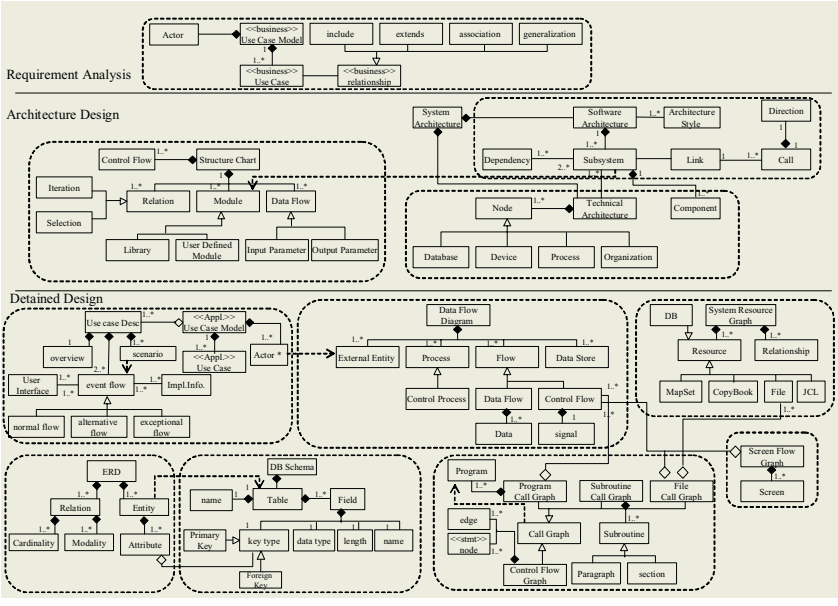


Fig. 2. Meta-model Architecture of Reverse Engineering Phase

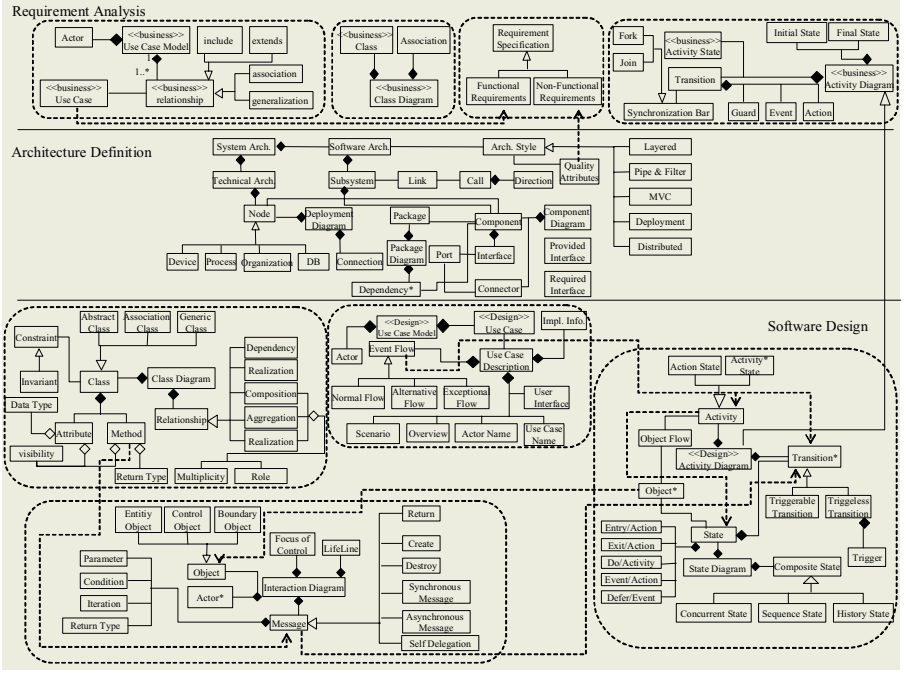


Fig. 3. Meta-model of CBD Phase

3.2 Meta-model of CBD Phase

A CBD phase develops component-based software through the steps of forward engineering. However, the difference between CBD phase and forward engineering is that it does not develop target system in all newly, but it transforms existing system into target system according to new paradigm. If the only functions of legacy system are transformed without adding new requirements, models of requirement specification in requirement analysis phase are not reflected. However, general reengineering strategy transforms legacy system by adding new requirements. In this case, all models suggested in Figure 3 are applied.

As shown in Figure 3, some models are duplicated with models represented in meta-models of reverse engineering phase. For example, models of architecture definition layer are in such that cases. However some models such as component diagram, package diagram, and so on are added in architecture definition of CBD phase.

3.3 Integrated Meta-model of Reverse Engineering and CBD

Figure 4 shows the integrated meta-model of both reverse engineering and CBD phase. The relationships that models extracted in reverse engineering are mapped and applied into models of CBD phase are represented with dependency relationship in Figure 4 We can recognize easily the connection between models of reverse engineering and models of CBD phase. Furthermore we can easily understand the information of models because of the hierarchy of representation level and of phase information.

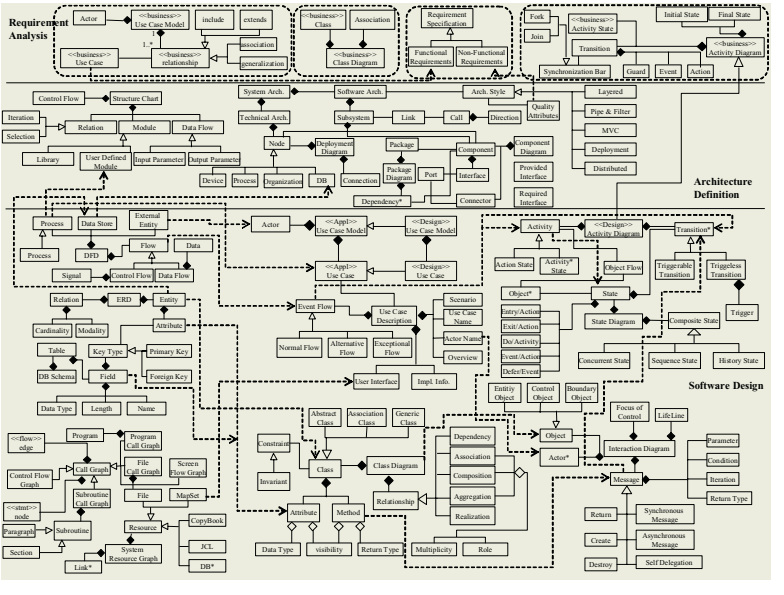


Fig. 4. Reverse + CBD = Integrated Meta-model of Reengineering

3.4 Relationship of Models in 3D Software Space

In this section, we suggest that how to apply proposed models are into which phase by applying 3D space concept, basis of integrated meta-model. Some of proposed models for each dimension are represented by reflecting 3D software space.

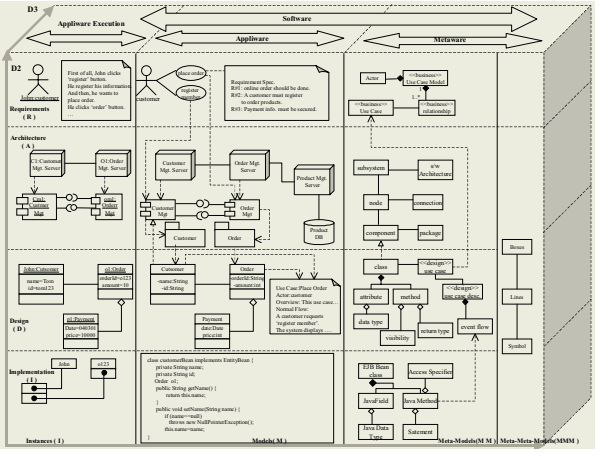


Fig. 5. Integrated Meta-model in 3D Software Space Concept

4 Experiments and Assessment

In this section, we present a case study of applying our approach into legacy COBOL system; capital business management system. This system is a COBOL-based system which being operated in IBM AS/400 Machine.

Figure 6 is a part of control flow graph extracted from COBOL source code. A control flow graph is a model representing control flow information for each program or each subroutine. This example depicts the control flow of a subroutine.

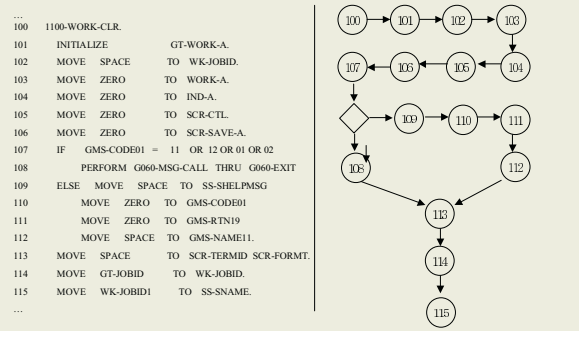


Fig. 6. Control Flow Graph

Figure 8 is an example of class diagram into which ERD(Entity Relationship Diagram) of Figure 7 is transformed in CBD phase. Entity of ERD is mapped into class, attributes of entity are mapped into attributes of class, and cardinality or modality of ERD also is transformed into multiplicity of class diagram. This example shows only parts transforming ERD into class diagram. However, if there are new requirements to add in CBD phase, new entities are added in ERD and class diagram.

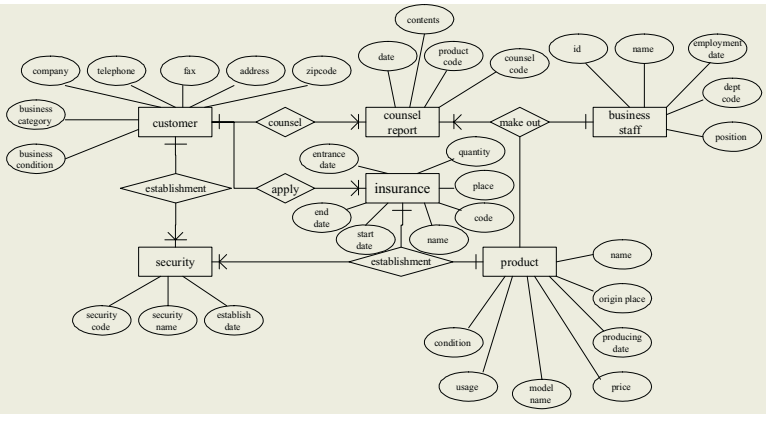


Fig. 7. An Example of Entity Relationship Diagram

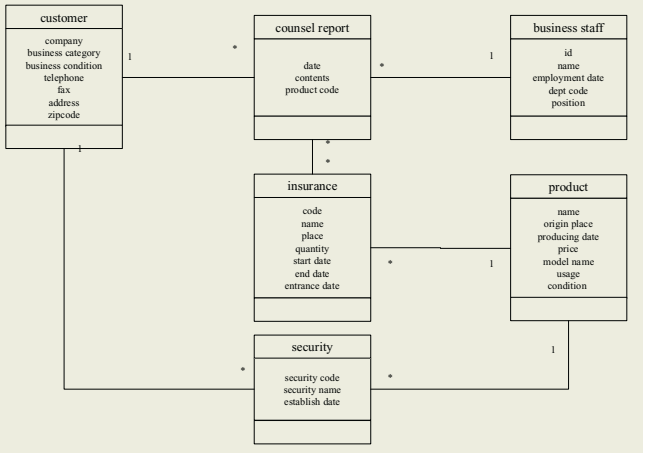


Fig. 8. An Example of Class Diagram

Figure 9 depicts a component diagram for target system to transform legacy system into component-based system based on identified components.

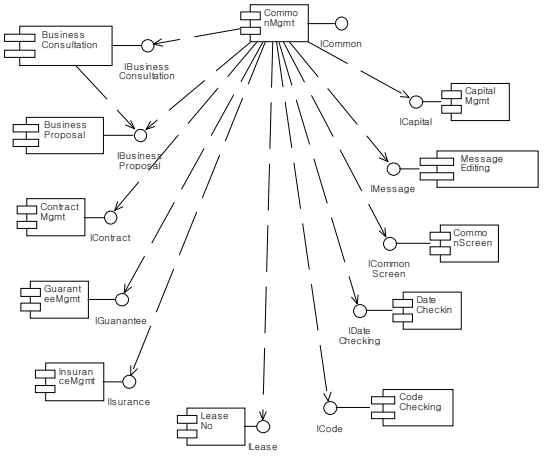


Fig. 9. An Example of Component Diagram

5 Assessment

In this section, we describe the feasibility of our approach; meta-model based reengineering approach by comparing proposed techniques with existing approaches in aspects of both quality attributes [8] and methodological view.

As shown in Figure 10, first of all, there is all the difference between code-based approach and model-based approach in the size of source code. Of course, this case does not reflect new requirements in legacy system. The complexity of system also is lower in model-based approach than in code-based approach.

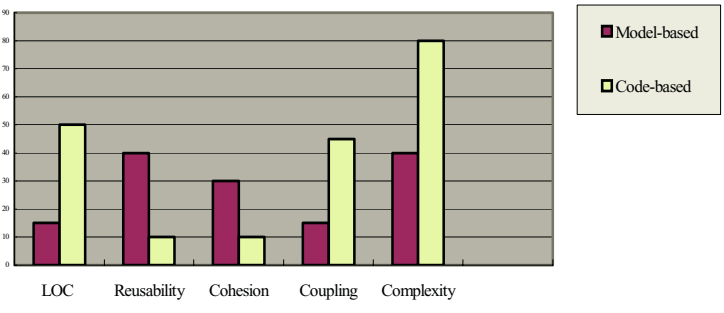


Fig. 10. Comparison Results of Quality Attributes

As depicted in Table 1, existing methodologies as like MORALE[9], Butterfly[10], or Sneed's methodology does not apply meta-model based reengineering approach. However, this paper proposes a meta-model based reengineering approach. This approach supports the integrity between models as well as understandability and

transformation of legacy system during reengineering process. Beside existing methodologies support only some models applied into specific phase, our approach supports models applied into whole phases; reverse engineering and forward engineering phase of reengineering process. Because proposed approach provides relationships between models, whole part of reengineering process can be supported.

Table 1. Comparison Results through Methodological Aspects

	MORALE[9]	Butterfly[10]	Sneed's Approach[3]	Our Approach
Reengineering Approach	Code-based	Data-based	Model-based	Meta-Model-based
Degree of Model	UI Model	Data Model	Reverse Engineering Model	Reverse/Forward Engineering Model
Architecture Information	System Architecture	Data Architecture	Software Architecture	System Architecture Software Architecture
Hierarchy of Model	-	-	Partially	Hierarchical Model
Abstraction Level	Low	Medium	Medium	High

6 Conclusion Remarks

Almost reengineering methodologies use its arbitrary models although different models provide the same semantics and representation degree. This situation raises confusion of models for reengineers as well as results in many difficulties in understanding relationships between models. Recognizing those problems of existing approaches, in this paper, we classify models with hierarchy and represent models based on meta-model by applying 3D software space concept. Therefore, the relationship between phase and model, the relationship between models, and the degree of representation of model are classified and suggested. We expect that this approach allows reengineers to select needed models according to reengineering strategy as well as to approach easily model mapping from legacy system to CBD system. Our future work is to specify meta-models with formal language and develop consistency checking technique and automation mechanism.

References

1. William Ulrich, *Legacy Systems : Transformation Strategies*, Prentice Hall, 2002.
2. Bennett, K, "Legacy Systems: Coping With Success", *IEEE Software*, Vol. 12, No. 1, pp. 19-23, 1995.
3. Sneed, H., Majnar, R., "A Case Study in Software Wrapping", *Int. Conf. in Software Maintenance,* Nov 16-20, Bethesda, Maryland, IEEE Computer Society Press, pp. 86-93, 1998.

4. Cimitile, A., De Lucia, A., Di Lucca, G., "An Experiment in Identifying Persistent Objects in Large Systems", *International Conference on Software Maintenance,* Nov 16-20, Bethesda, Maryland, *IEEE Computer Society Press*, pp. 122-130, 1998.
5. De Lucia, A., Di Lucca, G., Fasolino, A., Guerra, P., Petruzzelli, S., "Migrating Legacy Systems Towards Object-Oriented Platforms", *Int. Conf. in Software Maintenance,* pp. 122-129, 1997.
6. Jean-Marie Favre, "Meta-Model and Model Co-evolution within the 3D Software Space", ELISA'2003: International Workshop on Evolution of Large-scale Industrial Software Applications, Tuesday, 23 September 2003, Amsterdam, The Netherlands.
7. Nelson Weiderman, Dennis Smith, Scott Tilley, "Approaches to Legacy System Evolution", *CMU/SEI-97-TR-014*.
8. S.R. Chidamber and C.F. Kemerer, "A Metric Suite for Object-Oriented Design", IEEE Transactions on Software.
9. Abowd G. Goel A. Jerding D.F., McCracken M., Moore M., Murdock J.W., Potts C., Rugaber S., Wills L., "MORALE. Mission ORiented Architectural Legacy Evolution" *International Conference on Software Maintenance*, pp: 150 -159, 1997.
10. Bing Wu, et. al., "The Butterfly Methodology: A Gateway-free Approach for Migrating Legacy Information Systems" Proceedings of the 3rd IEEE Conference on Engineering of Complex Computer Systems(ICECCS97), Villa Olmo, Como, Italy, September 8-12, pp.200-205, 1997.

Behavior Modeling Technique Based on EFSM for Interoperability Testing

Hye-Min Noh, Ji-Hyun Lee, Cheol-Jung Yoo, and Ok-Bae Chang

Dept. of Computer Science, Chonbuk National University, 664-14 1ga, Duckjin-Dong, Duckjin-Gu, Jeonju, South Korea
{hmino, jhlee, cjyoo, okjang}@chonbuk.ac.kr

Abstract. With the rapid growth of network technology, two or more products from different vendors are integrated and interact with each other to perform a certain function in the latest systems. However, there are cases where products from different vendors or even from the same vendor often do not interoperate properly. Thus, interoperability testing is considered as an essential aspect for correctness of integrated systems. Interoperability testing tests the ability of software and hardware on different machines from different vendors to share data. Existing researches about interoperability testing are usually focused on an optimal test scenario generation applying graph and automata theory. Most of these researches model communication system behavior using EFSM(Extended Finite State Machines) and use EFSM as an input of test scenario generation algorithm. There are many studies on systematic and optimal test cast generation algorithms using EFSM, but in these researches, the study for generating EFSM model, which is a foundation of test scenario generation, is not sufficient. This paper proposes an EFSM generation technique, which is a foundation of test scenario generation for more complete interoperability testing based on use case specification. The generated EFSM through the proposed technique in this paper can be used as an input of EFSM-based test scenario generation algorithm proposed in other studies.

1 Introduction

With the rapid growth of network technology, two or more products from different vendors are integrated and interact with each other to perform a certain function in the latest systems. Also, systems or applications on mobile network tend to interact with systems or applications of the existing network. Therefore, interoperability testing is considered as an essential aspect of correctness of integrated systems. The existing researches about interoperability testing proposed the test scenario generation technique mainly for optimal testing. We cannot do complete testing if we use a wrong test scenario for interoperability testing. Due to the use of a wrong test, we cannot discover some of the errors within the system. Also, the efficiency of test will worsen because of redundant test cases. These problems can be summarized as two kinds of principles - completeness and irredundancy.

Completeness is that all the interoperations between the two systems are tested. Irredundancy is that all the redundant tests are removed to minimize the total number of

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 878 – 885, 2005.

tests. The tester must consider the completeness and irredundancy principles when he or she tests an integrated system.

The existing researches about interoperability testing are usually focused on an optimal test scenario generation applying graph and automata theory and localized at a specific domain[1][2][3]. Most of these researches model communication system behavior using EFSM(Extended Finite State Machines) and use EFSM as an input of test scenario generation algorithm.

For generation of a test scenario, which can fulfill the principle of completeness, we must identify states and change of states, because a state change of the system according to interoperations becomes a test case in interoperability testing. Use case specification can be used to identify the information about states, because use case specification is containing the information about system behaviors. EFSM describes state and state transitions of a system using behaviors of the system. EFSM is a foundation of test scenario generation algorithm. There are many studies on systematic and optimal test cast generation algorithms using EFSM. However, in these researches, the study for generating EFSM model, which is a foundation of test scenario generation, is not sufficient. This paper proposes an EFSM generation technique, which is a foundation of test scenario generation for more complete interoperability testing based on use case specification.

This paper is structured as follows. Section 2 opens with a presentation of interoperability testing and introduces related researches. Section 3 proposes a behavior modeling technique based on EFSM. Section 4 shows an example as an application of the proposed behavior modeling technique. Section 5 presents conclusions and future works.

2 Interoperability Testing

Interoperability testing is different than conformance testing. Interoperability testing tests the ability of software and hardware on different machines from different vendors to share data. The behavior of an interoperating system must be predicted from the specifications of each system [4]. Most of these researches model communication system behavior using EFSM(Extended Finite State Machines) and use EFSM as an input of test scenario generation algorithm[5].

There are many studies on test cast generation technique based on formal specification[6][7][8]. Formal specification is reliable description for system behavior. Formal specifications play an important role in software testing because it describes system behavior in detail. Therefore, there are many researches about automated test case generation technique from formal specification. However, all of the system is not designed using formal specification and formal specification is very difficult. Therefore, formal specification cannot be applied commonly.

Because of these reasons, this paper proposes an EFSM generation technique, which is a foundation of test scenario generation for more complete interoperability testing based on use case specification, which is not a formal specification.

3 Behavior Modeling Technique Based on EFSM

Behavior modeling technique based on EFSM is divided into two parts. The first is the state identification process based on use case specification. The second is a description of the state transition using identified states and writing the EFSM specification. Fig. 1 shows the states identification process with input and output.

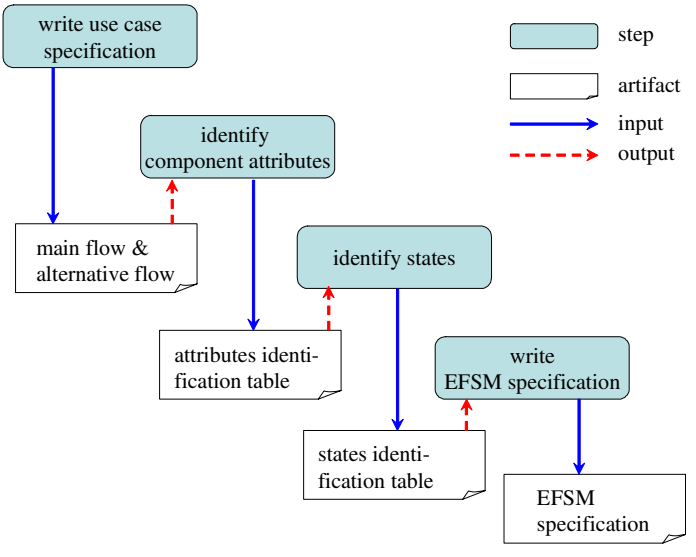


Fig. 1. State identification process

3.1 States Identification Process

State identification process is composed of use case specification analysis, component and attribute identification, and state identification using identified component attribute. After state identification process, we write the EFSM using identified information. Fig. 2 shows the process, which is behavior modeling based on EFSM. The use case specification is used as an input of state identification process. Each use case has to be described in a flow of events document. This textual document defines what the system has to do when the actor activates a use case. The structure of a use case specification can vary, but a typical description would contain

- ***Brief description***
- ***Actors*** involved
- ***Preconditions*** necessary for the use case to start
- Detailed description of flow of events that includes :
    - ***Main flow*** of events, that can be broken down to show :
        - Subflows of events (subflows can be further divided into smaller subflow improve document readability)
    - ***Alternative*** flows to define exceptional situations
- ***Postconditions*** that define the state of the system after the use case ends

```
Input : Main flow & alternate flow statements in use case specification
Output : States

Step 1 : Writing component attributes identification table
for all statements
{
    identify component;
    write the identified component name in corresponding column;
}
for all statements
{
    if(statement includes conditional statement)
        describe the corresponding condition in alternative thread column;
    identify the component attributes;
    identify the possible attribute value;
    write component attribute and attribute value in attribute column;
    if (attribute relate to alternative thread)
        use 'A' as prefix for attribute name;
    define the state name considering the component attribute value change;
    write defined state name and corresponding attribute value in component state
column;
}

step 2 writing state identification table using attribute identification table
for all identified components from Step 1
    write components in Comp. Name. column
for all identified attributes from Step 1
{
    if(attribute relate to alternative thread)
        continue;
    write attribute in attribute column;

}
refer to the behavior in use case statements and write possible attribute value in at-
tribute values column;
regard a column of attribute values as a state  and name the state;
```

Fig. 2. State identification algorithm

We use the detailed description (main flow and alternative flow) of use case specification for state identification. Three kinds of tables are written by state identification process. The composition of each table is as follows.

Component composition table

- component name: identified component name from the use case specification(main flow and alternative flow)
- component description: brief description for component

Component attributes identification table

- statement no.: the number for use case specification sentence
- component: the component to be including in the sentence

- behaviors: behavior of component
- alternative thread: the condition to be including in the sentence
- attributes: the attributes of component
- component state: state name of component when attribute value is change

States identification table

- component name: identified component name from the use case specification
- attribute: component attribute
- attribute value: the value of attribute
- state no.: identified state number

In the next section, we present the example of proposed technique

3.2 EFSM Specification

EFSM(Extended Finite State Machines) is quintuple as follows

$$M=(I,\ O,\ S,\ \vec{x}\ ,\ T)$$

where I, O, S, and T are finite sets input symbols, output symbols, states, variables, and transitions . Each transition t in the set T is a 6-tuple

$$t=(s, q, a, o, P, A)$$

where s, q, a and o are the start state, end state, input, and output. P is a predicate on the current variable values and A defines an action on variable values.

4 Application by Example of Behavior Modeling Technique Based on EFSM

This section shows an application by example of behavior modeling technique based on EFSM proposed in section 3. This example is to test interoperability between SP

Table 1. Use case specification

Use case	request for current location information
Brief description	This use case manages the records about request for current location information
Actors	SP
Preconditions	...
Main flow	1. SP create the XML document for getting the current location 2. SP send the message with SOAP form to OMC 3. OMC receive and analysis the message from SP 4. OMC check the service authentication 5. OMC return the error message if authentication is failed 6. OMC request the service processing to LGC through TCP/IP
Alternative Flows	...
Postconditions	...

and OMC. SP sends the message to OMC for getting the current location information. OMC receives the message and processes the request of SP. Table 1 shows the use case specification for request for current location information.

4.1 Components and Attributes Identification

We can write the component composition table and the component attributes identification table by step 1 of the states identification process. Table 2 and Table 3 show the component composition table and the component attributes identification table. Component composition table describes identified components. These components represent the subject of behavior. Component attributes identification table includes component behavior and component state name. Component name is defined according to the change of component attributes value.

Table 2. Component composition table use case specification

Component Name	Component Description
SP	LBS Application Service Provider
OMC	Platform Management System
LGC	Location Acquisition Gateway System

Table 3. Component attribute identification table

Stat. No.	Comp.	Behaviors	Alternative Thread	attributes	component state
1	SP	create XML(SP)		idle_SP{T, F} createXML_SP{T, F}	T : idle_SP, F : T : createXML_SP, F : XMLCreatFail_SP
2	SP	send message to OMC(SP)		sendMsg_SP{T, F}	T : sendMsg_SP, F : idle_SP
3	OMC	receive message from SP(OMC)		idle_OMC{T, F} receiveMsgFromSP_OMC{T, F}	T : idle_OMC, F : T : receiveMsgFromSP_OMC, F : idle_OMC
4	OMC	authenticate service(OMC)		AIauthServSucc_OMC{T, F}	
5	OMC	return error(OMC)	Whether authentication is success?	AIauthServSucc_OMC{T, F} retErr_OMC{T, F}	T : retErr_OMC(A1), F : requestServToLGC_OMC(A1)
6	OMC LGC	request service to LGC(OMC)		AIauthServSucc_OMC{T, F} reqServToLGC_OMC{T, F} idle_LGC{T, F} receiveReqFromOMC_LGC{T, F}	T : reqServToLGC_OMC(A1), F : T : idle_LGC, F : T : receiveReqFromOMC_LGC, F : idle_LGC

4.2 States Identification

We can writing the states identification table based on component attribute identification table by step 2 of states identification process. Table 4 shows the states identification table. States identification table represents the states name. The states name is defined according to the change of attribute value.

Table 4. States identification table

Comp.Name.	Attributes	Attribute Values				
SP	Idle_SP	T	F	F	F	F
	CreateXML_SP		T	T	T	T
	SendMsg_SP		F	T	T	T
OMC	Idle_OMC	T	T	F	F	T
	ReceiveMsgFromSP_OMC			T	T	T
	ReqServToLGC_OMC			F	A1=T	A1=F
					T	F
	retErr_OMC				F	T
State No.		S1	S2	S3	S4	S5

```
Initial state: S1(Idle_SPIdle_OMCIdle_LGCIdle_MODBIdle_DB)
Initial value of variables:
  ……
//Format
//From-State
//{Input}/{Output}/{Predicates}/{Actions}/{Color} Next-State
//
transition {
Idle_SPIdle_OMCIdle_LGCIdle_MODBIdle_Carrier
   {clientReq}/{}/{}/{Idle_SP=false, CreateXML_SP=true}/{White}
   S2(CreateXML_SP)
CreateXML_SP
   {XML}/{SOAPMsg}/{}/{sendMsg_SP=true; Idle_OMC=false; Re-
   ceiveMsgFromSP_OMC=true}/{Black}
   S3(SendMsg_SPReceiveMsgFromSP_OMC)
SendMsg_SPReceiveMsgFromSP_OMC
   SOAPMsg}/{reqServ}/{A1}/{ReqServToLGC_OMC=true; Idle_LGC=false;
   ReceiveReqFromOMC_LGC=true}/{Black}
   S4(ReqServToLGC_OMCReceiveReqFromOMC_LGC
    ……
}
```

Fig. 3. EFSM specification

4.3 EFSM Specification

Fig. 3 represents the EFSM specification, which describes transitions among the states. S1, S2 and S3 are states that are identified by the states identification process.

This EFSM specification describes what happened to the state transition during SP send XML documents to OMC for getting the current location information.

5 Conclusions and Future Works

With the rapid growth of network technology, two or more products from different vendors are integrated and interact with each other. Therefore, interoperability tests are considered as an essential aspect of correctness of integrated systems. However, the existing researches about interoperability testing are usually focused on an optimal test scenario generation applying graph and automata theory and localized at a specific domain. Also, the study for generating EFSM model, which is a foundation of test scenario generation, is not sufficient. Therefore, this paper proposes an EFSM generation technique, which is a foundation of test scenario generation for more complete interoperability testing based on use case specification.

The generated EFSM through the proposed technique in this paper can be used as an input of EFSM-based test scenario generation algorithm proposed in other studies. In future works, we will study on efficient algorithms for generating test scenarios based on EFSM, which is proposed in this paper. Also, we will research about automatic interoperability test case generation techniques and tools.

References

1. J. Gadre, Rohre C, C. Summers, and S. Symington. "A COS study of OSI interoperability," Computer Standards and Interface, vol. 9(3), pp. 217-237, 1990
2. G. S. Vermeer and H. Blik., "Interoperability testing: Basis for the acceptance of communicating systems," VI(C-19), Elsevier Science Publisher B.V., 1994
3. Pei Hsia and David Kung, "software requirements and acceptance testing," Annals of Software Engineering, vol. 3, pp. 291-317, 1997
4. Sungwon Kang, "Relating interoperability testing with conformance testing," Global Telecommunications Conference, vol. 6, pp. 3768-3773, 1998
5. Griffeth, N., Hao, R., Lee, D., Sinha, R.K., "Interoperability testing of VoIP systems," Global Telecommunications Conference, vol. 3, pp. 1565-1570, 2000
6. Miao Huaikou, Liu Ling, "A test class framework for generating test cases from Z specifications," Engineering of Complex Computer Systems, pp. 164-171, 2000
7. Cunning, S.J., Rozenblit, J.W., "Automatic test case generation from requirements specifications for real-time embedded systems," IEEE SMC '99 Conference Proceedings, vol. 5, pp. 784-789, 1999
8. Cunning, S.J., Rozenbiit, J.W., "Test scenario generation from a structured requirements specification," Engineering of Computer-Based Systems, pp. 166 - 172, 1999

Automatic Connector Creation for Component Assembly

Jung-Soo Han[1], Gui-Jung Kim[2], and Young-Jae Song[3]

[1] Division of Information and Communication, Cheonan University,
115, Anseo-Dong, Cheonan, Chungnam, Republic of Korea
jshan@cheonan.ac.kr
[2] Department of Computer Engineering, Konyang University,
26, Nae-Dong, Nonsan, Chungnam, Republic of Korea
gjkim@konyang.ac.kr
[3] Department of Computer Engineering, Kyunghee University,
Kihung-up YongIn-Shi, KyungKi-Do, Republic of Korea
yjsong@khu.ac.kr

Abstract. In this paper, we present the automatic connector creation technique which connects and assembles components without method call or component modification. The connector is automatically created by the definition of component port, specification and architecture. Through the created connector, components are assembled.

1 Introduction

Recently, software development process moves from the in-house type development to new paradigm such as standardized component, outsourcing, commercials off-the-shelf (COTS) components and so forth. That is, the developed final system does not progress toward a fixed system but toward open or flexible system that is developed on the basis of component and can be integrated into other system and upgraded by its own. In this system, since the system is upgraded in real time, it affects the system configuration. Therefore, this paradigm increases the efficiency in development but involves some risk that decreases the consistency or reliability of system configuration [1, 2].

Component configuration management (CCM) fulfills a function that increases the reliability by controlling the consistency among constituting elements within the system. When new component is added or substituted in the developed package on the basis of components, two problems should be solved. The first problem is generated when the existing component, which is connected with other package, is substituted. That is, when the new version component is substituted, the relation between changes and components is uncertain. The second problem, the dynamic behavior for the system under real time environment, is more serious. That is, when a component is substituted in real time, one package is running but the other package connected with previous component could not be running. The solution for these problems is component configuration management (CCM).

Also, for the fast development and maintenance of the complicated application program, component based development (CBD) technique that develops the compo-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 886 – 895, 2005.

nent as a software unit is spreading. And the research about component assembly for this is in progress [3]. Not only static component assembly but also dynamic component assembly is being studied [4]. Component assembly is performed through the specified interface composition and is deployed automatically by the specification context.

Therefore, in this paper, we present the automatic connector creation, which connects and assembles components without method call or component modification. The connector is automatically created by the definition of component port, specification and architecture structure.

2 CBD LifeCycle

In software engineering, a lot of techniques suggest software development life cycle with similar way and is used in the same way. But CBSE includes the component development and the system development using the component. The difference with existing techniques is the reusability. In order to develop and reuse the component in a lot of applications, specification, understandability, universality, adaptability and redeployment must be organized well. Two steps must be considered at design stage in component development. The first thing is the system architecture specification about the component function and the relation among components. This step makes us recognize the logical aspect of the system. The second thing is the system architecture specification constituting the component physically. Figure 1 shows the CBD life cycle compared with waterfall model. CBD life cycle means the process that searches the component, selects the correct component, and applies that to the existing system, tests and re-deploys it [5].

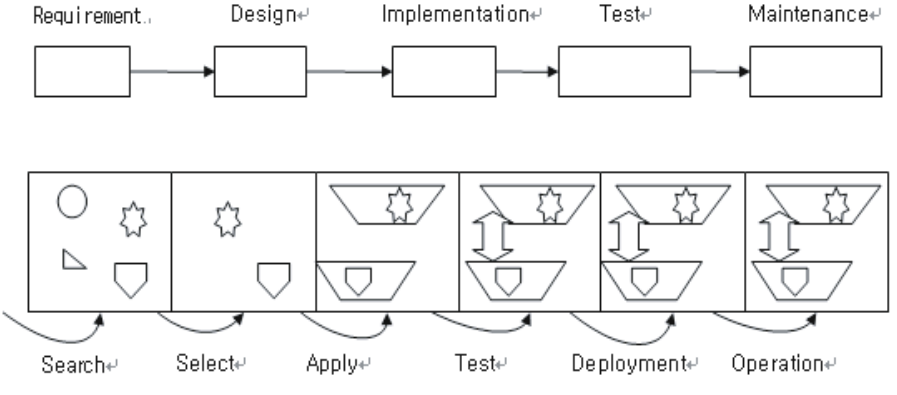


Fig. 1. Component life cycle compared with waterfall model

Figure 2 represents configuration management in the development stage and run time. In the development stage, a library is created from source code and the component is generated by assembling the library (dll file). And then the component-based

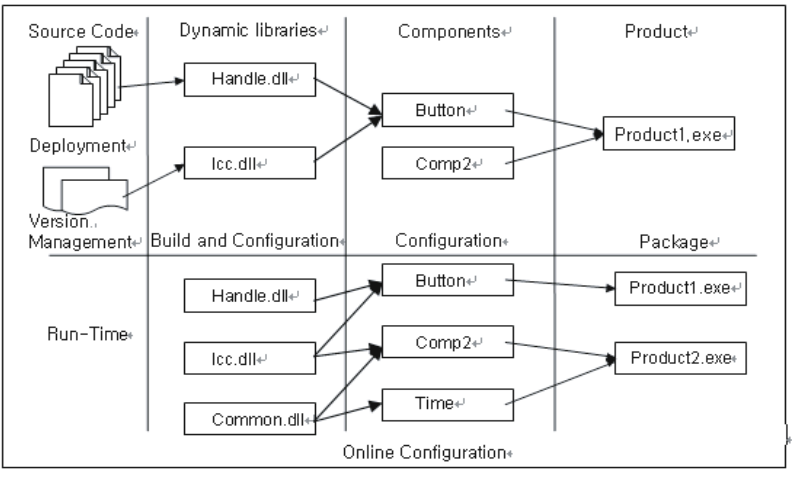


Fig. 2. Configuration management in component based product life cycle

product is implemented as the set of components. In this process, the source code of the first development stage is managed with the version information of the code. 'Build' progresses, connecting with the source code, by various CM tools (make and configuration tool) and then, finally, the product is created by assembling components. At this stage, since we control the source code and construct a whole system from this code, we can control the system configuration management. But, if a component is added or substituted, this control is impossible because we just know the function of the component partially. But, if the version control of the component is possible, we can control the version and configuration management at the same time and solve the uncertain relationship among components.

3 Library System Model

In this paper, we represent an automated library system model as an example to create a connector. Library consists of two components, 'Stock' and 'Library'. 'Stock' component has three operations, getTotalStock(), bookInput() and bookOutput(). 'Library' component has totalResult() operation and bookReturn() and bookBorrow() operation required by totalResult() operation. In order to connect 'Stock' component with 'Library' component, the connector uses bookReturn() and bookBorrow() operation required by 'Library' and each operation uses bookInput() and bookOutput() operation of 'Stock' component. Figure 3 shows Library system model. Here, the connector connecting 'Stock' component with 'Library' component is created automatically through component definition and specification and architecture structure.

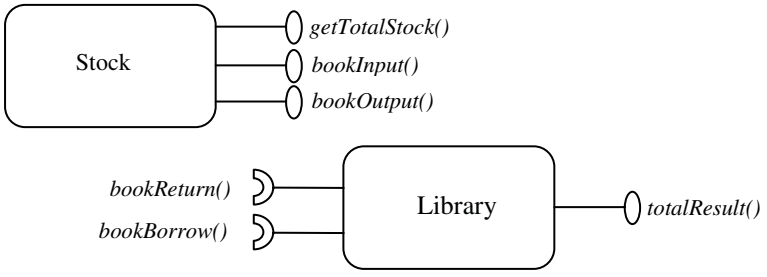


Fig. 3. Library system

4 Connector Definition and Specification

4.1 Definition and Specification

Components and ports. When a component is specified in UML class, the component name is expressed with stereotype <<component>>. The component consists of ports. Port (Operation) is a connecting part defined by a category. The basic type of ports is a operation port defined to require the provided or required operation of the component. Operation port is specified with the stereotype <<operationPort>>. Ports connect components by stereotype <<provided>> or <<required>>.

Connectors and assembly. Component assembly specification is to define the relation among components. The concept of 'required' and 'provided' port is to connect two ports. It can be used for the adaptation of ports without compatibility. 'Provided' and 'required' port specify the connector with stereotype << connect_type of ports>> as a UML dependency relation. Operation port connecting each component is specified with stereotype <<connectOperation>>.

Ports adaptation. Two ports can use connector. Library model requires two port composition and represents {adapt} value with a tag at the connecting time. Adaptation is processed in the mapping stage [6].

In figure 4, 'Stock' component describes getTotalStock(), bookInput() and bookOutput as 'provided', and 'Library' component describes totalResult() as 'provided' and bookReturn() and bookBorrow() as 'required'. With stereo type <<connectOperation>>, the connection between bookInput() of 'Stock' component and bookReturn() of 'Library' component is represented. We attached a tag {adapt} to apply these two operations. Also, with stereo type <<connectOperation>>, the connection between bookOutput() of 'Stock' component and bookBorrow() of 'Library' component is represented. We also attached a tag {adapt} to apply these two operations. Table 1 shows this structure of figure 4.

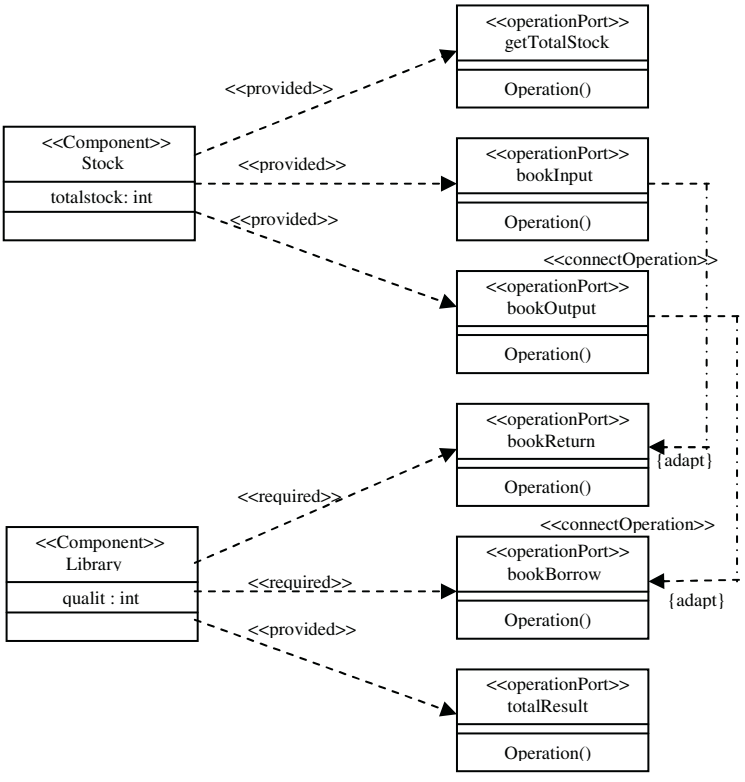


Fig. 4. Component specification, port and connector

Table 1. Component specification, port and connector

Component	perationPort	connect_type_of_ports	connectOperation
Stock	getTotalStock	provided	bookInput-book-Return
	bookInput	provided	
	bookOutput	provided	
	bookReturn	required	
Library	bookBorrow	required	bookOutput-bookBorrow
	totalReault	provided	

4.2 Architecture Structure

Architecture expressing structural rules can be added into an extended model. With Meta level constraint of OCL (object constraint language) meta-level, architecture structure is specified. For example, in order to represent that the operation connector (dependent state stereo type <<connectOperation>>) only expresses the combination of operation ports (class stereo type<operationPort>>), we add the rule that does not allow changing the UML dependency relation in OCL meta-model [6]. Figure 5

shows architecture structure. For each component connected by <<connectOperation>> defined with stereo type, 'supplier' that is the stereo type <<operationPort>> of 'Stock' component and 'client' that is the stereo type <<operationPort>> of 'Library' component are used. Client component uses the object of supplier by referencing the object of supplier component.

```
context Dependency
inv self.isStereoptyped("connectOperation") implies (( self.supplier → forAll(S:ModelElement
|  S.isStereotyped("operationPort"))  and  (self.client  →  forAll(C:ModelElement  |
C:.isStereotyped("operationPort"))
```

Fig. 5. OCL for Library system model

4.3 Automatic Connector Creation

Automatic connector creation is specified as figure 6 through the context dependency relation of figure 5. In the example of Library model, we first search the stereo type with {adapt} value. And then automatic connector creation rule specification is drawn up through the context expressed in figure 5. Using the model element of 'supplier' or the operation with stereotype <<operationPort>> and the model element of user or the operation of stereotype <<connectOperation>> including the operation with stereotype <<operationPort>>, it was represented as figure 6.

```
bookInput - bookReturn {adapt}
  supplier : Stock Stock.bookInput
  client : Library Library.bookReturn
bookOutput - bookBorrow {adapt}
  supplier : Stock Stock.bookOutput
  client : Library Library.bookBorrow
```

Fig. 6. Rule specification for automatic connector creation

Figure 7 is the interface storing the component information in the database (DB). The component name and the function of that component are inserted. Port type is inserted in Connect_type_of_ports in order to connect with other components. If user wants to assemble the component just stored, he can assemble the component by searching components that can be composed with this component and selecting the component. Table 2 is the component information in figure 7 stored in DB. Figure 8 represents 'Connnection' part. If user wants to assemble components, he prescribes the connecting state of components through the interface in figure 8 and clicks 'ok' button

and then it is stored as XML document. The name and function of supplied component is inserted in 'supplier' part, and the name and function of provided component is inserted in 'client' part. Program Code1 shows the XML document prescribed by the interface in figure 8.

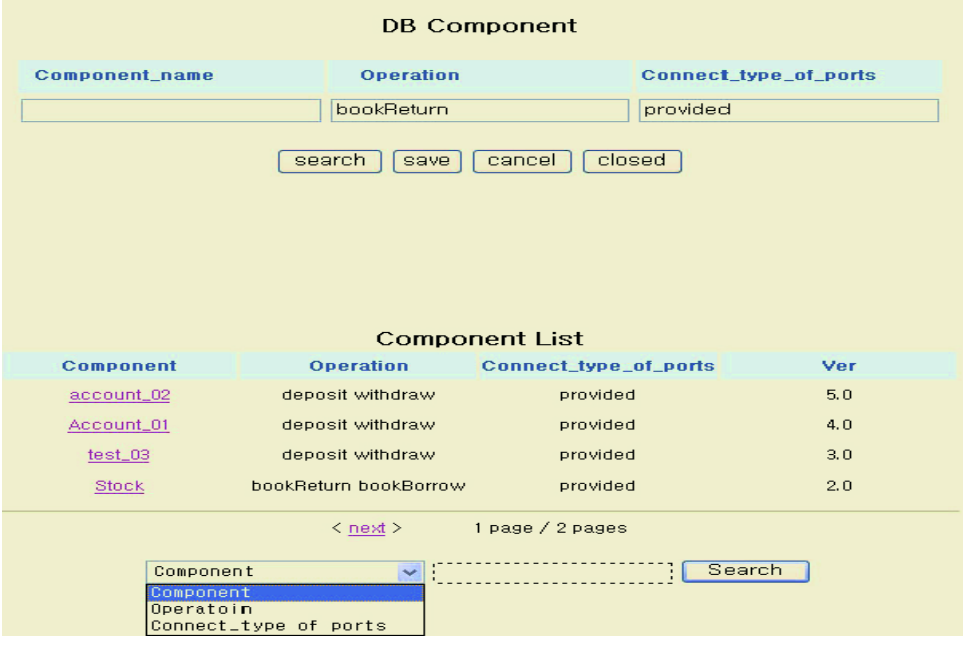


Fig. 7. DB component interface

Table 2. Component DB table

Component	operationPort	conncet_type of ports
Stock	getTotalStock	provided
Stock	bookInput	provided
Stock	bookOutput	provided

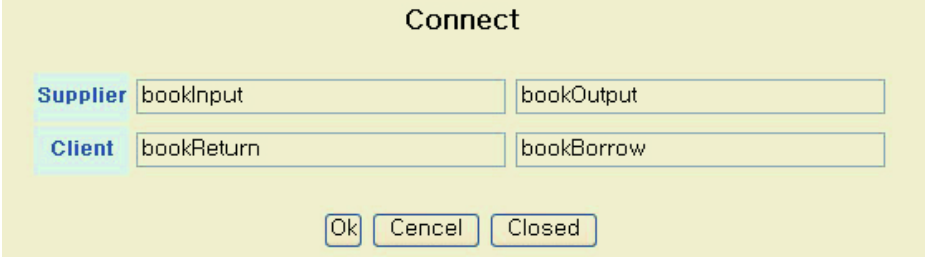


Fig. 8. Connector interface

Program Code 1. Connector specification

```
<?xml version="1.0" encoding="euc-kr"?>
<connectOperations>
<adapt>
 <supplier>
     <component> Stock </component>
     <operation>bookInput()</operation>
 </supplier>
 <client>
     <component> Library</component>
     <operation> bookReturn()</operation>
 </client>
</adapt>
<adapt>
 <supplier>
     <component> Stock </component>
     <operation> bookOutput()</operation>
 </supplier>
 <client>
     <component> Library </component>
     <operation>bookBorrow()</operation>
 </client>
</adapt>
</connectOperations>
```

The data from the interface in figure 8 enters into <adapt> as an element. That is, child element of <adapt> becomes <supplier> and <client> element and then <component> and <operation> element is created as an ingredient of each element. The data inputted through the interface is entered into <component> and <operation> element. The connector is generated automatically using the specification in figure 6. First, we search {adapt} and look up the dependent component by referencing to 'supplier' in operation expressed with stereo type <<connectOperation>> and then create the object of that component in the connector. And we perform the object referencing through stereotype <<operationPort>> of the generated object and then, by processing the task that finds the class using that object and the stereotype <<operationPort>> of that class, can use the component in dependent state. For the example of Library system, if the operation in stereo type <<connetOperation>> is bookInput() and bookOutput() operation of 'Stock' component, then it is represented as bookReturn() and bookBorrow() operation of 'Library' component. Since it is bookInput() and bookOutput() operation of 'Stock' component with stereo type <<operationPort>>

in 'supplier', 'Stock' component object is created. And, since it is bookReturn() and bookBorrow() operation of 'Library' component with stereo type <<operationPort>> in 'client', the connector gets to use those operations. Therefore, the connector is created like Program Code 2.

Program Code 2.Connector creation

```
public class connector{
 Stock Stock = new Stock();
  public void bookReturn(int quality){
    Stock.bookInput(quality);
}
  public void bookBorrow(int quality){
    Stock.bookOutput(quality);
 }
}
```

5 Conclusions

Recently, in order to reuse the system, application software moves into the component unit, and hence the research about the component assembly to reuse components is progressed vigorously. In this paper, we presented the automatic connector creation that connects components to assemble components and described how to define each component, set up ports, draw up the specification by defining 'Operation Port', 'Connect Port' and 'adaptor', and create the connector automatically. Since the connector is generated automatically, we can assemble components without method all or modification. As a future work, the component configuration management through the connector created automatically is required.

References

1. Magnus Larsson, Ivica Crnkovic, "New Challenges for Configuration Management", Ninth International Symposium on System Configuration Management (SCM-9), Toulouse, France, September 1999.
2. Alan W. Brown, K.C. Wallnau, "Engineering of Component-Based Systems", Proceedings of the 2nd IEEE International Conference on Complex Computer Systems, Oct. 1996.
3. Edward V. Berard, Essays " Object-Oriented software engineering", Volume 1, Prentice Hall, 1993.
4. J.Bosch,"Superimposition : A Coponent Adaptaion Technique", Information and Software Technology, Vol. 41, Issue 5, March 25, 1999.

5. George T. Heineman, William T. Councill, "Component-Based Software Engineering", Addison-Wesley, pp. 485-549, 2001.
6. George T. Heineman, "Adaptation and software architecture", Third International Workshop on Software Architecture, Orlando, Florida, November 1998.

MaRMI-RE: Systematic Componentization Process for Reengineering Legacy System

Jung-Eun Cha and Chul-Hong Kim

S/W Engineering Department, Electronics and Telecommunications Research Institute
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-350, Korea
{mary2743, kch}@etri.re.kr

Abstract. Most legacy systems have many problems to accommodate new technologies, or to be expanded or changed in accordance with complicated business requirements, since they are lack of standardization, openness, distributed architecture, and et al. Therefore, it is necessary to reengineer the legacy systems to maximize the utility thereof as an important asset of an organization. In this paper, we provide the a componentization process, that is the MaRMI-RE(Magic and Robust Methodology Integrated-ReEngineering), for reengineering legacy systems into component systems so that legacy systems can continue to be developed to comply with varying business and technical environments. We define and specify concrete procedures, work-products guidelines and considerations for transforming into component-based system with a well-defined architecture and more reusable assets.

1 Introduction

A legacy system is an application that was developed on older technology and is past its prime use, but still play an important part in current businesses[1]. Increasingly, legacy system is being viewed as an asset that represents an investment that grows in value rather than a liability whose value depreciated over the times[2]. Now legacy systems are being pressured to concurrently respond to increasing requirements. So, most legacy systems are faced with changing industry models such as e-Business and globalization, changing business models such as CRM, emerging information technologies such as Internet and open system, emerging new information architectures such as J2EE, Web service and component reuse[3]. That is, in order to utilize a legacy system as a reusable asset having the core value to an organization, it is required to reengineer the legacy system into a new target system having systematic architecture. Only by reengineering, the understandability and reusability of the system are improved, a flexible maintenance structure can be constructed, and thus a system evolution model capable of accommodating later system variations can be obtained[4]. In particular, the necessity of reengineering legacy systems into component-based systems with better design construction and architecture has been further emphasized, as the Internet becomes ubiquitous not only as an information sharing medium for people and organizations but also as a core technology for businesses, and as Component Based Development (CBD)

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 896 – 905, 2005.

based on pre-developed interoperable independent components becomes the dominant software (S/W) development paradigm.

However, conventional reengineering methodologies are not provided with support systems and standard guidelines allowing users to select or repeat reengineering procedures and techniques to satisfy their intentions, and therefore it is unavoidable to depend on users' subjective judgments at the time of important decision. Moreover, insufficient efforts have been made to concretely define the procedures and techniques of reengineering, so that a great number of organizations have repeatedly undergone similar trial and error in promoting reengineering projects.

In this paper, we propose the MaRMI-RE as reengineering process for transforming legacy system into component systems so that legacy systems can continue to be developed to comply with varying business and technical environments; and, more particularly, to a reengineering process which presents procedures, techniques and work products required to systematically transform legacy systems into component systems.

2 Related Works

As a conventional reengineering methodology that has been most widely referred to, there is Common Object-based Reengineering Unified Model (CORUM)[5] that is developed at the Carnegie Mellon University (CMU) Software Engineering Institute (SEI). This methodology collects and arranges requirements from various standpoints to integrate architecture-based reengineering tools with code-based reengineering tools, and provides a framework required to prepare solutions that meet the requirements. It presents an integrated model of an architecture-based reengineering process and a forward engineering process. However, this method presents neither a detailed work process that is concretely applicable to the execution of a reengineering project, nor the guidelines and techniques of tasks that are required for the execution of the process. Mission ORiented Architecture Legacy Evolution (MORALE)[6], developed at the Georgia Institute of Technology to improve a system by reflecting a new requirement (user-configurable view) in Mosaic Web browser, detects and effectively analyzes risk elements for the initial change of the evolution of the system, and then extracts components that can be used in a new system, with an emphasis on the mission of an organization rather than technical elements.

However, according to most of the above-described research, a reengineer is charged with a risk of information loss or deformation, which may occur during the transformation of the legacy system into a target system, without a support from systematized task procedures or work products. Accordingly, a reengineering methodology, capable of providing processes and techniques to systematically transform and integrate a large-scale legacy system into a component-based system, is required. Recently, researches that aim at identification and transformation of architecture are suggested, but they not only don't provide concrete and practical reengineering methods[7].

3 Basic Concept

3.1 Conceptual Model

Our MaRMI-RE is built on the Horse Shoes metamodel of CORUMII. But MaRMI-RE includes Planning steps for deriving the best transformation strategy and process.

(Figure 1) illustrates a conceptual model of MaRMI-RE, which is an architecture-based componentization process for reengineering a legacy system. It is a component-based reengineering process to transform an AS-IS model, which a legacy system has, into a "TO-BE model", which a target system includes, and is an architecture-based reengineering process capable of accommodating temporary change requirements[8].

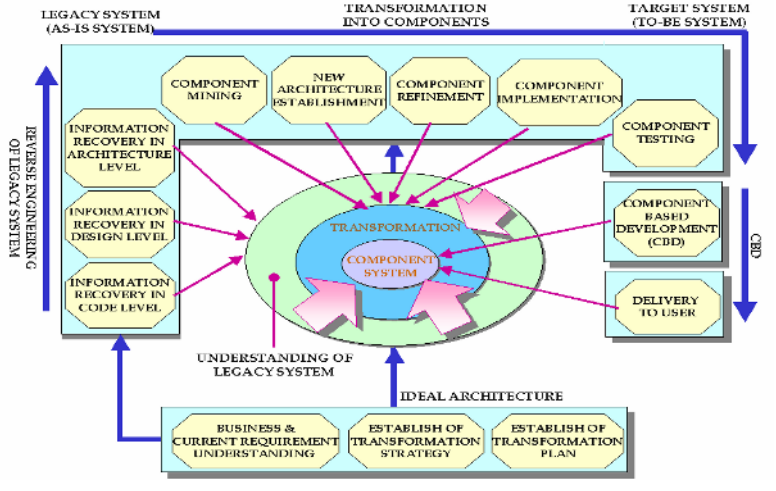


Fig. 1. Concept Model of MaRMI-RE

3.2 Customizing of Reengineering Process

Our process allow to customize a reengineering process in parallel and selectively, unlike a sequential or synchronized development process provided by conventional

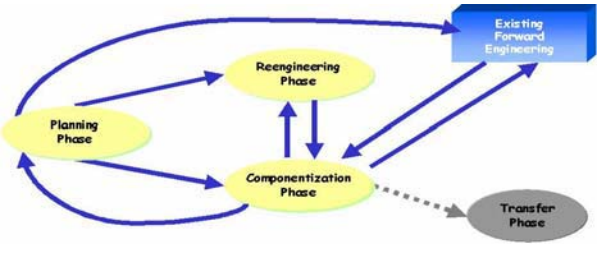


Fig. 2. Basic Process of MaRMI-RE

methodologies, thus supporting continuous expansion, assembly and customization process on the basis of target architecture[9]. (Figure 2) is the basic process for performing reengineering project through adapting MaRMI-RE.

Table 1. Individual development scenarios

Type	Scenario
1	Planning→Reverse Engineering Componentization→ Transfer
2	Planning→ Componentization Reverse Engineering → Transfer
3	Planning→ Componentization → Forward Methodology → Componentization → Delivery
4	Planning→Forward Methodology→ Componentization → Transfer
5	Planning→Reverse Engineering Componentization→Forward Methodology → Componentization→ Transfer
6	Planning →{Reverse Engineering | Forward Methodology}→ Componentization → Transfer

The (Table 1) is shown the basic scenarios based on (Figure 2). Type 3 scenario directly proceeds to the componentization phase without the tasks of the reverse engineering phase, required activities of conventional forward engineering The MaRMI-Ⅲ[10], are performed so as to generate newly required business components, and the results thereof are integrated with the work products of the componentization phase, thus performing the project. At the primary analysis of the planning phase, if most parts must be newly changed without being greatly influenced by the legacy system, required components are first generated through the tasks of the conventional forward engineering methodology, and then the type 4 scenario proceeds to the componentization phase, so that componentization tasks based on the vision and strategy of the reengineering project are performed. The type 4 is a combination of type 1 and type 3, which is used when the target system requires businesses other than businesses included in the category of the legacy system. Lastly, by type 6, a reengineering project established in the planning phase is executed by performing a procedure of integrating obtained component information with components of newly added businesses in the componentization phase, after the components of newly added businesses are generated through the tasks of conventional forward engineering methodology at the same time that information on components to be extracted from the resources of the legacy system are obtained through the reverse engineering phase.

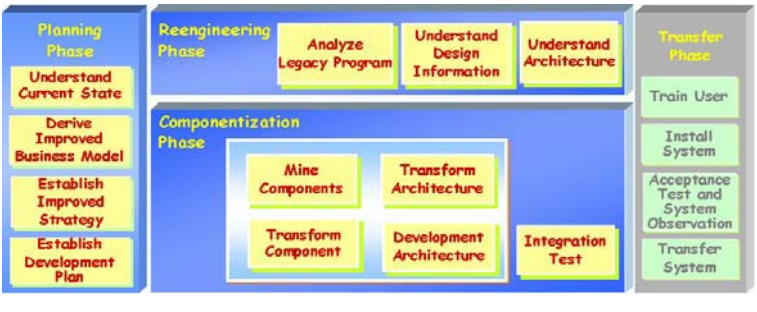


Fig. 3. Phase and activities in MaRMI-RE

4 Activities and Tasks in Each Phase

(Figure 3) illustrates a view showing 4 phases and 16 activities constituting the entire process of our process. Each phase is performed independently or sequentially.

4.1 Planning Phase

The planning phase is to determine whether to proceed to componentization through the entire analysis of the legacy system, and to present a reengineering direction for subsequent phases. The phase is shown in (Figure 4). Through the tasks, problems of legacy system are grasped, a business direction is analyzed to determine a suitable improvement direction, and the purpose, target and scope of a project are fixed, thus drafting a development plan.

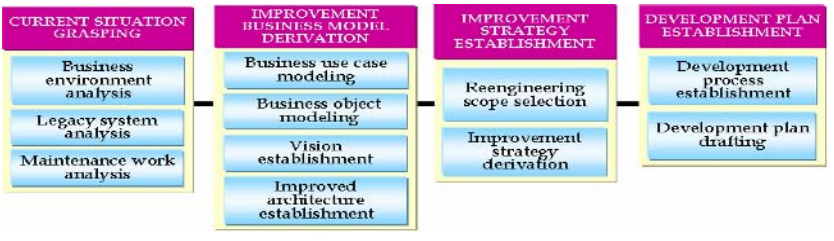


Fig. 4. Activities and tasks in Planning Phase

(1) Current situation grasping activity
This activity is to grasp the configuration of an organization, the workflow and the greatest issues that the organization faces, through the analysis of entire and general information about the work, and to understand the function of the work and the functions of sub-systems for each work unit. Also, it is to analyze information about the maintenance and management of the legacy system.

(2) Improvement business model derivation activity
The goals are to clearly grasp the requirements of parties concerned with the activity through business use case modeling and business object modeling, and to present an improvement business model, which is to be a target later. On the basis of the improvement business model, the purpose and scope of the project are determined. The architecture generated in this case is the ideal model of the business area, which presents an aim to set architecture information recovery in the reverse engineering phase or the target architecture in the componentization phase.

(3) Improvement strategy establishment activity
It provides an optimal approaching method to perform a reengineering project. For this activity, reengineering work to be improved is selected, and technical elements are analyzed from the standpoint of the business value and system for the work to determine reengineering priority, and an optimal transformation strategy for componentization is established with respect to each work unit. The strategy established here is compared to analysis results in an architecture transformation activity of the componentization phase of (Figure. 6).

(4) Development plan establishment activity
The activity is to select items of the procedure and work product of the process to be actually applied to the development by establishing a development process on the basis of a determined component transformation strategy, and to draft a development plan by collecting and arranging the work products obtained from previous tasks. For a reengineering scenario suitable for the user's requirement and the organization's capability based on the strategy determined through the activity, one of the scenarios derived from the basic process of (Table 1) is selected.

4.2 Reverse Engineering Phase

(Figure 5) shows the activities and tasks of this phase to improve the understanding of static and dynamic action information by analyzing the work products of the legacy system. Architecture information is understood and abstracted through recognition of relations between the legacy system elements, so that a preparation task for componentization is performed, and a modeling task for abstracting the analysis results of the code semantics in the form of design information is performed.

(1) Program analysis activity
In this activity, the syntax information and semantic information of the legacy program are analyzed and extracted at system and unit program level by source code restructuring and analysis, and pieces of analyzed information are normalized using a relationship diagram between data and control flows, a call graph between modules, etc. In this activity, efficiency can be increased through the use of automated tools.

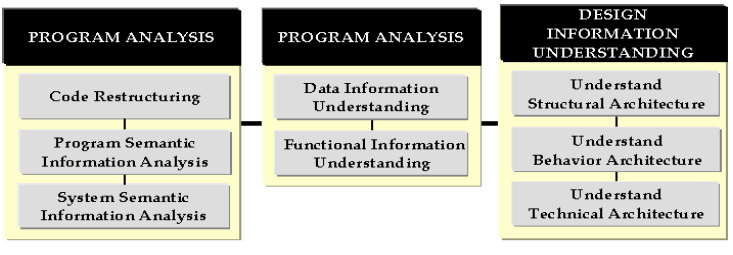


Fig. 5. Activities and tasks in Reverse Engineering Phase

(2) Design information understanding activity
This is to identify functional unit processes on the basis of program analysis information, specify control flow between the unit processes and data flow between the unit processes and related tables, and provide system design information for architecture understanding, which is a subsequent activity. This activity is used to obtain higher understanding by modeling the design information of the legacy system and abstracting the modeling results in the form of a structural diagram.

(3) Architecture understanding Activity
This is to improve the understandability for the legacy system through information recovery for structural, technical and behavioral architecture constituting the legacy system. That is, Modules of the legacy system, are further abstracted and identified by

the unit of independent component (sub- system), and interdependence between the elements is expressed; How call relations between components are made is understood on the basis of sub-systems or components constituting the structural architecture so as to grasp the entire behavior of the legacy system.

4.3 Componentization Phase

The main activities of this phase are transforming legacy information including all outputs as likes source codes, design modeling et al., to new forms required in target environment. (Figure 6) illustrates the activities and tasks of the componentization phase. This phase groups parts with higher relevance and identifies the grouping results as component candidates so as to componentized system entities with higher semantic relevance on the basis of the information extracted through the reverse

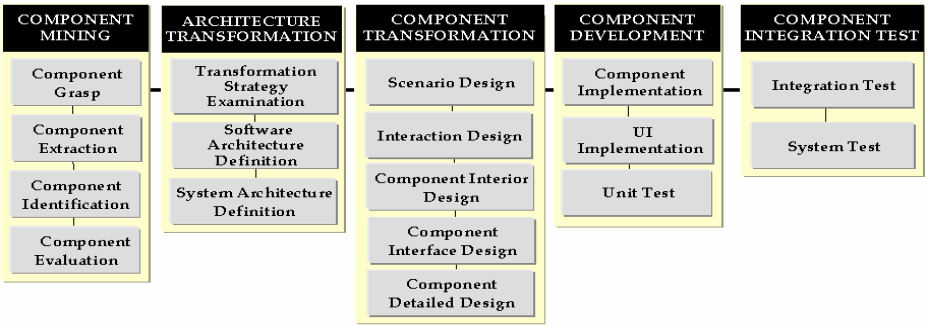


Fig. 6. Activities and tasks in Componentization Phase

Table 2. Tasks and Detailed Procedures of The Component Mining Activity

Task	Summary	Procedure	Work product
Component grasp	Component candidates performing independent business functions are selected, and system entities constituting each of the candidates are traced and grasped with respect to each candidate.	(1)use case related system entity grasp (2)use case analysis (3)component candidate grasp	-interrelation modeling table -use case analysis table -component entity description report
Component extraction	Components are extracted on the basis of the system entities constituting each component, and interrelations and interactions there between are grasped.	(1)sharing element grasp (2)component extraction (3)grasp of interrelations/interactions between components	-component list table/ interaction table/entity description report -application use case/component correspondence table
Component identification	Components performing independent functions not included in the legacy system are identified as components and extracted on the basis of business use cases constructed in the planning phase.	(1)component candidate grasp (2)component extraction	-component list table -business use case/component correspondence table
Component evaluation	A utility method related to how the extracted components are to be utilized is established and evaluated, in which interrelations and interactions between components are readjusted/the system is expressed on the basis of the interrelations between the extracted components.	(1)establishment of component utility strategy, and evaluation criteria for utility strategy (2)component evaluation (3)readjustment of interrelations/interactions between components	-component list table - component interaction table -{application use case/component | business use case/component correspondence table} }

engineering process in the legacy system. Further, the reengineering methods of the legacy system and the strategy are determined, and S/W, component and system architectures are defined to componentized extracted reusable components. Also, the interfaces of the extracted components are identified; the static and dynamic structures of the components are created and are transformed into system-manageable programs newly defined on the system architecture.

(1) Component mining activity
This activity used to execute a task of transforming the legacy system into a system having new architecture. Therefore, the legacy system is divided into several parts according to units performing a business function, and the division parts are allowed to correspond to respective components and then grasped and extracted. (Table 2) summarizes the tasks and detailed procedures of this activity.

(2) Architecture transformation activity
Our goals in this activity are to confirm a method of reengineering the legacy system and a strategy of successfully performing the reengineering, and fixing a technique of componentizing the extracted reusable components. Tasks constituting this activity are summarized and described in (Table 3).

Table 3. Tasks and Detailed Procedures of the Architecture Transformation Activity

Task	Summary	Procedure	Work product
Transformation strategy examination	Reengineering scope and method are determined, strategy and technique of componentizing extracted components are defined, and the appropriateness thereof is examined. That is, reengineering requirements and transformation types are analyzed, and transformation strategy is established.	(1)transformation strategy and component utility strategy comparison/analysis (2)transformation strategy readjustment (3)refinement of improvement strategy establishment report of target system	-transformation strategy examination report -improvement strategy establishment report (refinement)
Software architecture definition	Pieces of architecture analysis information obtained from various standpoints are examined to identify the functional requirements and quality attributes of a target system, and the architecture structure of the target system is set, thus defining the software architecture of the target system.	(1)architecture analysis information examination (2)definition of functional requirements of target system (3)quality attribute derivation (4)architecture structure setting (5)s/w architecture definition	-architecture information analysis report -architecture functionality list table -architecture quality attributes list table -quality scenario
System architecture definition	The technical architecture and component architecture of the target system are defined, and defined components are arranged in a physical environment, thus deriving the system architecture of the target system.	(1)technical architecture definition (2)component architecture definition (3)system architecture definition	-technical architecture -component architecture -system architecture

(3) Component transformation activity
The component transformation activity is to identify the interfaces of extracted components, design the internal structure of the components, and identify the operations of the component interfaces on the basis of dynamic message flow information between the internal classes of the components.

The detailed procedures and main work products are summarized in the (Table 4).

(4) Component development activity
This activity is to implement applications to comply with a component platform on the basis of the component detailed design information. That is, through the use of the pieces of design information extracted through the component transformation activity, components are mapped to comply with an implementation technology platform and then implemented thereby. And, it includes a unit test implemented components.

Table 4. Tasks and Detailed Procedures of the Component Transformation Activity

Task	Summary	Procedure	Work product
Scenario design	Scenarios of a task flow related to how respective use cases identified in the plan and reverse engineering phases must be operated in a new target system are analyzed and designed.	(1)drafting of normal scenario according to use case (2)drafting of selective scenario according to use case (3)drafting of exceptional scenario according to use case	-use case specification
Interaction design	Which interaction is performed between entities in order for each use case to perform a corresponding task in the system is modeled on the basis of information of use cases and entities.	(1)use case selection and actor placement (2)object or entity arrangement (3)message identification (4)interaction diagram drafting	-component interaction diagram
Component interior design	Internal elements of each component are identified, and the internal structures of the component are designed.	(1)class extraction (2)method and attribute grasp (3)relation setting (4)class allocation according to component	-class diagram -component diagram
Component interface design	Services to be provided according to component are defined by grasping interfaces thereof, and required services according to interface are extracted through operations.	(1)use case-based interface identification (2)data-based interface identification (3)interface refinement (4)interface details (5)component details	-component specification -component diagram (refinement)
Component detailed design	In order to describe components in detail in conjunction with a specific platform (J2EE), mapping to beans and interfaces are defined, and the design of parts related to continuity, transaction and security is performed.	(1)packaging definition (2)EJB mapping definition (3)continuity design (4)transaction design (5)security design (6)arrangement design	-component detailed design report

(5) Component integration test activity component integration test activity
This activity is to integrate developed individual components with each other through the construction of prototyping, thus determining whether the entire functionality of the legacy system is exhibited, and analyzing and examining restriction items. For this activity, extracted components are arranged on the architecture of the reengineering system and integrated with each other on the basis of a transaction strategy, thus examining whether the implemented components normally communicate with other components. Further, whether the component architecture and business requirements are sufficiently defined and implemented is examined.

5 Current Status and Conclusion

The MaRMI-RE features an architecture-based approach to deriving new application structure, a reverse engineering technique for extracting architectural information from exist code and business domain knowledge, an approach to component system

generation that system's architectural components can be reused in the evolved version, and a specific technique for dealing with the difficulties that arise when extracting components from legacy system and deciding transformation strategies and process. So, Therefore, this process is advantageous in that an organization recognizes a legacy system thereof as a reusable asset, and the creation of continuous values can be performed on the basis of a component system even though business and technical environments related to the system are changed.

In particular, it presents a core reengineering process, detailed procedures and guidelines thereof, and work products required to execute the reengineering process, so that organizations intending to perform reengineering can utilize the process, detailed procedures and guidelines, and work products as ideal reference tools to obtain the reengineering effect that the organizations expect.

Currently, we achieves MaRMI-RE version 1.0, and have finished two case studies adapted out MaRMI-RE. One case is textile-manufacturing system constructed with IBM CICS COBOL, and another case is the Business Management System as a part of banking business operated on AS 400. The Business Management System as legacy system is consisted of 124 COBOL programs and has 3 versions with different development environments and different operational environments. Each system is operated independently on different company.

In Future, we intend to maximize a usability by completing development process supporting full life cycle integrated forward engineering to reverse engineering, and constructing supporting tools, acquire a reliability through case studies in various business fields.

References

[1] Dolly M, Neumann, "Evolution Process for Legacy System Transformation", IEEE Technical Applications Conference, Washington, November, 1996, pp57-62

[2] Nelson Weiderman, Dennis Smith, Scott Tilley, "Approaches to Legacy System Evolution", CMU/SEI-97-TR-014, 1997

[3] William Ulrich, Legacy Systems : "Transformation Strategies", Prentice Hall, 2002

[4] SEI Reengineering Center "Perspectives on Legacy System Reengineering, 1995

[5] Rick Kazman, Steven G. Woods, S. Jeromy Carriere, "Requirements for Integrating Software Architecture and Reengineering Models: CORUM II", Fifth Working Conference on Reverse Engineering, Honolulu, Hawaii, Oct 1998, pp: 154-163

[6] Abowd G. Goel A. Jerding D.F., McCracken M., Moore M., Murdock J.W., Potts C., Rugaber S., Wills L., "MORALE. Mission ORiented Architectural Legacy Evolution" International Conference on Software Maintenance, Bari, ITALY, October, 1997, pp150 – 159

[7] Jochen Seemann, Jürgen Wolff von Gudenberg, "Pattern-Based Design Recovery of Java Software", *Communications of the ACM*, Vol. 38, No. 10, pp65-74, October 1995.

[8] Jung-Eun Cha, et al., "Reengineering Process for Componentization of Legacy System", Journal of the Korea Society of System Integration, Vol.2, No.1, pp 111 – 122, May, 2003

[9] Jung-Eun Cha, et al., "Establishment of Strategies and Processes for Reengineering of Legacy System", proceedings of the 20th KIPS Fall Conference, Vol.10, No.2, Nov. 2003

[10] Jin-Sam Kim, et. al., MaRMI-Ⅲ User manual, 2003, ETRI(Electronics and Telecommunications Research Institute)

A Study on the Mechanism for Mobile Embedded Agent Development Based on Product Line

Haeng-Kon Kim

Department of Computer Information and Communication Engineering,
Catholic University of Daegu, Kyungbuk, 712-702, Rep. of Korea
hangkon@cu.ac.kr

Abstract. In most mobile embedded agent systems (MEAS), agents are required to achieve their own goals. An agent's goals, however, can conflict with others either when agents compete with each other to achieve a common goal or when they have to use a set of limited resources to accomplish agents' divergent goals. In either case, agents need to be designed to reach a mutual acceptable state where they can avoid any goal conflicts through negotiation with others to achieve their goals. In this paper, we consider that a ABCD architecture, as a shorthand for Architecture platform, Business, Common, and Domain is a core component of agents' mental attitudes and represent resource-bounded ABCD agents in logic programming framework. We propose an algorithm in which ABCD agents with different goals solve their problems through negotiation resolving goal conflicts. Finally, we develop a negotiation meta-language to show the effectiveness of the negotiation system proposed in this paper.

Keywords: Multi-Agent Systems, ABCD Agents, Negotiation, ACL, Logic Programming.

1 Introduction

Research in MEAS is concerned with the behaviors of a collection of autonomous agents trying to solve given problems that are beyond their individual capabilities. These agents are autonomous and may be heterogeneous in nature. Agents in MEAS need to interact with others since there exist inherent interdependencies among them, but inter-agent conflicts may arise because of two basic reasons: different agents have contrasting goals and they have inconsistent knowledge. Negotiation might be a promising mechanism for resolving these conflicts. The Contract Net Protocol (CNP) has present negotiation as an organizing principle used to effectively match tasks and agents. Parsons *et al*. [1] have proposed a formal framework, based on a system of argumentation, which permits agents to negotiate to establish acceptable ways to solve problems. Jennings *et al*. [2] have studied on the automated negotiation, and viewed negotiation as a distributed search through a space of potential agreements. Fatima *et al*. [3] have introduced bilateral negotiation between agents with time constraints and incomplete information. In general, there exist goal conflicts among agents in MEAS and it is usually not possible that an agent has complete knowledge

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 906 – 915, 2005.

about others. When agents' goals conflict with others' in such a state that each possesses partial knowledge about other agents, research on how to negotiate is important like in human society [4]. Therefore, agents in MEAS need to communicate with other agents to achieve goals and need to be designed to negotiate with others and to reach a mutual acceptable state where agents can avoid any goal conflicts due to theirs interdependencies upon others. In this paper, we represent agents' *beliefs, calls,* and *desires* (ABCD) for MEAS in logic programming and develop a negotiation mechanism for ABCD agents. We then introduce an ACL and a negotiation protocol for ABCD agents and show how goal conflicts through negotiation in MEAS are resolved using our ACL and negotiation mechanism in practice.

2 Related Works

2.1 Product Lines

The software market - just like other markets - has a great demand for variety in products. The entirety of product-variants of one software is also referred to as software system family or software product line[8]. Manufacturing was the first discipline that provided an answer how to efficiently build varying products. Instead of building single system family members, interchangeable parts were assembled to products. Reuse implies that the asset base contains artifacts which can be reused in more than only one single product. This means that the asset base contains common parts which do not change between product line members and variable parts that feature different functionalities from member to member. Variability can be realized on run-time or development time. Product line engineering is concerned with development time variabilities. A first approach to supporting commonalities and variabilities at the design levels refers to parameterization as in fig.1. The underlying idea is to make a component flexible so that it can be tailored according to the context it is put in. To this end the component is equiped with parameters that specify its behaviour. The common parts of the component are not parameterized while the varying parts are.

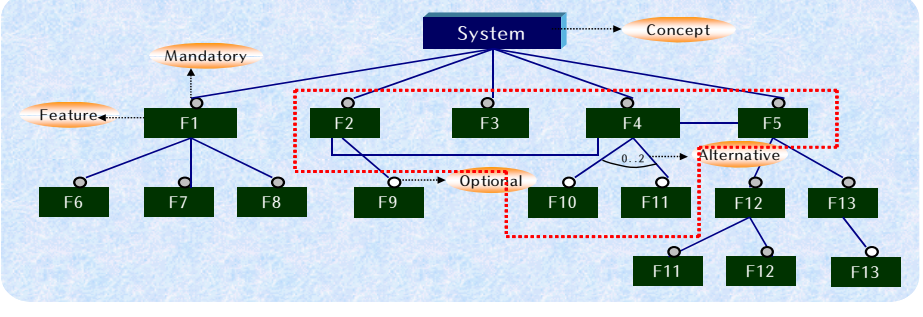


Fig. 1. Techniques for modeling variability

2.2 Negotiation Techniques for Resolving Conflicts

2.2.1 Contract Net Protocol (CNP)

The CNP for decentralized task allocation is one of the important paradigms developed in distributed artificial intelligence. A contract net problem-solver is a collection of agents in manager and worker roles. A top level task is allocated to a manager agent, which generates subtasks and issuing task announcements for them to some subset of potential worker agents. Workers bid on tasks they desire and are qualified for. The manager selects the highest rated bid and allocates the task to that worker, possibly monitoring the worker's progress toward solution. When several workers supply final reports of individual subtask results, the manager is responsible for integrating them and supplying a global solution.

2.2.2 Argumentation Systems

An argument is classified into one of a number of classes based upon its acceptability, which is determined by examining the arguments for and against given proposition to see whether any of the steps in the arguments themselves have any arguments against them. The argumentation systems is based on a well-ground framework for describing the reasoning process of negotiation agents. An originating agent puts forward an initial proposal. The recipient agents evaluate the proposal by constructing arguments for and against it. This system uses ABCD modalities and argumentation consequence relations (ACR). For example, one of the ACR constructs an argument for q from arguments for p and $p \rightarrow q$ where Δ is a knowledge base, A and B are labels indicating fact p and rule $p \rightarrow q$, respectively.

2.2.3 Negotiation in Electronic Commerce Systems

There are agent-mediated electronic commerce systems that assist a customer in negotiating the terms of transaction. AuctionBot users create new auctions to sell products by choosing from a selection of auction types and then specifying its parameters. Buyers and sellers can then bid according to the multilateral distributive negotiation protocols of the created auction. AuctionBot then manages and enforces buyer bidding according to the auction protocols and parameters. On the other hand, Anthony *et al*. [4] have presented the design of autonomous agent that can participate across multiple online auctions, in particular, English auction and Dutch auction. The agent makes decisions on behalf of a consumer and endeavors to guarantee the delivery of the item according to the user's preferences.

3 ABCD Agent Architecture for Negotiation

3.1 Definition of Conflicts

In this paper, we assume that agents would recognize conflict situations in which mental attitudes of agents in conflict by checking two unacceptable conditions. The unacceptable conditions are the cases that one agent requests resources from another agent, but the recipient should possess the resources to achieve its goal, and the following *rebutting* occurs:

Definition 1. Let s and s^* be sentences of the form $M^1(p_1) \wedge \ldots \wedge M^n(p_n)$ and $M^{*1}(q_1) \wedge \ldots \wedge M^{*m}(q_m)$ from some knowledge base Δ, respectively where M^i and M^{*j} $(1 \le i \le n, 1 \le *j \le m)$ are belief, desire, or intention predicates. s *rebuts* s^* if $p_i \equiv \neg q_j$ for some p_i and q_j.

The rebutting means that for example, one agent's request directly conflicts with another agent's mental attitudes. It might occur when each agent has conflicting beliefs, desires, or intentions, i.e., $B_i(a) : B_j(\neg a)$, $D_i(p) : D_j(\neg p)$, or $I_i(p) : I_j(\neg p)$. It might also occur when one agent has erroneous assumptions about another agent's knowledge, i.e., $B_i(B_j(a)) : B_j(\neg a)$, $B_i(D_j(p)) : D_j(\neg p)$, or $B_i(I_j(p)) : I_j(\neg p)$.

3.2 Axioms for Resource-Bounded ABCD

We consider that the axiomatization for beliefs is the standard *KD*45, *D* and *K* axioms for desires and intentions, and Modus Ponens inference rule. We also consider the following axioms for resource-bounded ABCD:

$$B_i(have(X, Z) \wedge give(X, Y, Z)) \rightarrow B_i(have(Y, Z)).$$
$$B_j(\neg have(j, Z) \wedge give(X, j, Z)) \rightarrow B_j(have(j, Z)).$$
$$B_i(have(X, Z) \wedge give(X, Y, Z)) \rightarrow B_i(\neg have(X, Z)).$$
$$B_i(have(i, Z)) \wedge \neg B_i(holdon(i, Z)) \wedge request(X, i, give(i, X, Z)) \rightarrow give(i, X, Z).$$
$$B_i(have(i, Z)) \wedge B_i(holdon(i, Z)) \wedge request(X, i, give(i, X, Z)) \rightarrow \neg give(i, X, Z).$$

3.3 A ABCD Agent Architecture

We consider a ABCD agent architecture in Figure 2 containing the following two major components: *knowledge base* and *negotiation library* including *planner*, *monitor*, and *communicator*. The knowledge base is a set of logical sentences that includes knowledge about the agent's capabilities and other agents', and rules for problem decomposition. The factors in the knowledge base are represented by predicates that mean mental attitudes of the agent. On the other hand, the negotiation library is responsible for deciding how to solve each task, supervising the execution of tasks, and handling incoming and outgoing messages.

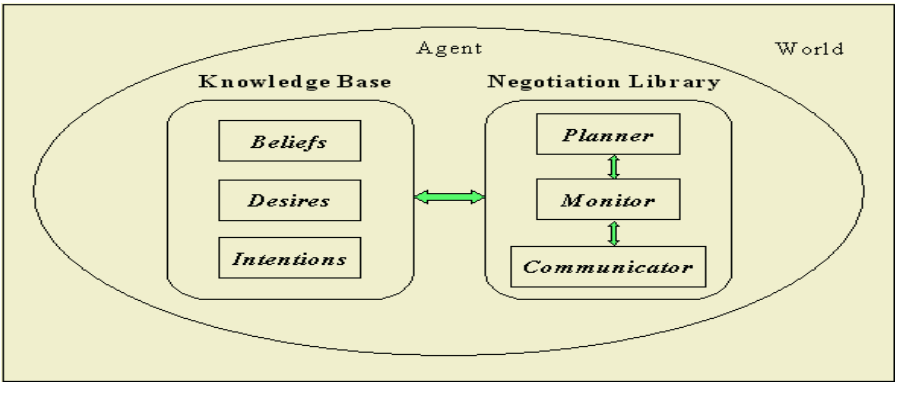


Fig. 2. A ABCD agent architecture for negotiation

4 A Negotiation Model for ABCD Agents

4.1 An ACL for ABCD Agents

Negotiation is achieved through the exchange of messages in a shared communication language. We use a variant of the KQML [5], FIPA ACL specification [6], and the negotiation meta-language [7] as our negotiation language for ABCD agents to resolve goal conflicts. The actual exchange of messages is driven by the participating agents' own needs, goals, or mental attitudes. We represent the set of beliefs as B, the set of desires as D, and the set of intentions as I. Each agent has a unique identifier and we denote the set of identifiers of the agents involved in negotiation as *Agents*. Assumed ABCD agents are negotiating about the allocation of deficient resources, agents require the allocation of deficient resources to achieve their goals. We denote a set of goals as *Goals* and a set of the resources as *Resources*. In this case, we can define a communication language *CL* for ABCD agents as follows:

Definition 2. Given $a1, a2 \in Agents$, $g \in Goals$, $r \in Resources$, and $m \in B$, D, or I, we define a *CL*:

$request(a1, a2, g, r) \in CL$ $\quad$ $ask\_if(a1, a2, m) \in CL$

$inform(a1, a2, m) \in CL$ $\quad$ $give(a1, a2, r) \in CL$

$reject(a1, a2, g, r) \in CL$ $\quad$ $alternative(a1, a2, g, subgoals) \in CL$

$achieved\_goal(a1, a2) \in CL$

4.2 A Negotiation Protocol for ABCD Agents

In order to simplify protocol analysis, we assume that two ABCD agents are involved in our negotiation protocol. The sequence of our negotiation protocol for ABCD agents can be shown in Figure 3 as a finite state diagram. In Figure 3, S0~S7 represent different negotiation states during a negotiation process. S0 is the initial state and S7 is the terminal state in which an agreement or disagreement is reached. The process of negotiation starts when an agent generates a *request* message. Other agents then either accept it, reject it, or make a counter-request.

4.3 Interpretation and Generation of Messages

The interpretation of messages implements a negotiation-state transition and the generation of messages determines a message to be taken in a particular state. However, the actual effect of a message depends upon the agent's interpretation. This interpretation process is highly domain-specific and is also dependent upon the internal structures presented in the agent architecture. How an agent chooses which message to utter depends upon many factors: agent's theory, the active goals of the agent, or the history of the negotiation; and it also depends upon the way that particular agent interprets those messages. We define a message generation function G under the condition of negotiating agent about deficient resources.

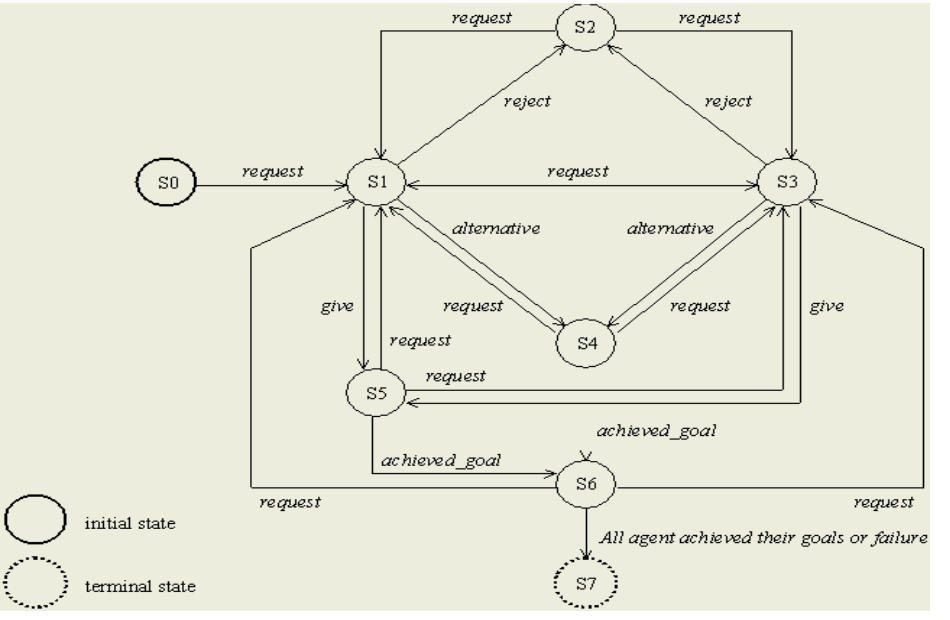


Fig. 3. A Negotiation protocol for ABCD agents

Definition 3. We suppose that agent $a2$ has received a message from agent $a1$. Given an appropriate CL, $a2$ generates a message depending upon the following message generation function G:

If goal g satisfies the unacceptable conditions and there exists an alternative to achieving g,
$G(request(a1, a2, g, r)) = alternative(a2, a1, g, subgoals)$
If g satisfies the unacceptable conditions and there does not exist an alternative,
$G(request(a1, a2, g, r)) = reject(a2, a1, g', r)$ or *make a counter-request*
If g does not satisfy the unacceptable conditions, $G(request(a1, a2, g, r)) = give(a2, a1, r)$
If $a2$ has received an alternative, $G(alternative(a1, a2, g, subgoals)) =$ *make a replanning*
If $a2$ has received resources r, $G(give(a1, a2, r)) =$ *continue planning*
If a request was rejected by $a1$, $G(reject(a1, a2, g, r)) =$ *search an alternative*
If $a1$ has notified its goal achievement, $G(achieved\_goal(a1, a2)) =$ *make a replanning*
If $a2$ was asked whether p is true or not, $G(ask\_if(a1, a2, p)) = inform(a2,a1,p)$
or $inform\ (a2,a1,not\ p)$
If $a1$ informs $a2$ of the truth of p, $G(inform(a1, a2, p)) =$ *continue planning*

5 Experiments and Comparisons

5.1 Experiments

5.1.1 Knowledge Bases of ABCD Agents

To show how our negotiation mechanism might work in practice, we consider three home improvement ABCD agents with different objectives and resources. Agent $a1$

has the intention of hanging a picture, *intend*(*do*(*a*1, *hang_picture*)), and believes that it has in its possession a picture, a screw, a hammer, a hanger nail, and a screwdriver. It also believes that its name is *a*1 and agent *a*3 has a hanger.

Agent *a*2 has the intention of hanging a mirror, *intend*(*do*(*a*2, *hang_mirror*)), and believes that it has a mirror and a nail. It also believes that its name is *a*2 and agent *a*1 has a screw, a hammer, and a screwdriver. Finally, agent *a*3 has the intention of hanging a clock, *intend*(*do*(*a*3, *hang_clock*)), and believes that it has a clock and a hanger. It also believes that its name is *a*3 and agent *a*1 has a hanger nail.

5.1.2 Implementation of ABCD Agents

We construct ABCD agents and allow them to negotiate with each other from the top-level window. NegotiationWindow class is responsible for creating this window,

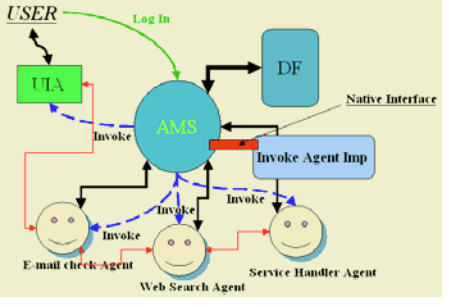


Fig. 4(a). Architecture of MESA

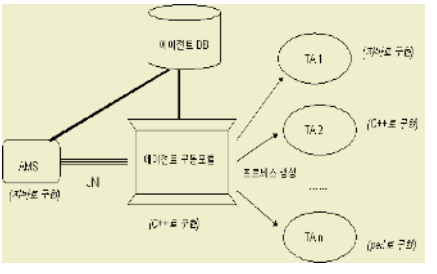


Fig. 4(b). Product Line for MESA

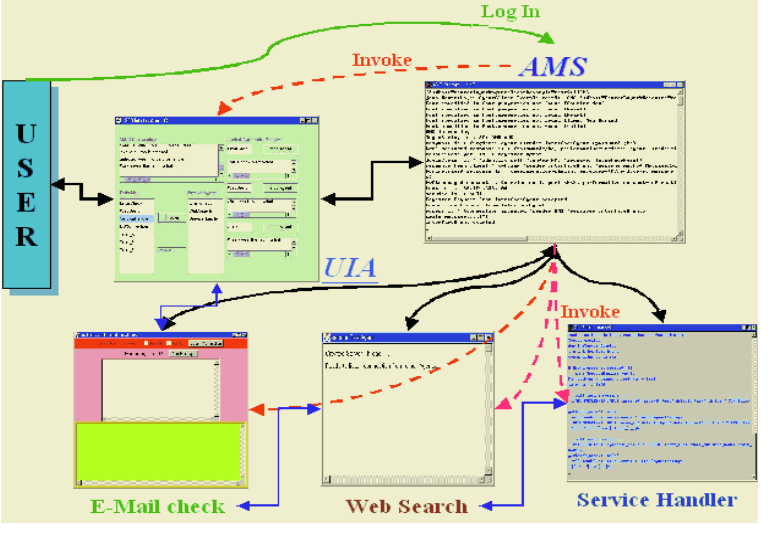


Fig. 4(c). Implementation of MESA

constructing the ABCD agents, and starting a negotiation with the agents. Figure 3.a shows the architecture for MEAS based on ABCD to implement and Figure 3.b illustrates the product line framework for MESA Figure 3.c demonstrates the execution examples on MESA. Table 1. Illustrates the roles of the methods in NegotiationWindow class.

Table 1. The methods in NegotiationWindow class

Method names	Roles
negotiationWindow	Constructor to create the instance
constructWindowContents	Creates a 600 × 500 pane
constructMenu	Creates menubar and menu, handles action events
*createSimpleABCDAgent*1	Creates agent1's window with Prolog engine
*createSimpleABCDAgent*2	Creates agent2's window with Prolog engine
*createSimpleABCDAgent*3	Creates agent3's window with Prolog engine
addItemtoMenu	Adds items to menu
main	Displays system information and take an array as a default Prolog engine

5.1.3 Negotiation Processes of ABCD Agents

Agent *a*1 tries to achieve its goal, *intend*(*do*(*a*1, *hang_picture*)). Thus, it tries to solve the query, ?-*solve*(*b*(*do*(*a*1, *hang_picture*))). This query is transferred to the monitor of *a*1 and again passed on to the planner of *a*1. The planner decomposes the query and comes to know that *a*1 needs a nail. *a*1 however has no knowledge about nail so that it asks *a*2 and *a*3 if they have a nail, *ask_if*(*a*1, *a*2, *b*(*have*(*a*2, *nail*))) and *ask_if*(*a*1, *a*3, *b*(*have*(*a*3, *nail*))), and then it waits for a reply through the communicator. On the other hand, *a*2 which receives an *ask_if* message from *a*1 checks its knowledge base, informs *a*1 that it has a nail, *inform*(*a*2, *a*1, *b*(*have*(*a*2, *nail*))), and waits for a reply. Now, *a*1 comes to know that *a*2 has a nail, requests a nail from *a*2, *request*(*a*1, *a*2, *hang_picture*, *nail*), and waits for a reply. This *request* message means that *a*1 requests a nail from *a*2 to achieve its goal, hanging a picture.

*a*2 which receives the *request* message from *a*1 first checks the unacceptable conditions of the resources in the message. It knows that the resources conflict occurs so that tries to achieve its goal, *intend*(*do*(*a*2, *hang_mirror*)). It decomposes the goal, but comes to know that it lacks a hammer to hang a mirror on its own. It requests a hammer from *a*1, *request*(*a*2, *a*1, *hang_mirror*, *hammer*). Agents continue to negotiate with one another by sending and receiving messages. Now, *a*1 believes that *a*2 and *a*3 have achieved their goals. It replans the process of goal achievement to achieve its own goal, requests a nail from *a*2, *request*(*a*1, *a*2, *hang_picture*, *nail*). *a*2 gives a nail to *a*1, *a*1 then achieves its goal and notifies *a*2 and *a*3 of the fact. Finally, *a*1, *a*2, and *a*3 reach a mutual agreement state through negotiation and achieve their conflicting goals.

5.2 Comparisions

The CNP does not allow agents to cause goal conflicts and a manager agent is responsible for all negotiations. Unlike the CNP, our negotiation mechanism for ABCD

agents supposes that agents may have conflicts. These conflicts occur when each agent has conflicting beliefs, desires, or intentions and when one agent has erroneous assumptions about another agent's knowledge. In this state, we show the capability of negotiation among agents without an agent playing manager role.

In this paper, we suppose the environments in which agents negotiate about accomplishment of conflicting goals and allocation of the deficient resources. Kasbah, an electronic commerce system implemented by LISP, creates a buying agent and a selling agent substituting for users. These agents negotiate with one another on Kasbah market. Kasbah agents may have conflicting goals and negotiate, changing suggested price using three functions of price decay or raise over time while our negotiation system negotiates changing agents' beliefs.

6 Conclusions and Future Research

In this paper, we have represented agents' beliefs, calls and desires for MEAS in logic programming environments and we have introduced a negotiation mechanism for ABCD agents. We have defined some axioms for resource-bounded ABCD agents. These axioms help agent to act rationally, to reason about the environments, and to plan its goal achievement. We have then established a ABCD agent architecture with four components: knowledge base, planner, monitor, and communicator. We have shown how goal conflicts through negotiation in MEAS are resolved using our ACL and negotiation mechanism in practice. Through the home improvement ABCD agents example, we have shown the possibility of negotiation among agents without complete knowledge about other agents. In this particular example, ABCD agents communicate with each other using our ACL and they negotiate based on our negotiation protocol. They evaluate a received message as checking the unacceptable conditions and generate a response message by the message generation function. A number of issues raised in this paper require further investigation. Firstly, negotiation strategies of agents vary according to the environments that they belong to. Secondly, we could carry out a type of hypothetical inference about abilities and actions of other agents, adding a function of abductive inference to the structure of ABCD agents. Finally, we have considered the ACL containing some communication actions or performatives.

References

[1] S. Parsons, C. Sierra, and N. R. Jennings. ″Agents that Reason and Negotiate by Arguing″. *Journal of Logic and Computation*, 8(3):261-292, 1998.

[2] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra, and M. Wooldridge. ″Automated Negotiation: Prospects, Methods and Challenges″. *International Journal of Group Decision and Negotiation*, 10(2):199-215, 2001.

[3] S. S. Fatima, M. Wooldridge, and N. R. Jennings. ″Multi-issue Negotiation under Time Constraints″. In *Proceedings of the First International Conference on Autonomous Agents and Multiagent Systems*, 2002.

[4] P. Anthony, W. Hall, V. Dang, and N. R. Jennings. ″Autonomous Agents for Participating in Multiple On-line Auctions″. In *Proceedings of the International Joint Conferences on Artificial Intelligence Workshop on E-Business and the Intelligent Web*, pages 54-64, 2001.

[5] Y. Labrou and T. Finin. ″A Proposal for a New KQML Specification″. *Technical Report CS*-97-03, Computer Science Department, University of Maryland Baltimore County, 1997.

[6] FIPA Communicative Act Library Specification. *http*://*www.fipa.org*/*specs*/*fipa*00037/. *The Foundation for Intelligent Physical Agents* (FIPA), 2000.

[7] M. Wooldridge and S. Parsons. ″Languages for Negotiation″. In *Proceedings of the Fourteenth European Conference on Artificial Intelligence*. 2000.

[8] Atkinson, J. Bayer, Etc., Component-based Product Line Engineering with UML, Addison Wesley, 2002.

Frameworks for Model-Driven Software Architecture

Soung Won Kim[1], Myoung Soo Kim[2] , and Haeng Kon Kim[3]

[1] Department of Electrical & Electronics Engineering, Anyang University, Anyang, Kyunggi, 430-714, South Korea
swkim@anyang.ac.kr
[2] Dept. of Computer Engineering, Korea Third Military Academy, Yung Chun, Kyungbuk, 770-849, South Korea
profkms@hanmail.net
[3] Department of Computer Information & Communication Engineering, Catholic University of Daegu, Kyungbuk, 712-702, South Korea
hangkon@cu.ac.kr

Abstract. On every new development new issues are taken into account to cope with either environmental or the stakeholders needs that are evolving over time. Several approaches have faced this problem. Some of them exhibit this evolution ability by using static design/compile-time techniques whereas others introduce this ability in the system at run-time. Nevertheless, in both cases evolving requirements give rise to the need of adaptability which is inherent to every software development. This paper sketches our work in this field in which we are concerned about the MDSAD (Model Driven Software Architecture Development) methodology to guide the reflexive development of architectures from the software requirements. In particular, we are detailing the first step of this methodology, i.e., the definition of the goals model whose constituents are the fundamental basis for the overall process defined in MDSAD proving its suitability for obtaining traceable architectural models. It provides our work either to its ability to specify and manage positive and negative interactions among goals or to its capability to trace low-level details back to high-level concerns.

1 Introduction

Dynamism and evolution are currently two main software concerns. On every new development new issues appear related to the customization of the software. They attempt to accomplish either environmental or the stakeholders, needs that are evolving over time. How to overcome deficiencies and limits exhibited by traditional development methodologies is a clear challenge that has been faced by several techniques. Some of them try to introduce this evolution ability by using static design/compile-time techniques whereas others introduce this ability in a way that evolution appears at run-time. Aspect Oriented Software Development (AOSD) [1] is related to the first type of techniques providing advantages in expressiveness by the separation of concerns. Both functional and non-functional needs, such as performance or Compatibility of the system's behaviour can be separately acquired and specified across the development lifecycle. A methodology MDSAD (Model Driven Soft-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 916 – 926, 2005.

ware Architecture Development) illustrates a process to concurrently define requirements and architectural artifacts to obtain architectural artifacts that provide ability for dynamism and evolution by means of AOSD and reflexive techniques. In this work, we are detailing the first step of this methodology, i.e., the definition of the goals model whose constituents are the fundamental basis for the overall process defined in MDSAD.

2 The MDSAD Methodology

This methodology is concerned with the definition of software architectures from functional (ER) and non-functional requirements (NFR). With this aim, it provides the analyst with the guidance along the process from an initial set of requirements to an architectural instance. MDSAD is a not a product oriented approach but a process oriented one. It does not describe a set of metrics to determine if the requirements, functional and non-functional, are implemented into the final software. On the contrary, requirements and architectural design decisions are defined concurrently in such the way that every decision is make to achieve a given requirement. Iterative and incremental architectural development is the two main characteristics taken into account in its conception. This concurrent description of requirements and software architecture is not as straight-forward as defining all the artifacts at once. Therefore, it seems natural to provide the analyst with an iterative process in order to incrementally develop the software artifacts. It allows him/her to reason on partial models and focus on different architectural views. MDSAD iterates over a set of five steps(Figure 1) which are described next.

Step 1. Goals Model (GM) Definition
In this step, the set of goals to be accomplished by the software system are defined by means of both functional and non-functional requirements. An informal set of requirements, stated in natural language, is the input to trigger the model definition (see section 3). The elaboration of the Goals and Scenarios Models (see step 2) are two intertwined processes. The GM is operationalized at the next step by means of the Scenarios Model (SM). Furthermore, the analysis information provide by the SM helps us to refine and identify new goals at the next iteration. Therefore, both models are coupled with a meaningful advantage in terms of traceability.

Step 2. Scenarios Model (SM) Definition
The analyst has to identify the set of scenarios which operationalize the established goals and compose them to form an iterating and branching model of the system's behaviour. Each scenario depicts the elements that interact to satisfy a specific goal and their level of responsibility in achieving a given task. These elements are *shallow-components*, i.e., a rough description of the components that appear into the final software architecture. Use Cases(UC) and Message Sequence Charts[2,3,4,] are employed for the construction of the SM. The former provide us with a visual metaphor notation for scenarios, where the identified shallow components are the involved actors, Nevertheless, when we speak about Software Architecture [5], we

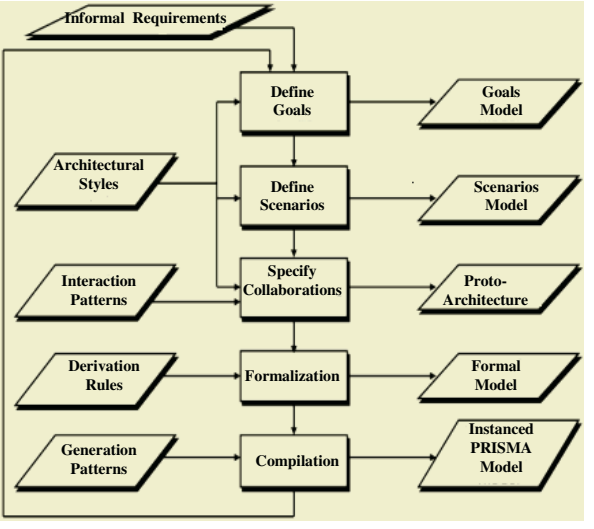


Fig. 1. MDSAD

MSCs identify the coordination structure through the temporal sequences of interaction events. Several alternative scenarios can operationalize the same goal, just like several alternative programs can implement the same specification. In order to offer the best approach, they have to be analysed caring about conflicts that may arise among the operationalized goals.

Step 3. Collaborations Definition
The third step is aware of the collaborations among the identified shallow components. The collaborations realize to UCs through collaboration diagrams. The objective is to obtain a proto-architectural styles, interaction patterns and the architect's ability. Additionally, the connectors are defined according to the components interactions. These first class citizens are required to achieve a loose coupling between the components.

Step 4. Formalization
A semantic check and analysis of the models and the proto-architecture is required to identify eventual conflicts, e.g. different scenarios resulting in incompatible architectural configurations, and obtain the best alternative to avoid (or minimize) them. The artifacts, defined in the previous steps, are specified in the APL language [6], whose interpreter helps us to validate them.

A set of derivation rules are provided to generate a formal specification from scenarios, goal and collaborations, in order to assist and speed up the formalization process. This specification is validated and then used for an automatic compilation process in the next step.

Step 5. Compilation
Using the formal model and a set of generation patterns, the translation from the requirements model to an instantiated PRISMA model specification is accomplished.

ABCD [6] Architectural Models has been selected as the reference model to specify architectural elements by integrating Component-Based Software Development (CBDS) [7] and AOSD [1] approaches. This architectural model is able to be compiled into a concrete target system preserving its compiled into a concrete target system preserving its reflexive properties. This step is assisted by the engineer in order to refine the architectural elements. As MDSAD is intended to be iterative and incremental, a feedback is provided from Step 5 to 1. In this way, all the models are up-to-date all over the process.

3 Goals Model: An Artifact for RE

In the context of Requirements Engineering, the *Goal-Driven Requirements Engineering* paradigm [8] has proven useful to elicit and define requirements. More traditional systems analysis tool, such as UCs, focus on establishing the features (i.e. activities and entities) that a system will support. Additionally, we have to bear in mind that architectural models are a bridge between requirements and the system-to-be providing us with a lower abstraction level. They are used as intermediate artifacts to analyse whether the requirements are met or not. Therefore, this paradigm has two advantages that make it appropriate to systematically guide the selection among several architectural design alternatives:

- Its ability to specify and manage positive and negative interactions among goals [9, 10] allows the analyst to reason about design alternatives.
- Its capability to trace low-level details back to high-level concerns is very appropriate to bridge the gap between architectural models and requirements.

These are the main reasons why the *Goals Model* has been introduced as an MDSAD artifact to identify and describe the users' needs and expectations, their relationships and how these can be met for the target system.

3.1 Building Components for the Goals Model

It is stereotyped as *Functional* or *Non-functional*, according to the type of need or expectation it refers to:

- *Functional goals* describe services that the system provides, i.e., the transformations the system performs on the inputs.
- *Non-functional goals* refer to how the system will do these transformations, for instance, in term of performance, adaptation, security, etc. We are highlighting them because they are especially meaningful in terms of software quality [11].

Additionally, other aspects have to be stated when a goal is defined. For instance, a set of *preconditions* and *postconditions* has to be identified, preconditions establish which situations must hold before some operation is performed. Postconditions define the situations that have to be achieved after some operation. Their evaluations help us to determine the best design alternative among those that satisfy the postconditions for the established goals. Moreover, each goal has to be classified according to its priority, from *very high* to *very low*, for the system-to-be. This classification helps the

analyst to focus on the important issues. These priorities can arise from several factors: organizational ones when they are critical to the success of the development, constraints on the development resources, etc. A textual notation for goals definition, along with a visual notation (Figure 2), is provided to deal with the previous aspects:

GOAL GoalName
ID identifier
TYPE [functional | Nonfunctional]
DESCRIPTION ShortDescription
PRIORITY [Very High | High | Normal | Low, Very Low]
AUTHOR autherName
PRE conditions
POST conditions

Aside from goals, other components for the GM are the *Operationalizations*. When an analyst has refined the initial set of goals, he/she must offer a set of solutions that allow the system to achieve the established goals. These solutions provide architectural design choices for the target system which meet the user's needs and expectations. They are called operationalizations because they describe the operation of the system, i.e., the system behaviour, to meet functional and non-functional requirements. Their textual notation, together with their visual notation in figure 2, is:

OPERATIONALIZATION *Opname*
ID identifier
DESCRIPTION ShortDescription
AUTHOR authorName
PRIORITY [Very High | High | Normal | Low, Very Low]

It can be noticed that there is no section in the textual notation to define the alternative solutions to satisfy a given goal. In contrast, it is in the SM where these solutions are expressed. However, entities known as operationalizations are introduced in the GM to represent conceptually each solution so that relationships among the different alternatives can be established within the GM. Operationalizations imply a coupling between the SM and the GM, and traceability between operationalizations and a specific view of the SM is achieved by means of *identifiers*. Additionally, operationalizations have a special property called *Priority* to express how meaningful a given solution is for the system, i.e., its relevance for the system development.

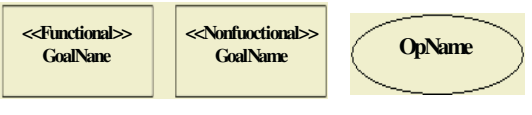


Fig. 2. Visual notation for goals and operationalizations

3.2 Relationships: An Element in the Refinement Process

The stated components, goals and operationalizations, are inter-related by means of a set of *relationships.* They are in charge of gluing the different elements to complete

the model and enhance its cohesion. Moreover, their relevance is not only restricted to this gluing but also they allow the analyst to introduce the rationale of the system design. The decomposition of goals or how an operationalization positively or negatively contributes to a goal, can be defined via relationships. There are two types of refinements that can be applied: intentional and operational. The former describes how a goal can be reduced into a set of subgoals via *AND/OR* relationships. The latter depicts how a set of solutions address a goal by means of *AND OPERATIONAIZE/OR OPERATIONALIZE* relationship. Both, components and relationships are structured as an acyclic goal graph, where the refinement is achieved along the structure, from the higher to the lover level, by applying intentional and operational refinements.

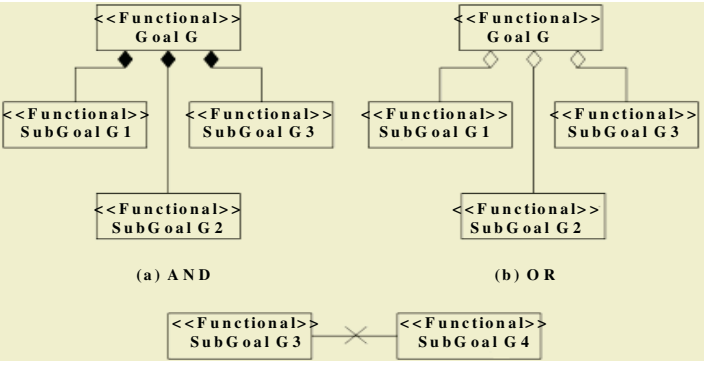


Fig. 3. Visual notation for Goals Relationships

An AND relationship (Figure 3 (a)) between a goal *GoalX* and a set of sub-goals G_1, …, G_N is established if the sub-goals have tobe satisfied in order to satisfy *GoalX*. Its textual notation is described as:

AND(*Goal1* {, *GoalN*}) ACHIEVE *GoalX*

A goal *GoalX* is related to a set of sub-goals G_1, …, G_N via an *OR* relationship (Figure 3 (b)) if *GoalX* is satisfied if at least a sub-goal is satisfied. Its textual notation is described as:

OR(*Goal1* {, *GoalN*}) ACHIEVE *GoalX*

Every goal, which is too coarse-grained to be directly addressed by a solution, is refined in a set of subgoals which are a decomposition of the original one. Whenever a sub-goal is needed to achieve a goal, and AND relationship is established between them. On the contrary, if the sub-goal may optionally appear, then an OR relationship is established between them.

Additionally, a CONFLICT relationship (Figure 4) can be set up among two goals if an incompatibility appears between them, in other words, whenever the satisfaction of a goal prevents the satisfaction of another goal. It is described as:

CONFLICT Goal1, Goal2
DESCRIPTION shortDescription

An operational refinement deals with operationalizations and goals. The alternative solutions for each goal are established by means of this decomposition. There can be a

large number of valid operationalization methods that are applicable to a goal. In such a case, it is up to the analyst to examine the impact of such methods on other types of requirements and decide on what and how many operationalizating methods must be applied via AND OPERATIONALIZE/OR OPERATIONALIZE relationships. An *AND OPERATIONALIZE* relationship (Figure 4 (a)) relates the set of mandatory solutions for a goal. On the other hand, whenever several alternative solutions can be provided for a goal, the analyst can introduce them in terms of the *OR OPERATIONALIZE* (Figure 4 (b)) relationship. Their textual notation is:

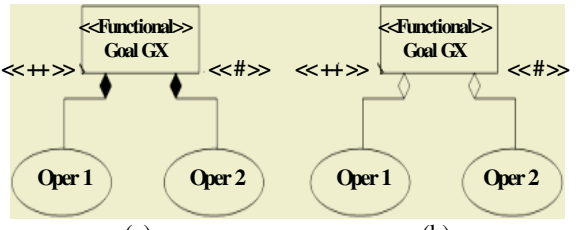


Fig. 4. Visual notation for Operationalize Rel

It is shown in this description that a set of symbols [++ | + | # | - | -- |] are used to stereotype the relationship. They denote how an operationlization collaborates to achieve a goal. Symbols ++ and + describe a positive collaboration, i.e., it provides a sufficient or partially sufficient solution, respectively, to satisfy the related goal. On the other hand, symbols – and – describe a negative collaboration, i.e., the operationalization prevents or partially prevents, respectively, the satisfaction of the related goal. The # symbol is introduced to specify operationalizations whose impact (positive or negative) is unknown at the moment and is the default stereotype.

OPERATIONALIZE relationships do not only relate operationalizations to goals but also to other solutions. It is used to refine the operationalizations down to other simpler ones, i.e., to describe how a solutions. AND OPERATIONALIZE relationships relate a set of operationalizations whose composition satisfies another one. OR OPERATIONALIZE relationships are used when alternative sub-operationalizations exist. Again, these relationships are positively or negatively stereotyped.

3.3 Goals Model and Aspects: How to Solve the Tangled Components

As was stated above, AOSD provides a mechanism to specify and relate concerns (solving the cresscutting). In this respect, these can be handled adequately avoiding that they appear tangled over the system and providing significant advantages in terms of understanding, maintenance and evolution. Concerns may or not result in concrete aspects in the end system. We need to bear in mind that PRISMA, the selected model to instantiate architectures, follows an aspect-oriented approach. Components (and connectors) are defined by a gluing of aspects, i.e., they are the *melting pot* for them. Therefore, it is necessary to identify what aspects are and which components(and connectors) are crosscut by them. It is difficult to answer this question

because there is not a well-established understanding of the notion of concern, In AOSD, the term concern is understood as the so-called functional and non-functional goals. This answer may help to understand why the GM is so important in this work. The concerns are being established when we are identifying the goals (and sub-goals) of the system, by means of the iterative process stated above. The concerns of the system are specified as nods of the goals graph. Moreover, the GM assists us in identifying one of the main issues in AOSD: which concerns are being crosscut by others, i.e., which concerns are in conflict. The initial set of goals(concerns) could seem that interact in synergy, without any conflict. But the iterative refinement process can help us to identify if crosscutting appears. In this case, a *conflict relationship* is established between each pair of confliction sub-goals. A trouble may appear when several sub-goals show an equal priority. In such a case, the analyst must define a trade-off among them. Several works [12] have stated that concerns may or not be aspects in the end system. This is determined within MDSAD by means of the SM which scatters the aspects over the components and connectors by using the UCs and the MSCs. The UCs identify as actors the components, the subsystems or systems which are collaboration to operationalize a goal, i.e., a concern. As a result, the UCs where the involved actors are only components will be identified as aspects in such a way, the set of UCs, where a component G_1 is involved, is the set of aspects that compose G_1.

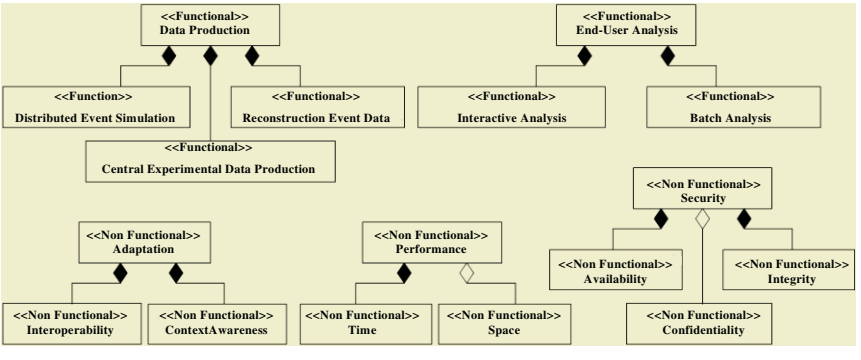


Fig. 5. Iterated Goals Model

4 Applying MDSAD to a Case Study

Iteration 1 – Step 1 Goals Model. The initial set of Requirements is presented as five nodes in Figure 6 obtained from the previous sentences. *Data production* deals with experiments design. *End-User analysis* with the required process to evaluate generated data. The *Performance* requirement is due to high efficiency required from this kind of applications. *Adaptation* appears because of the self-configurability or auto-adaptability that is needed. Finally, *Security* is introduced to deal with information protection.

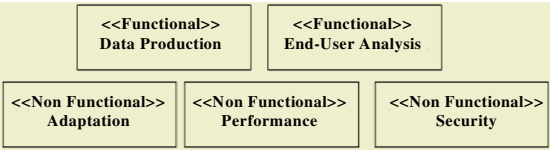


Fig. 6. Original Goals Model

Iteration 2 – Step 1 Goals Model Definition. Several subgoals can be identified in order to satisfy the previous ones, as it is shown in Figure 7. For instance the *Adaptation* goal can be intentionally refined up to *ContextAwareness* and *Interoperability.* Because both sub-goals are required to achieve the Adaptation goal, and AND relationship has been established between them. In a similar manner, *Time* and *Space* are the intentional refinement of Performance. However, Space and Performance are OR-related because there is no high restriction is this sense.

Iteration 3 – Step 1 Goals Model Definition. Some of the identified goals can be operationally refined. *ContextAwareness* and *ContextAcquistion* are AND OPERATIONALIZE-related because the latter is a required solution for the former (Figure 7).

Iteration 3 – Step 2 Scenarios Model Definition. In order to operationalize the subgoal *ContextAware,* three actors can be identified: *GridComputingNode, Cluster Manager* and *LoadBalancer*. The first one has to gather context information and distribute it to the LoadBalancer and the ClusterManager in order to optimize the resources. UC and the MSC, for this sub-goal are showed in Figure 8.

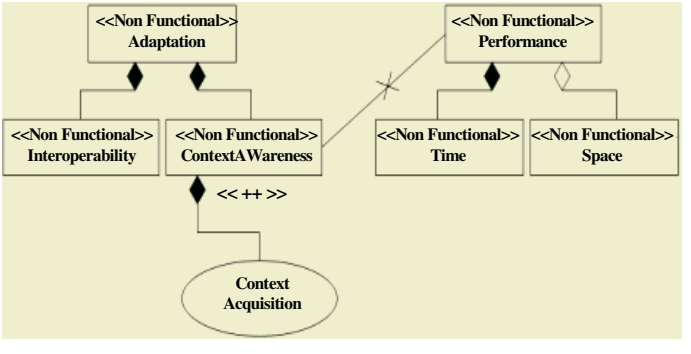


Fig. 7. Iterated Goals Model (part of)

Iteration 3 – Step 3 Collaborations Definition. Because MSCs are closely related to collaboration diagrams it is very easy to get the latter. The result for the ContextAware goal can be observed in Figure 8. It is appreciated that the three components (GridComputingNode, LoadBalancer and Cluster Manager) have to be disguised with this goal (aspect). Moreover, to get a loosely coupling between the involved components, a connector is defined to aggregate and analyse the relationship between them.

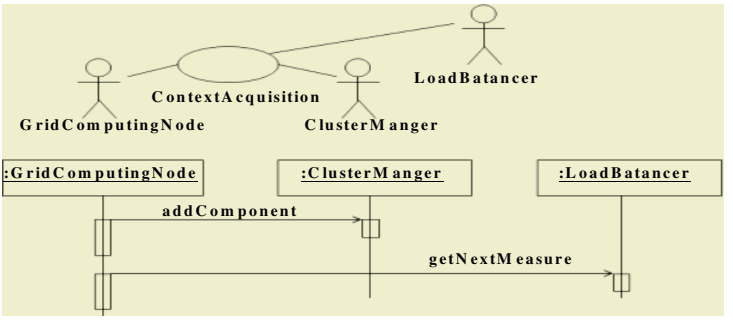


Fig. 8. Iterated Scenarios Model (part of)

Iteration 3 – Step 4 Formalization. The defined models will be formalized in this step, but a detailed description is outside of the scope of this paper.

5 Conclusions

Summarizing, in this paper we show how to address the iterative development of requirements and architectures during the development of software systems. A methodology, MDSAD, that guides the analyst, from an initial set of requirements to an instantiated, architecture has been presented. It uses the strength provided by the coupling of scenarios and goals to systematically guide through the iterative process. Moreover, it allows the traceability among both artifacts to avoid lacks of consistency.

Additionally, requirements and architectures can evolve iteratively and concurrently, in such a way that running-systems can dynamically adapt their composition and/or topology to meet their evolving requirements. It is granted by the reflexive properties of PRISMA and by the relationships between requirements and architectures that have been established along the process.

There are questions and problems that remain open: the completeness of each scenario and the completeness of the set of scenarios; the compatibility and consistency of the architectural styles; the creation of the architectural structure or its transformation on the basis of NFRs. All these subjects constitute our future work.

References

1. Maja D'Hondt and Theo D'Hondt, "The Tyranny of the Dominant Model Decomposition," OOPSLA Workshop on Generative Techniques in the Context of Model-Driven Architecture, Seattle, Washington, November (2002)
2. Tzilla Elrad, Omar Aldawud, Atef Bader, "Aspect-Oriented Modeling: Bridging the Gap between Implementation and Design," Generative Programming and Component Engineering (GPCE), Pittsburgh, Pennsylvania, October (2002) 189-201
3. Robert Filman and Dan Friedman, "Aspect-Oriented Programming is Quantification and Obliviousness," OOPSLA Workshop on Advanced Separation of Concerns, Minneapolis, Minnesota, October (2000)

4. David Frankel, Model Driven Architecture: Applying MDA to Enterprise Computing, John Wiley & Sons (2003)
5. Jeff Gray, Ted Bapty, Sandeep Neema, and James Tuck, “Handling Crosscutting Constraints in Domain-Specific Modeling,” Communications of the ACM, October (2001) 87-93
6. Jeff Gray, Janos Sztipanovits, Ted Bapty, and Sandeep Neema, Two-Level Weaving to Support Changeability in Model-Based Program Synthesis, in Aspect-Oriented Software Development, Addison-Wesley (2003)
7. ISO/IEC 9126 "Information Technology -Software Quality Characteristics and metrics-Part 1,2,3.
8. P. Grunbacher, A. Egyed and N.Medvidovic: “Reconciling Software Requirements and Architectures: The CBSP Approach”. Proc 5th IEEE Int. Symp, Toronto, Canada, August (2001) 202-211, RE, 27-31
9. G. Kiczales, J. Lanping, A. Mendhekar, C. Maeda, C. Lopes, J. M. Loingtier, and J. Irwin, “Aspect-Oriented Programming”, Proc, European Conference on Object-Oriented Programming, Finland. Springer-Verlag LNCS 1241, Jun (1997)
10. Van Lamsweerde. “From System Goals to Software Architecture”, Formal Methods for Software Architecture, LNCS 2804, Springer-Verlag (2003) 25-43
11. Nuseibeh, “Weaving the Software Development Process Between Requirements and Architecture”, Proc. 1st Int. Workshop From Software Requirements to Architectures (collocated ICSE), Toronto, Ontario, Canada, May (2001) 12-19
12. F. Tisato, A. Savigni, W. Cazzola, and A. Sosio, “Architectural reflection realising software architectures via reflective activities”, EDO, LNCS, Berlin (2000)

Parallel and Distributed Components with Java

Chang-Moon Hyun

Department of Computer Game Development , Tamna University,
San 70, Hawon, Seogwipo, Cheju. Korea
cmhyun@cheju.tamna.ac.kr

Abstract. This paper presents an environment for supporting distributed application using shared component in Java within a heterogeneous environment. In particular, It offers a set of components such as Lists, Queues, Stacks that can be shared across a network of heterogeneous machine in the same way as DSM systems. Shared is achieved without recourse to Java RMI or component proxies as in other component systems. An implementation of the environment is provided together with performance timings.

Keywords: Java RMI, Distributed Component, Distributed Systems, Shared Distributed Memory.

1 Introduction

Computers became almost a necessity in every domain due to their ability to manipulate a large amount of data. Over two decades, from 1960 to 1980, mainframes were predominant and were used to achieve any required processing. They usually consist of a central computer linked to a set of terminals without any processing power. In the same decade of 1980's, distributed systems and Local Area Networks (LAN) have emerged to allow several machines to be linked and thus form a powerful machine which is cheaper and sometimes more powerful than a super mainframe [9].

Computer networks allows as well linking heterogeneous computers whether they are local or distant geographically. This is one essential characteristic of LAN's and WAN's (Wide Area Networks). In this way, it is possible to link computers of different architectures running different operation systems, but communicating with a common protocol such as TCP/IP. Tanembaum [9] outlined several advantages of distributed systems, over centralized systems such as: economy, performance, cost, incremental, scalability, flexibility. For this reason there is a net swift from centralized systems to distributed ones.

However, software for distributed systems is not an easy task. Writing programs on distributed systems requires knowledge on computer networks, communication protocols, concurrency, synchronization, distribution, load balancing … etc. The component technology has proven to be a suitable paradigm for programming distributed systems. Several systems have been devised. At present, the technologies of components and distributed programming are unified giving rise to the distributed compo-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 927 – 937, 2005.

nent technology (DOT) [7]. The web technology is also integrated with the DOT. One good example is Java [2] of Sun Microsystems. Java is a an component oriented language that give support to distributed programming through the concept of distributed components. Characteristics of Java include: portability of Java code on different platforms (e.g. Unix, Windows, Mac), support of web programming. Sharing in Distributed systems is one essential characteristic, but it is a difficult task to achieve. Several DSM systems have been provided [8] but are limited and complex. One interesting issue is to investigate sharing across different platforms, and in particular in a transparent form is the usage of the concept of an component. For this purpose, we devise a library of shared component that can be accessed across any node of a heterogeneous system.

This paper is organized as follows: Section 2 gives the basic concepts and outline existing systems such as Java and DSM. Section 3 introduces our library. Section 4 describes the implementation and gives some performance measures. Section 5 outlines some concluding remarks.

2 Related Works

2.1 Distributed Systems and Sharing

The progress in the technology of VLSI design, cheap processors, and high speed network made it possible to interconnect several machines to form very powerful computer systems. These machines have the componenti of supporting applications which consume a lot of processing power as in the case of visualization applications, CAD/CAM, Distributed database, etc [9]. There are two types of systems: multiprocessors and multicomputers. Multiprocessor systems exhibit a semantic of previsible performance, but have a very complex hardware and difficult to construct. Programming these machines is simple; the programs share a common share a common address space. Multicomputer systems consists of a set of processors; each with its own memory and is connected to another through a high speed network. These systems are easy to built physically but the software is more difficult since it requires the programmer structuring its application as intercommunicating processes. With the intention of combining characteristics of multiprocessors and mulicomputers, DSM (Distributed Shared Memory) systems appeared. These provide a global virtual memory accessed by all processors. However, the software is difficult to built and has an imprevisible performance [7].

2.2 Component and Distribution

The component oriented technology has been pointed to be a promising one to control the complexity generated by distributed systems. The concepts of component orientation such as modularity, access to data through the interface makes the use of components appropriate to model distributed systems. This approach has been enforced and well accepted with the development of several distributed environments that use the concept of components (e.g. CORBA of OMG[12], DCOM and OLE of Microsoft.

Another aspect that should be taken into consideration is that component oriented applications are programmed in terms of communicating components which is a natural form to program distributed systems [7,10]. The technology of distributed components (DOT) is revolutionizing the programming methodology in distributed systems. In fact it includes three technologies of: components, distribution and Web. The component technology (OT) was introduced in the seventies by Adele Goldberg and Alan Kay in a language called Smalltalk. In thin model, components encapsulate structures and behaviour; each is an instance of a class that belong to a class hierarchy. The OT has matured quickly and became well accepted as a technology that encapsulates complexities, improves maintenance, promotes reuse and reduces the cost of life cycle of software.

The focus of DoT has been to do the atmosphere of more transparent computation with regard to the use of computation engines and computation components. With the Web technology (WT), born in the nineties, it caused a quick explosion in the use of the Internet. In 1995, Java came and gave more initiative for the usage of DOT. The key concept behind DOT is the interconnection. Many abstraction levels can be used to describe the connection between machines and net [7]. DOT brought the concept of components with the notion of transparency of the distributed computation [10]. Components became a skilled technology for distributed processing.

2.3 Java

The Java language, of Sun Microsystems, is revolutionizing programming of distributed applications. The use of Java in WEB pages, turns them more flexible and interactive, facilitating small and big companies put their services and products on-line, besides enriching the form of knowledge presentation in the Internet.

The language Java [6] was projected for the development of systems that you run in distributed contexts and heterogeneous environments. It offers safety to applications, consumes few resources of the system and can run on several hardware platforms and software [5]. Java is not just a language, but it has an entire philosophy of advanced software engineering. The Java system presents several advantages, because it is simple, component oriented, robust, portable and architecture neutral; it adopts modern methodologies (client/server), interpreted, besides it is multithread.

One of the most interesting characteristics of Java is its portability. This facilitate the development of applications for heterogeneous environments in a network. This implies that Java application can run on different hardware architectures without need of fittings. Instead of creating an executable code for a certain machine, Java generates bytecodes, that are interpreted and executed by a virtual machine (neutral architecture), that is written for several types of platforms. Although Java is interpreted, it may present good performance, since the creation of the code does not have the cycle of compile-link-load-test-crash-debug of other languages; now it is simply complied and run. One in the most usual ways of communication among processes in a distributed system is the remote procedures call (RPC) [12]. Java is one language that uses this mechanism and uses RMI (Remote Method Invocation). Figure 1 illustrates the RMI system. The RMI system is divided into four parts [12]:

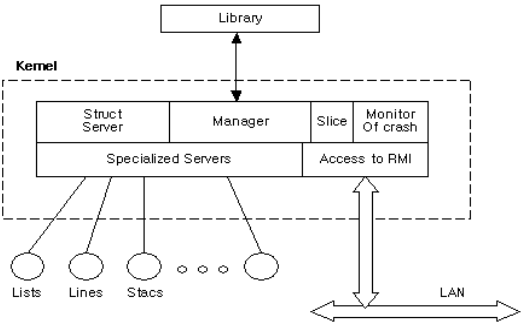


Fig. 1. RMI Architecture

Stub/skeleton: the stubs are used by clients (proxies) and the skeletons are used for the servers (dispatchers);
Remote reference: reference behavior and invocation (using unicast or multicast);
Transport: configuration and management of the connection;

Distributed garbage collection: reference of remote components. Java supports multithreading. Monitors are used to prevent inconsistencies o data due to multiple accesses which may arise. According to Andrews [1], a monitor is a mechanism of abstraction of data: they encapsulate the representations of the resources and they supply procedures that act in these resources. The mutual exclusion is guaranteed since a procedure is just executed one at a time (similar CCR) and conditional synchronization is obtained through conditional variables.

3 The Model of Our Environment

Our environment provides a collection of components (data structures) that can be shared across a heterogeneous network. It has the following characteristics: Access to elements of components (data structure elements) are done through methods implemented within components. For example, insert to put an element to an component, and delete to remove an element from the component. There is no difference in accessing a local or remote component. Provides method for creating components: local or remote. Components are shared between nodes of the network.

The proposed environment is illustrated in Figure 2. The environment is presented with three heterogeneous computers; however, the system may have any number of nodes. The architecture has the following components:

Component space: Each node has its own component space or local repository where the components are stored. The sum of all repositories is called the global repository which is managed by the ma*nager component*. It is implemented by the language Java through RMI.

Kernel: Each node has its own kernel. The role of the kernel is to keep track the allocation of data structures. The kernel provides functions to create and remove components. It also allows other kernels to access local data structures in a safe and efficient form. A kernel consists of sub components: server structures, specialized serv-

ers, slice and RMI access. The server structure is responsible for creating and deleting the structures. The specialized server (e.g. List server, stack server, queue server) are responsible for servicing specific component requests. The slice component is responsible to compute the load in each node every five minutes. The load is used by the manager to best distribute components over the network. There are two types of kernels: a kernel with a manager figure 3-a and a kernel without a manager figure 3-b. In a network there is only one kernel with a manager all of other nodes consist a one kernel without a manager. The manager is responsible for decision such as where an component should be created. This decision is taken depending on the load of each node. For this purpose, the manager maintains a table of information containing a load of each node, estimate time slice of each process, amount of memory available, machines addresses (Ips). The crash monitor component verifies which machine has crashed. If it happens, the manager is called to take appropriate action such as taking the machine out of the system.

Library: provides a set of functions (methods) for manipulating the components (i.e. data structures).

The functions are provided as part of the interface of each component. There are several types of components:

Lists: Lists can be created in three forms: local, remote and specified location. A local list is created on the same node where the create method is invoked. The remote list is created on a remote node chosen by the system. A list created in a specified location is created on a node specified by the user (IP specified).

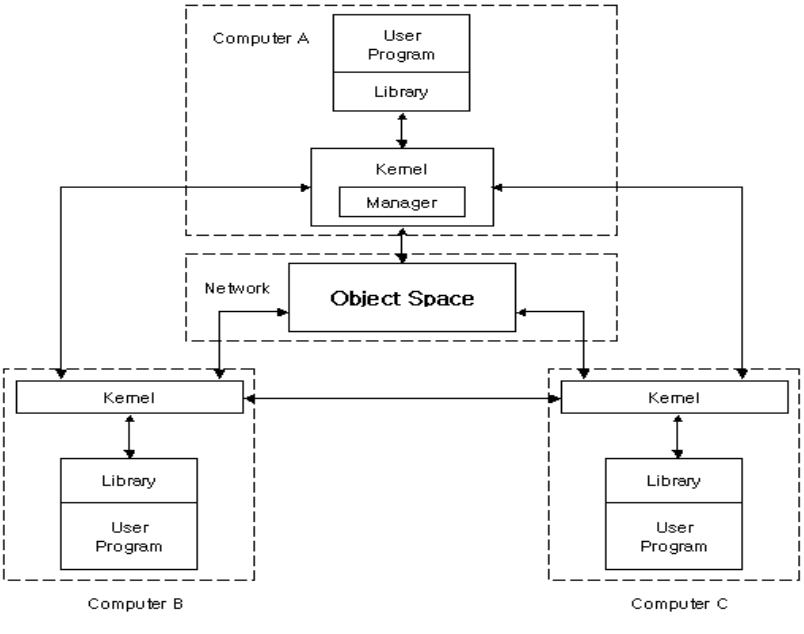


Fig. 2. Architecture of our Environment

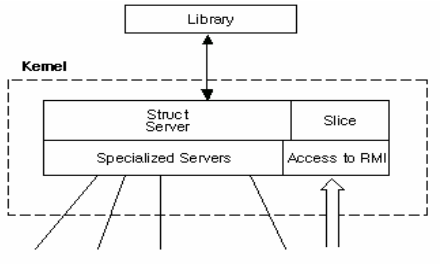


Fig. 3(a). Kernel without manager

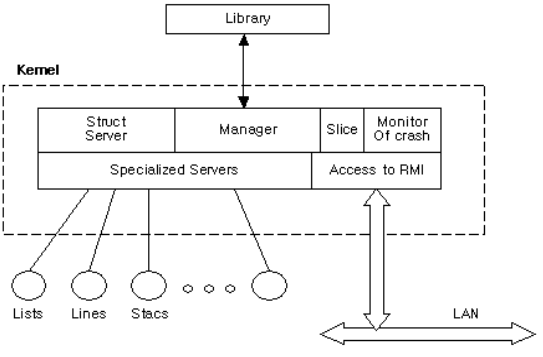


Fig. 3(b). Kernel without Manager

Stacks: three types of tacks may be created as well in the same manner as component Lists. However, the Stacks are accessed in LIFO through operations push and pop.

Queues: Components queue may be local remote or location specified. Accesses to such component is in FIFO through operation of insert and remove.

3.1 Manipulating Components

The usage of the library is very simple. Creating and invoking methods on shared component can be done in the same way as in Java. A shared component may be created as local, remote or with specified location. Constructors are provided for appropriate types of components. List_local(Name) for initialising a local component with *Name*, List_remote(Name) to initialise a remote component with *Name* and List_remote(Name, IP) to initialise a remote component specified at the node IP with Name. A programmer may create a remote component and initialise it as follows:
List_remote list – new List_remote (“hello”);

This has the effect of creating a remote component list of type List and initialise it with an element *hello*.Manipulating an component is also simple. For example, inserting an element world to the created component may be done as follows:

list.insert("world"); // insert the an element world to the component List list. As it can be seen, programming with shared component is very simple. The programmer does not need to get involved with details of communication, RMI, component localization etc.

4 Evaluation of the Proposed Environment

The proposed environment is implemented on a network of three computers running different operating systems.

Platform A: consists of a PC, Pentium-S 150 MHz processor, running Windows 95.
Platform B: consists of a PC, Pentium II MMX 266 processor, running Linux
Platform C: consists of a PC, Pentium II MMX 266 processor, running Windows 98.
The three platform were running the Java VisualAge 3.0 of IBM.
Test have been made to measure:
The times for creating a kernel with a manager
The times for creating a kernel without a manager
The times for creating data structures
The times for accessing data structures
The tests measurements were developed with a small application that require to share data stored in shared components of type Queue, Stack, and List. It was observed that the measurements were close with any shared component type. For this reason we present, results from experiments with shared components of type List.

The data shared by the application are component instances of the following class:

```
Public class Info_Func Implements Serializable

{
String name;
Int age;
Boolean sex;
String identity;
String cpf;
String address;
String profession;
Int registers;
Int wage;
}
```

Methods of this class are omitted intentionally. The class is defined as Serializable for marshalling and un marshalling purposes. In the first instance, we proceeded with test measurements for creating and accessing the kernels. Tables 1,2 and 3 exhibit the results obtained. Note that the manager in all tests is held in platform C.

Note also that the tests were carried out several times. The values presented in the tables are average values.

Table 1. Results obtained with Kernel of the platform A (IBM-PC Windows 95)

KERNEL Timings (T in ms)		
Action	T $_{minimo}$ (ms)	T maximo (ms)
creation of Kernel *	2.100	33.679
access to Kernel	29	54

* without manager

Table 2. Results obtained with Kernel of the platform B (IBM-PC Linux)

KERNEL Timings (T in ms)		
Action	T $_{minimo}$ (ms)	T maximo (ms)
creation of Kernel ^	5.397	44.770
access to Kernel	14	41

* without manager

Table 3. Results obtained with Kernel of the platform C (IBM-PC Windows)

KERNEL Timings (T in ms)		
Action	T $_{minimo}$ (ms)	T maximo (ms)
creation of Kernel ^	1.682	11.904
creation of Kernel ^^	24.814	32.857
access to Kernel	16	32

We can notice that there are differences in timing concerning the creation of the Kernels on each platform. This is due to the differences in platforms and the differences in softwares. However, the measurement for accessing the kernels, the timings and very close (i.e. all are inferior to 50 ms). The second serie of measurements if concerned with the manipulation of shared components. In particular, creating a shared list, and accessing it. Tables 445 and 6 illustrate the appropriate timing. Analysing the results shown in tables 4, 5 and 6, we can notice that local access are cheaper then remote access. This is logical in that remote method invocations are more expensive than local invocations. We can notice as well that accesses with a specified location are cheaper than accesses without any specified location. This is due to the fact that if a location is specified, then there is no need to consult the manager of where about is the list. Table 5 – Results obtained from treatment of a list in platform B.

Table 4. Results obtained from treatment of a list in platform A

DATA STRUCTURE: LIST Timings (T in ms)		
Action	T $_{minimo}$ (ms)	T maximo (ms)
creation of a list_local	1,506	1,552
Insert_front() *	13	38
Insert_back() *	15	28
Remove_front() *	18	1401
Remove_back() *	17	19
Creation of a list_remote	1,943	26,103
Insert_front() **	12	60
Insert_back() **	12	28
Remove_front() **	14	1,401
Remove_back() **	14	1,423
Creation of a specified list_remote	10,420	11,841
Insert_front() ***	18	56
Insert_back() ***	17	57
Remove_front() ***	17	20
Remove_back() ***	18	24

* *executed on a local list*
** *executed on a remote list*
*** *executed on a specified remote list*

Table 5. Results obtained from treatment of a list in platform B

STRUCTURE OF DATA: LIST Measure parameter: time (T)		
Action	T $_{minimo}$ (ms)	T maximo (ms)
creation of a list_local	842	1,414
Insert_front() *	7	28
Insert_back() *	8	22
Remove_front() *	9	79
Remove_back() *	8	39
Creation of a list_remote	668	20,523
Insert_front() **	8	25
Insert_back() **	8	13
Remove_front() **	8	28
Remove_back() **	8	18
Creation of a specified list_remote	2,634	20,998
Insert_front() ***	9	59
Insert_back() ***	12	22
Remove_front() ***	13	24
Remove_back() ***	15	1,141

Table 6. Results obtained from treatment of a list in platform C

DATA STRUCTURE:: LIST Timing: (T in ms)		
Action	T mínimo (ms)	T maximo (ms)
creation of a list_local	324	412
Insert_front() *	9	26
Insert_back() *	9	14
Remove_front() *	10	12
Remove_back() *	10	13
Creation of a list_remote	306	3,650
Insert_front() **	9	1,939
Insert_back() **	8	297
Remove_front() **	9	18
Remove_back() **	11	18
Creation of a specified list remote	5,984	6,747
Insert_front() ***	9	28
Insert_back() ***	13	25
Remove_front() ***	14	23
Remove_back() ***	12	23

* *executed on a local list*
** *executed on a remote list*
*** *executed on a specified remote list*

5 Conclusion

In this paper, we have presented an environment for supporting distributed applications within a heterogeneous network of machines, In particular, if offers a set of shared components that can be accessed from any machine in the same way as any other component without recourse to RIM or component proxies. This will facilitate the programmers task. An implementation using a centralized model has been presented together with some timings. The results can be improved if a decentralized strategy is used. As part of future development, we are in a process of distributing the component physically.

References

1. ANDREWS, Gregory R. Concurrent Programming – Principles and Practice. The Benjamin/Cummings Publishing Company, Inc. Printed in United States of America, 1991.
2. BRUCE, Eckel. Thinking in Java. Prentice Hall PTR. Printed in United States of America, 1998.
3. COULORIS, George e et. al. Distributed Systems – concepts and design. Addison-Wesley, 2ª edition, 1994.
4. DEITEL, H. M. e DEITEL, P. J. Java – How to Program. Prentice Hall. Second Edition. Printed in United States of America, 1997.

5. GONG, Li. Java Security: Present and Near Future. IEEE Micro, 1997.
6. GOSLING, James e HENRY, MCGilton. The Java Language Environment – A White-Paper. JavaSoft. U.S.A, 1996.
7. NORTHROP, Linda e et. al. Distributed Component Technology With CORBA and Java: Key Concepts and Implications. Technical Report. Printed in the United Stated of America. 1997.
8. PROTIC, Jelica e et. al. Distributed Shared Memory: Concepts and Systems. IEEE Parallel & Distributed Technology: Systems & Applications. 1996.
9. TANENBAUM, Andrew S. Modern Operation Systems. Prentice Hall, Inc. Printed in the United States of America, 1992.
10. TANENBAUM, Andrew S., KAASHOEK, Frans. M.. Programming a Distributed System Using Shared Components, Proceedings on 2nd International Symposium on High Performance Distributed Computing, pp 5-12, IEEE CS Press 1993.
11. WALDO, J, WYANG, G. A Note on Distributed Computing. On-line in http://www.sunlabs.com/techrep/1994/abstract-29.html/
12. WOLLRATH, Ann e et. al. Java – Centric Distributed Computing. IEEE Micro, 1997.
13. CRISTIAN, Flaviu. Abstractions for Fault-Tolerance. 13th IFIP World Computer Congress, Aug 28-Seb 2, Hamburg, 1994.

CEB: Class Quality Evaluator for BlueJ

Yu-Kyung Kang[1], Suk-Hyung Hwang[2], Hae-Sool Yang[3], Jung-Bae Lee[2], Hee-Chul Choi[2], Hyun-Wook Wee[2], and Dong-Soon Kim[2]

[1] Dept. of Computer Science, Graduate School, SunMoon University, 100 Kal-San-Ri, Tang-Jeong-Myeon, A-San, Chung-Nam, 336-708 Korea
aquamint99@sunmoon.ac.kr
[2] Division of Computer and Information Science, SunMoon University, 100 Kal-San-Ri, Tang-Jeong-Myeon, A-San, Chung-Nam, 336-708 Korea
{shwang, jblee}@sunmoon.ac.kr
[3] Graduate School of Venture, Hoseo University, 29-1 Se-Chul-Ri, Bae-Bang-Myeon, A-San, Chung-Nam, 336-795 Korea
hsyang@office.hoseo.ac.kr

Abstract. Today most programming courses for beginners use an object-oriented language. This has led to many new learning tools, theories and methods for teaching object orientation. One such new tool is BlueJ which is developed to teach Java and object orientation for beginners. BlueJ does not have any support for quality evaluation for program and this paper describes the development of "CEB", a BlueJ extension for class quality evaluation based on the CK metrics. The CEB is developed as a plug-in for BlueJ and designed to be used at hand by Java programming beginners. This paper presents the BlueJ and its extensions, and describes the design and implementation of the CEB.

1 Introduction

The object-oriented software paradigm has been widely adopted and acclaimed by both industry and educational institutions. In the context of object-oriented software, the class hierarchies are a cornerstone of frameworks i.e. of adaptable and reusable object-oriented architectures and, therefore its quality is very important. Any kind of automated help in evaluating the quality of classes or class hierarchies can thus be of interest and have applications in both industry and educational institutions.

Recently, developers prefer object orientation for application development and education. And developers attach importance to software development of higher quality, not fast software development. Therefore, necessity of quality evaluation metrics suitable to object oriented is increasing. The research for metrics for correct quality evaluation and progress of quality are actively proceeding[1,2,3]. Although there are many object-oriented development environments, lots of them are not suitable for students and instructors to learn and teach OO programming such as Java. In order to settle the above problems, BlueJ was designed and implemented by the BlueJ team at Deakin University,

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 938–944, 2005.

Melbourne, Australia, and the University of Kent at Canterbury, UK for the purpose of teaching object orientation with Java and is provided free[4].

In this paper, we designed and implemented a BlueJ Extension, CEB(Class Evaluator for BlueJ) based on the C&K Suite Metrics[1] for a class quality evaluation of object oriented program. By using CEB, object oriented programming beginner becomes aware of the importance of software quality from the beginning stage of object-oriented programming. Students can easily evaluate a class quality, by using CEB. Therefore, high quality object oriented program can be easily built by using CEB within BlueJ environment.

The structure of the rest of paper is as follows: In section 2 some related works, class quality evaluation metrics proposed by Chidamber and Kemerer as well as BlueJ are briefly presented. In section 3, BlueJ extension is described and design and implement of the BlueJ extension for class quality evaluation is presented. Finally, our research result and future work are presented in section 4.

2 Related Works

2.1 Object-Oriented Metics

Numerous software metrics related to software quality have been proposed in the past and are still being proposed[5,6,7]. Among them, we should concentrate on the essential and basic metrics for object-oriented software. Chidamber and Kemerer defined a suite of six metrics, to qualify the coupling, cohesion, inheritance relationship and complexity of class in an OO system[1].

WMC(Weighted Methods per Class): WMC is the sum of the weights of all methods in the class.

DIT(Depth in Inheritance Tree): DIT is depth of the class in the inheritance tree. In cases involving multiple inheritance, the DIT will be the maximum length from the node to the root of the tree.

NOC(Number Of Children): NOC is number of immediate sub-classes subordinated to a class in the class hierarchy.

CBO(Coupling between Objects): CBO is a count of the number of other classes to which it is couple.

RFC(Response For a Class): RFC is the number of methods in the set of all methods that can be invoked in response to a message sent to an object of a class.

LCOM(Lack of Cohesion in Methods): LCOM indicates as the number of pairs of methods in a class that don't have at least one field in common minus the number of pairs of methods in the class that do share at least one field.

The above CK metrics are essential and fundamental metrics when object oriented programming. Therefore, our research focused the above metrics.

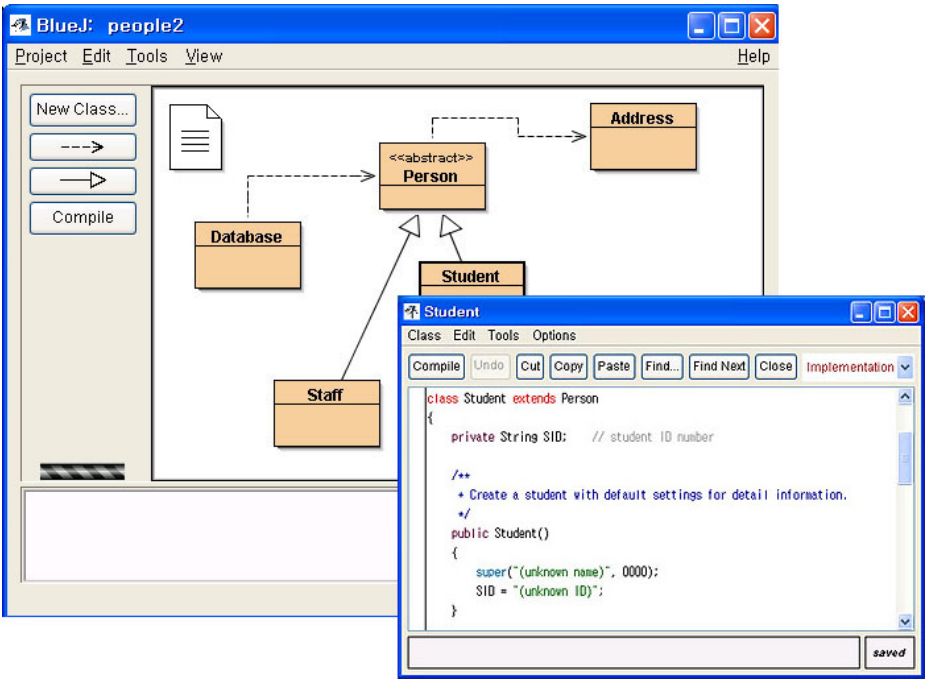


Fig. 1. Screenshot of BlueJ

2.2 BlueJ

BlueJ is a visual programming environment designed to teach object-oriented programming, using Java as the implementation language. It was designed and developed by Michael Kölling and John Rosenberg. BlueJ allows students to concentrate on solving programming problems without becoming distracted by the mechanics of compiling, executing and testing Java programs. It also help students to develop a clear understanding of object-oriented concepts such as objects and classes, message passing, method invocation, and parameter passing. The testing of classes is greatly simplified. The editor is language sensitive and assists with debugging. Moreover, BlueJ supports the following features for the beginner of an object-oriented programming[4](See figure 1):

- fully integrated environment
- graphical class structure display
- built-in editor, complier, virtual machine, debugger, etc.
- easy-to-use interface, ideal for beginners
- interactive object creation
- interactive object calls
- interactive testing
- incremental application development

3 Design and Implementation of CEB

3.1 BlueJ Extension

The BlueJ contains mechanism for extension. That is, BlueJ offers an extension API. Extensions offer additional functionality not included in the core system:

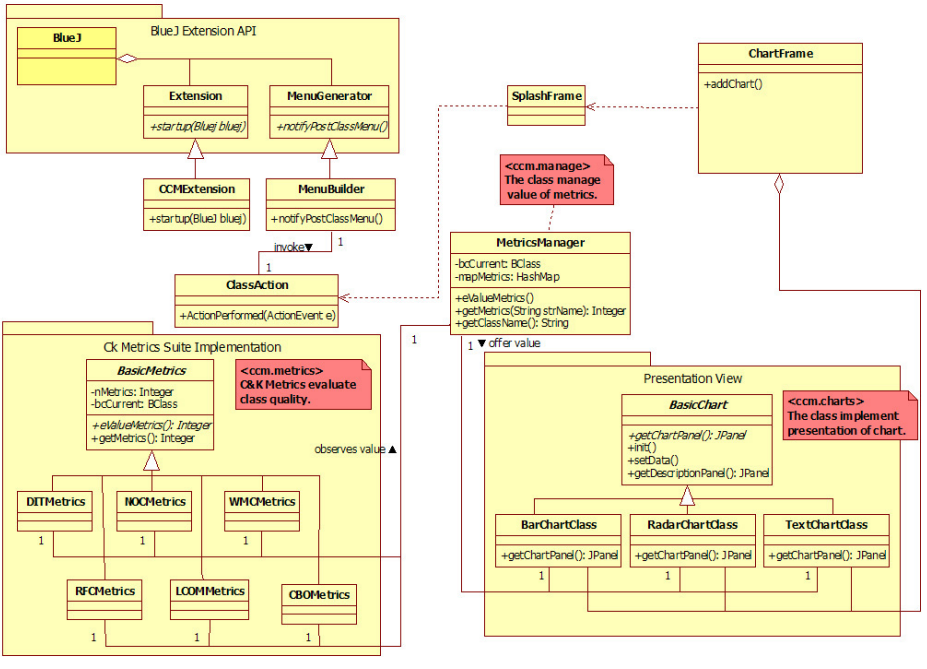


Fig. 2. Brief class diagram for the extension(CEB)

Submitter: Allows submission of projects to specified address via email, ftp, http or file copy. Highly configurable.
Checkstyle: Allows automated checking of coding styles. Coding styles can be specified flexibly in an external file.
Sequence Diagram Editor: An editor that allows the manual creation of sequence diagrams, and links in with BlueJ.
Remote File Manager: Extension to open, edit, and save project from/to a server with a sftp daemon without having to know how to use Unix. The extension tries to mimic traditional file dialogs.
Jeliot 3: Jelot 3 is a Program Visualization application from the University of Joensuu, Finland, and the Weizmann Institute, Israel. This extension integrates Jeliot in BlueJ.
BlueJ Extension Manager: The BlueJ Extension manager checks whether updates for BlueJ or any installed extensions are available. It can also check for availability of new extensions, and lets you install or delete extensions.

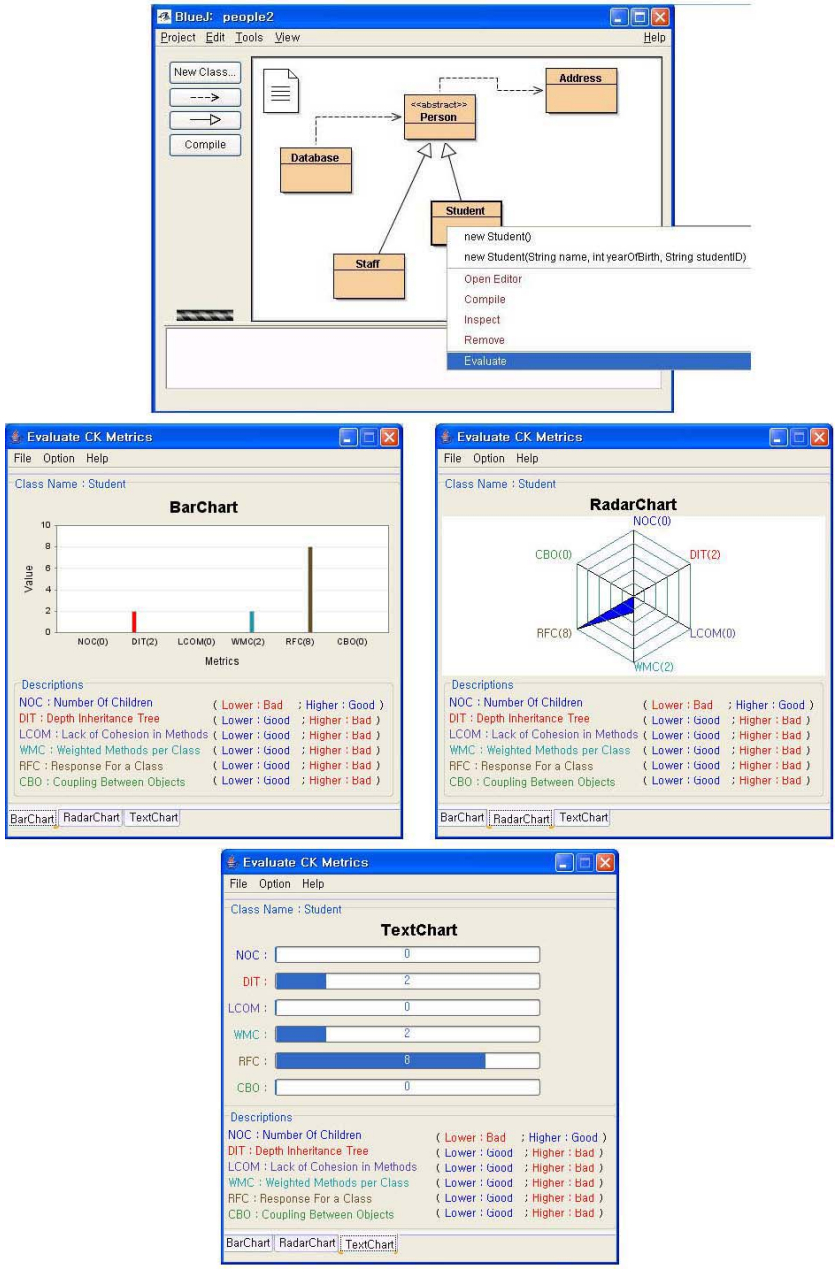


Fig. 3. Screenshots of Class Evaluator for BlueJ

The BlueJ extensions mechanism is a way of adding new functionality to BlueJ as and when it is needed, avoiding user interface clutter and user confusion. The BlueJ Extensions API provides access for extensions to the BlueJ application

via a proxy object, and to the classes and objects which BlueJ is manipulating via a number of wrapper classes. Also, the BlueJ proxy object generates events when the user performs significant actions within BlueJ, and provides access to its projects, packages, classes and objects via a set of wrapper classes. When it is needed new function using this function, we can use added functions[4].

3.2 CEB: Class Quality Evaluator for BlueJ

Our research design and implement the following features for adding class quality evaluation function to BlueJ(See Figure 2).

1. **BlueJ Extension API**
   BlueJ offers an extension API for extension. Our research added a class quality evaluation function and a menu item to the BlueJ by using Extension API.
2. **C&K Metrics Suite Implementation**
   NOC and DIT metrics are simply implemented by using the editor of BlueJ Extension. And, RFC, LCOM and CBO, WMC are implemented by using construction analyzer(JavaCC). In particularly, WMC(Weighted Methods per Class) is implemented by applying Cyclomatic Complexity to method, not by using a factor of weight by hand.
3. **Presentation View**
   In order to display the results of Class Quality Evaluation, we use JFreeChart Component to present Bar Chart, Radar Chart and Text Chart in our Extension.

"MetricsManager" class provides internal information of the target class for C&K Metrics Suite implementation and Presentation View. Based on the information, we can achieve quality evaluation and result presentation via various charts(See figure 3).

4 Conclusions and Future Research

In this paper, we designed and implemented a class quality evaluation function helpful to object oriented programming beginner. Our extension (Class Evaluator for BlueJ) provides "easy to use" for the object oriented programming beginner. We believe, the object oriented beginner can make his/her program much better whenever CEB used. If professor and/or TA use CEB at programming education when they teaching or training, the students can evaluate his/her program by himself. Then, CEB provides immediate feedback for efficient improvement and supplementation for their program. Therefore student can have a good chance to make his/her program with higher quality. Also, instructor can use class quality evaluation function as a supporting tool to checking report and a guideline of programming education.

The CEB is now available from BlueJ website(http://www.bluej.org). We are currently working on the CEB and some BlueJ extensions that help programmers and designers to evaluate the quality of object-oriented program. Various quality evaluation metrics for class and/or class hierarchy and useful design patterns for object oriented program design will be included in the next version.

Acknowledgements

We give special thanks to KBUG(Korean BlueJ Users Group) members for their technical supports.

This research was supported both by the MIC(Ministry of Information and Communication), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Assessment), and by a grant of the Korea Health 21 R&D Project(0412-M102-0404-0002), Ministry of Health & Welfare, Republic of Korea.

References

1. Shyam R. Chidamber and Chirs F. Kemerer, "A Metrics Suite for Object Oriented Design", IEEE Transactions of Software Engineering, Vol.20, No.6, pp.476-493, 1994.
2. M. Dao, M. Huchard, H. Leblanc, T. Libourel, C. Roume, "Towards a Metric Suite for Evaluating Factorization and Generalization in Class Hierarchies", In Quantitative Approaches in Object-Oriented Software Engineering (Workshop ECOOP'01), pp.85-97, 2001.
3. S. Hwang, et al, "A Metrics-Based Approach to the Reorganization of Class Hierarchy Structures", Trancs. of the KIPS, Val 10-D, No.5, pp859-872, 2003.
4. BlueJ homesite (http://www.bluej.org)
5. Victor R. Basil, Lionel C. Briand and WalCelio L. Melo, "A Validation of Object-Oriented Design Metrics as Quality Indicators", IEEE Transactions of Software Engineering, Vol.22, No.10, pp.751-761, 1996.
6. Ralf ReiBing, "Towards a Model for Object-Oriented Design Measurement", Proc. Of 5th International ECOOP Workshop on Quantitiative Approaches in Object-Oriented Software Engineering(QAOOSE 2001), 2001.
7. Tom Mens and Michele Lanza, "A Graph-Based Metamodel for Object-Oriented Metrics", Electronic Notes in The oretial Computer Science, Elsevier Science, Vol.72, No.2, 2002.

Workflow Modeling Based on Extended Activity Diagram Using ASM Semantics

Eun-Jung Ko, Sang-Young Lee, Hye-Min Noh, Cheol-Jung Yoo, and Ok-Bae Chang

Dept. of Computer Science, Chonbuk National University, 664-14 1ga, Duckjin-Dong, Duckjin-Gu, Jeonju, South Korea
{ejko, sylee, hmino, cjyoo, okjang}@chonbuk.ac.kr

Abstract. The Unified Modeling Language(UML) provides rich notations for representing and analyzing architecture and behaviors of systems. Among these notations, UML activity diagram is well-known for describing systems' dynamic behaviors, and it is useful to model business process and workflow. Currently, the UML semantics is informally defined in plain text and it is often unclear, ambiguous or it contains contradictory assertions. It is difficult to present the precise semantics which are taken as important in workflow system with the guide provided by OMG to the UML activity diagram. In this thesis, the alternative approach of using Abstract State Machines to formalize UML activity diagrams is presented. We propose the workflow modeling methodology by applying ASM semantics to the activity diagram. Through the exact definition to formal semantics based on ASM, it is possible to effectively model the workflow.

1 Introduction

Recently in the field of information technology, data-driven information technology has changed to process-driven information technology. The process-driven information technology is based on the common core concept that is the business process. The workflow technology enables the process management to automate.

UML provides the only OMG standard notation for modeling software[1], among these notations, the UML Activity Diagram is well-known for describing systems' dynamic behaviors. The activity diagram is used for workflow modeling including the business process[2].

It must be understandable, general and easy work to develop the workflow modeling. The method of modeling must contain the formal concept and be able to be analyzed. If the activity diagram is used for workflow modeling and is expressed using the semantics, it is represented formally, and is possible to analyze.

Although the OMG provides a textual semantics for UML activity diagram, some ambiguities can be caused due to the lack of rigorous semantics[3]. So far, there are no formal semantics for the activity diagram, informal OMG semantics are not suitable to apply the workflow modeling[4].

The recent UML is short of the formal and common semantic descriptions[5]. The lack of semantic interferes with the use of UML, is due to time spent on understanding the modeling of the artifact.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 945 – 953, 2005.

The research community tries to formalize the UML activity diagram in various ways. For example, pi-calculus[4], FSP(Finite State Processes)[6] is used for representing semantics of activity diagrams. OCL(Object Constraint Languages) [7] can be used to describe the constraints between model elements. Using ASM(Abstract State Machine)[8] formalizes the activity diagram.

This paper presents a UML formalization approach using ASM. In the paper, we provide the ASM semantic expression for workflow modeling based on UML activity diagram.

The rest of the paper is organized as follows: In Section 2 we present some related work and section 3 gives the ASM semantics expression for workflow modeling. Section 4 presents the extension of the activity diagram in workflow modeling. Section 5 presents a case study and evaluation. We end in Section 6 by drawing some conclusions and presenting future work directions.

2 Related Works

The workflow specification describes how the workflow systems work. Therefore it is necessary to define the semantics of the activity diagram and to give the meaning related to the workflow system.

The research community tries to formalize UML activity diagram in various ways. The method of expression of the semantics of activity diagram are OCL(Object Constraint Language)[7], FSP(Finite State Processes)[6], pi-calculus[4], and ASM(Abstract State Machine)[8]. ASM was first proposed ten years ago. Since then, they have been successfully used in specifying and verifying many software systems[9]. The ASM theory is that any algorithm can be modeled at its natural abstraction level by an appropriate ASM[10]. The ASM methodology has the following desirable characteristics. The ASM uses classical mathematical structure to describe states of a computation that distinguished ASM from informal methodologies. ASM uses extremely simple syntax, which can be readable and writable. Existing methods are useful in a specific domain, while ASM is useful in a wide variety of domains.

We define the semantics for workflow modeling based on activity diagram using ASM, which has advantages and provides the definition of the semantics with basic formalization.

3 ASM Semantics Expression for Workflow Modeling

3.1 The Necessary Information for Workflow Modeling

The existing workflow systems are not suitable for modeling the category of systems that involve timeliness as a factor. Indeed, they are not suitable for modeling modern business processes, either. In particular if these are to be suitable for modeling large interconnected real time asynchronous systems one needs an ability to model[Han].

A workflow model made for the anticipated changes in the environment, exceptions, dynamic change, and quantum of workflow that can be chained together to develop the full workflow model, and timing factors. A workflow model can also contain conditions or constraints because activities cannot be executed in an arbitrary

way. Two types of conditions can be applied in a workflow model. They are pre-conditions and post-conditions.

Moreover, in a workflow modeling the information should be associated with each activity. The information includes that who has control over the activity through the assignment of activities to qualified users or application function, what other activities are required to complete the activity, the information is input or output of the activity, and the data and control information required for task accomplishment. The necessary information for workflow modeling is presented in table 1 according to workflow characteristics.

Table 1. Workflow elements and definition

elements	Definition
data	represents the data that will be used or produced in an action
conditions	represents the conditions of an action
resources	represents the resources which will be required to do an action
participants	represents employees who will do an action
action	represents a piece of work. It can be a manual action or automatic action
timing	represents time elements which are time or deadline to accomplish the workflow action

3.2 The Expression of the ASM Semantics

By adapting the idea from workflow elements and definition, we try to express the extended ASM semantics to extend the existing UML activity diagram.

The existing action node lacks representation about the state of the action node and does not contain the timing element. Therefore the timing element and the state element of the action node are added, and the action node is extended.

The expression of ASM semantics for extended action node is presented in table 2.

Table 2. The expression of ASM semantics for extended action node

```
if currTarget is node(in,in_data,A,out,out_data,
                      isDynamic,dynArgs,dynMult,action_type,time)
then if ¬isDynamic then A
     elseif [dynArgs] = dynMult then
          for all L_i ∈ dynArgs
               A(L_i)
     elseif time ≥ deadline
          then active := exit()
     elseif action_type is manual then
          active := participaint
          active :=→ out
```

The action nodes are of form node(in, in_data, A, out, out_data, isDynamic, dynArgs, dynMult, action_type, time) where the parameter in and out denotes the incoming and outcoming arc, A an atomic action, isDynamic if A may be executed dynMult times in parallel, each time with an argument from a set dynArgs of sequences of objects where the parameter out_data is created after action node A is accomplished.

The parameter action_type can be a manual action or an automatic action, if action_type is manual, it involves the participant, and the parameter time is exited if time is over the deadline.

The condition element is added and then is extended using ASM semantics. The pre-conditions need to be satisfied so that an action can be enacted. The expression of ASM semantics for extended pre-conditions node is presented in table 3.

Table 3. The expression of ASM semantics for extended pre-condition node

```
if currTarget is node(PRE, out, isExistent)
then if isExistent then PRE
     elseif PRE = true then
          active :=→ out
```

If the pre-conditions are existent, pre-conditions are accomplished. If the pre-conditions are true, out is active. The post-conditions are conditions that need to be satisfied to tell if an action has been finished.

The expression of ASM semantics for extended post-conditions node is presented in table 4.

Table 4. The expression of semantics for extended post-condition node

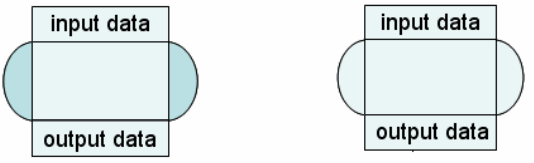
```
if currTarget is node(in, POST, isExistent)
then if isExistent then POST
     elseif POST = true then
          active := nextnode(in, in _data, A, out, out _data,
                             isDynamic, dynArgs, dynMult, action _type, time)
```

After the action is accomplished, if post-conditions are existent, the post-conditions are accomplished. If post-conditions are true, the next action can be accomplished.

Finally, the resources node is required to do an action that describes the required machine, tool and device to accomplish the work. Resources node is considered as a special action node without parameter such as initial node and final node.

4 The Extension of Activity Diagram in Workflow Modeling

Activity diagram is extended applying the concepts of extended ASM semantics. The activity diagram notation of extended action node is presented in fig. 1.



(a) manual action node (b) automatic action node

Fig. 1. Extended action nodes

In the fig.1, To distinguish the manual action type from the automatic action type, if the action type is manual, the action node is filled by color such as (a), and otherwise it is not filled such as (b).

Up and down the action, the rectangle parts present the input data, which is needed by the action node, and output data, which is the outcome from the action node. Also the item of the timing element is assigned in the action attributes, and is extended in the visual notation. The pre-condition and post-condition notation nodes are presented in fig. 2.

(a) pre-condition node (b) post-condition node

Fig. 2. Extended condition node

Also, resources node is a special node, the resources node is presented in fig. 3.

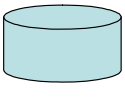


Fig. 3. Extended resources node

5 Case Study and Evaluation

5.1 Case Study

In this paper, we define that the workflow is the part of the retail trade, which is a sort of the process of ordering and purchasing the items. The business process that combine the automatic system with human system is modeled using activity diagram.

In the case study, we apply the workflow to the example of a shipment company that returns an item ordered by post. The workflow modeling for treating the return items is presented in fig. 4.

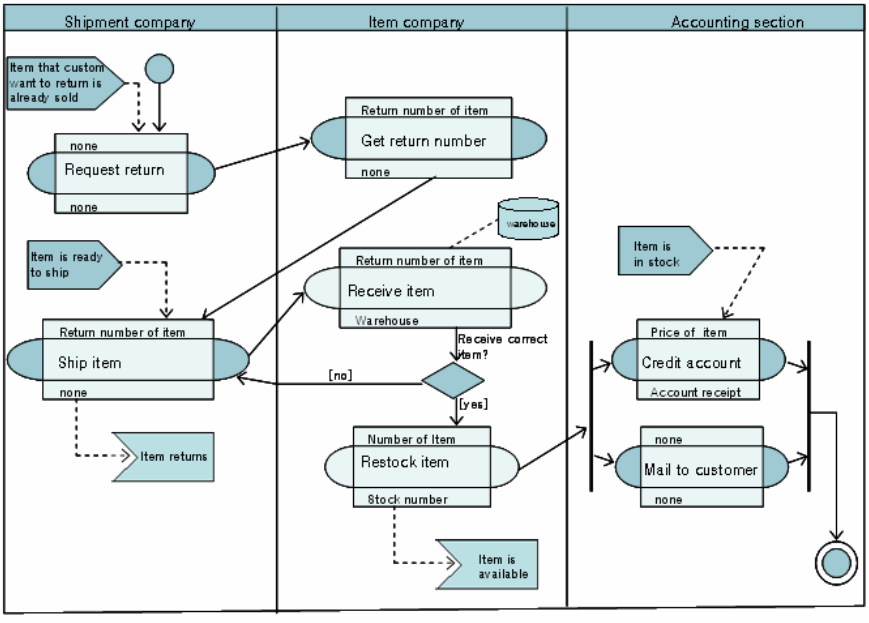


Fig. 4. The workflow modeling for treating return items

5.2 The Expression of ASM Semantics About the Workflow Modeling

In the case study, activity diagram nodes are expressed using the semantics in section 4. Not only these nodes can be expressed, but the activity diagram can also be expressed using the ASM semantics.

Fig. 4 shows the steps presenting the workflow modeling for treating return items. The steps can be expressed the ASM semantics in table 5.

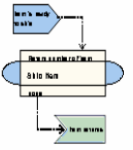
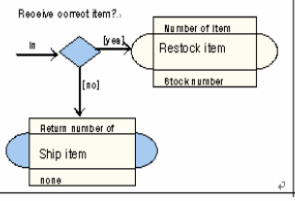
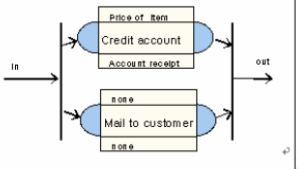
5.3 Evaluation

This paper presents the workflow modeling technology using activity diagram applying ASM semantics

The existing methods of applying Pi-calculus and FSP semantics express only the semantics of UML activity diagram itself and can not be expressed corresponding to the workflow characteristics.

Therefore, there is no study of the activity diagram extending and customizing based on the semantics, which has expressed the semantics of the UML activity diagram itself. It has been assumed that only the part of the UML activity diagram is expressed using the semantics.

Table 5. The steps expressing the ASM semantics

step	semantics
	$makeGraph(PREPPOST, in, out)$ $makeGraph(PRE, out)$ $makeGraph(P, in, out)$ $makeGraph(POST, in)$
	$makeGraph(\text{IF } a \text{ THEN } P \text{ ELSE } Q, in, out)$ $let\beta = new(BRANCHNODE)$ $yes = makeLink(\beta, a)$ $no = makeLink(\neg\beta, a)$ $connect(in, \beta)$ $makeGraph(P, yes, out)$ $makeGraph(Q, yes, out)$
	$makeGraph(P_1, P_2, in, out)$ $let\ \alpha = new(FORKNODE)$ $\alpha_1 = makeLink(\alpha, transition)$ $\alpha_2 = makeLink(\alpha, transition)$ $\beta = new(JOINNODE)$ $\beta_1 = new(transition)$ $\beta_2 = new(transition)$ $connect(in, \alpha)$ $makeGraph(P_1, \alpha_1, \beta_1)$ $makeGraph(P_2, \alpha_2, \beta_2)$ $connect(\beta_1, \beta)$ $connect(\beta_2, \beta)$ $connect(\beta, out)$

Jablonski and Bussler discussed the main components of a workflow model. They discussed them from the perspective of the workflow model. There are five perspectives, which are important that every workflow model should adhere to[11].

The function perspective describes the functional units of a work process that will be executed. The operation perspective describes how a workflow operation is implemented. The behavior perspective describes the control flow of a workflow, and the information perspective describes the data flow of a workflow. The organization perspective describes who has to execute a workflow or a workflow application[11].

In this paper, four perspectives were evaluated, with the operation perspective as exception because the operation perspective describes how a workflow operation is implemented and put a focus on the implementation. The four used were function, behavior, information, and organization.

The evaluation is presented in table 6. We compare our method with Petri-net method proposed Eshuis's study[12] and activity diagram method applying Pi-calculus semantics proposed Yang[4].

Table 6. Comparison our method with the other methods

perspective \ methods		our method	Eshuis's study (Petri-Net)	Yang's study (Pi-calculus)
function	dynamic	support	support	support
	goal-driven	the exit of activity diagram is not goal approach	the insufficiency of the understanding goal achievement	the exit of activity diagram is not goal approach
behavior	conditions and constraints	express the pre-condition and post-condition	the lack of the expression	express the constraints using transition
	timing	define the semantics but node notation is not supported	support the perfect timing elements	the lack of the timing elements
information	data flow	data flow using input and output data	data flow is not supported	data flow is not supported
	event-driven	AD itself is event-driven	event-driven because event modeling	AD itself is event-driven
organization	participants	express by swimlane	participants is not supported	express by swimlane
	resources	use the resources node	resources is not supported	don't use the resources notation

6 Conclusion and Future Works

The UML activity diagram is well-known for describing the systems that have a dynamic behavior. So it is useful to model business process and workflow.

In this paper, the alternative approach of using Abstract State Machines(ASM) to formalize UML activity diagrams is presented. Therefore, activity models can have rich process semantics. It is also suggested that ASM semantics corresponding to the workflow system characteristics are extended.

Through the exact definition to formal semantics based on ASM, it is possible to effectively model the workflow.

If our method is used, it would be for its understandable and ease of use without professional view, we can effectively treat the exception and various variations during workflow processing.

Future works need us to study the implementation of the workflow modeling tool using the extended activity diagram notation based on ASM semantics proposed by us.

References

1. OMG, Object Management Group, Unified Modeling Language Specification 1.4, 2001.
2. KEJ, Ko Eun Jung, Lee Sang Young, Yoo Cheol Jung, and Chang Ok Bae, Expression of ASM Semantics for Workflow Modeling Based on Activity Diagram, Proceedings of The 30th KISS Fall Conference, pp.331-333, 2003.
3. Han Y., Sheth A., Blussler C., A Taxanomy of Adaptive Workflow Management, Proceedings of the CSCW-98 Workshop Towards Adaptive Workflow Systems, 1998.
4. Yang Dong and Zhang ShenSheg Using pi-calculus to Formalize UML Activity Diagram for Business Process Modeling, IEEE ECBS, 2003
5. Jan Hendrik Hausmann, Reiko Heckel, and Stefan Sauer, Toward Dynamic Meta Modeling of UML Extensions: An Extensible Semantics for UML Sequence Diagrams, IEEE 2001, pp. 80-87, 2001.
6. Roberto W. S. Rodrigues, formalising UML Activity Diagrams using Finite State Processes, UML2000 Workshop, 2000.
7. Martin Gogolla and Mark Richters, Expressing UM Class Diagrams Properties with OCL, LNCS 2263, pp. 85-96, 2002.
8. Borger E., Cavarra A., and Riccobene E., An ASM Semantics for UML Activity Diagram, In: T. Rust(Ed.), Proc. AMAST 2000, LNCS, Vol. 1816, pp. 292-308, 2000.
9. Wuwei Shen, Kevin Compton, and James K. Huggins, A Toolset for Supporting UML Static and Dynamic Model Checking, 26th International Computer Software and Applications Conference(COMPSAC 2002), IEEE Computer Society 2002, pp. 147-152, 2002.
10. Hammer D. K., Hanish A. A., R., and Dillon T. S., Modeling Behavior and Dependability of Object-oriented Real-time Systems, Journal of Computer Systems Science and Engineering , Vol. 13(3), pp. 139-150, 1998.
11. Jablonski S., Bussler C., Workflow Management-Modeling Concepts Architecture and Implementation, Int. Thomson Publishing, London, 1996.
12. Eshuis R., and Wieringa R., A Comparison of Petri-net and Activity Diagram Variants, In Proceedings 2nd International Colloquium on Petri-net Technologies for Modeling Communication Based Systems, Berlin, Germany, 2001.

Unification of XML DTD for XML Documents with Similar Structure

Chun-Sik Yoo, Seon-Mi Woo, and Yong-Sung Kim

Division of Electronics and Information Engineering, Chonbuk National University,
664-14 1ga Duckjin-Dong, Duckjin-Gu, Jeonju, Jeonbuk, 561-756, Republic of Korea
{csyoo, smwoo, yskim}@chonbuk.ac.kr

Abstract. There are many cases that XML documents have different DTDs in spite of having a similar structure and being logically the same kind of document. For this reason, a problem may occur in which these XML documents will have different database schema and are stored in different databases, and we have to access all database that is concerned to process the queries of users. Consequently, it decreases seriously the efficiency of retrieval. To solve this problem, we propose an algorithm that unifies DTDs of these XML documents using the finite automata and the tree structure. The finite automata are suitable for representing repetition operators and connectors of DTD, and are simple representation method for DTD. By using the finite automata, we are able to reduce the complexity of algorithm. And we apply a proposed algorithm to unify DTDs of science journals.

1 Introduction

XML documents declare DTD (Document Type Definition) to define the structure of document, and perform the strict document structure validation using this DTD. Even if documents belong to the same kind logically, DTDs are varied by document manufacturers, manufacture time, etc. If DTDs are different, the schemas of database storing XML documents are different, and it is required to perform the same query in many times on several databases. Thereby, the retrieval time and the complexity are increased, and the quality of retrieval result is decreased. For example, the papers of the scientific journals have very similar structure [8]. But, in transformation process into XML documents, these documents are regarded as the different kinds of documents and have own DTD between journals. Despite these papers are logically of the same kind, documents are stored in the different databases and users have to access all related databases to process one's query. As a result, the cost of database construction and the number of database access is increased, and also the efficiency of retrieval decreases fairly. Therefore, it is needed a research that unifies DTDs of XML documents belonging to conceptually same kind and having similar structures, and that constructs XML document database using unified DTD [2, 3, 4, 6, 7, 8].

In this paper, we propose an algorithm that represents DTDs using tree and finite automata and unifies these DTDs of documents having similar structure. Using

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 954 – 963, 2005.

proposed algorithm, we expect that the cost of XML document database construction and that of database retrieval will be reduced, and the efficiency of retrieval will be improved. Also, the users can get more effective management and administration environment for XML document database.

2 Related Works

[1], [5] are researches that model DTDs of SGML documents using finite automata and tree, and [2], [6] are doing for XML documents. [1] creates DTD for same document class. In this research, a method is described that constructs finite automata which input a document inserted tags manually, merges finite automata, produces the regular expression for that document class using merged finite automata, and create DTD using regular expression. [5] introduces the automata theory and the set theory to create a content model of SGML documents. [2] did modeling XML DTD using regular tree, and it proposed an XML query algebra based on tree automata model. [6] proposed a method to find mapping automatically between elements of information sources and those of mediated schema. This research constructs trees about XML DTD, and applies term matching technique using thesaurus.

But, in most cases, these researches integrate structural similarity using simple DTD modeling technique and other researches integrate DTD to use as ontology for query interface. And some researches are researches about the integration of tagged documents manually. Thereby, they have weaknesses to accomplish the purpose of true DTD integration. Therefore, a research that integrates DTDs of documents belonging to conceptually same kind is needed [8].

3 XML DTD and Finite Automata

The structure of XML document is defined by element type declaration, and element type declaration can include entity. Therefore, we focus on elements and entities in XML DTD unification.

To unify XML DTDs, DTDs are transformed into tree and finite automata. The transformation method of XML DTD into tree is trivial. That is, the hierarchical structures of parent elements and child elements are represented to tree in the order that they are defined in mixed contents and content models. Connectors are of two kinds: ',' and '|'. And repetition operators are of three kinds: '?', '', and '*'.

4 DTD Unification Algorithm

The simplest method to unify DTDs of XML documents having similar structure is a method that contents of elements are declared as "ANY". But, this method is not suitable because it takes no account of the structure of documents and, as a result, the loss of meaning will occur. To solve this problem, a method will be used to connect DTDs of each XML document by OR ('|'). This method keeps the structures of each

document. But it is not able to connect simply each DTD by OR if the number of documents is more than two or the structure of each document is complex. Also, it is difficult to determine elements that are connected, and the content model of element connected becomes very complicate and becomes very long, and ambiguity problem can occur in DTD parsing.

To solve these problems, DTD unification method is needed that takes into account of the structure of each XML document and minimizes the loss of meaning.

4.1 Unification Algorithm for Similar DTDs

DTD unification algorithm proposed in this paper does not consider attribute-list declaration and relation of XML DTD and XML document instance.

<Definition 1> defines DTD unification of XML documents having similar structure. Common identifiers are elements that exist in more than two similar DTDs together.

<Definition 1> "DTD Unification"
DTD Unification is a technique that creates single unified DTD by using common identifiers, tree, and finite automata and preserving text elements of each similar DTD and minimizing the loss of meaning.

By definition, a unified DTD plays a role of global conceptual schema for common subject domain represented by XML documents having similar DTD.

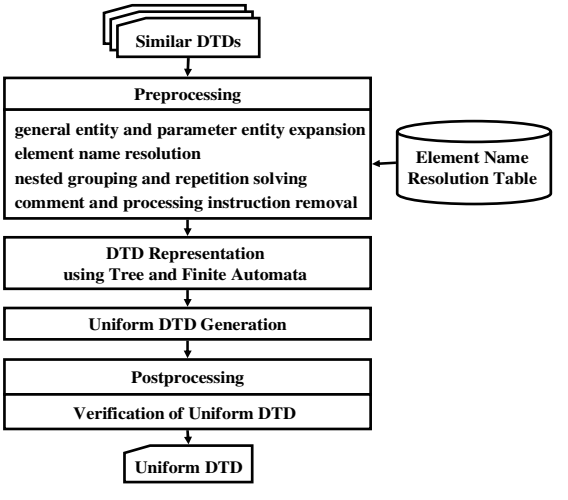


Fig. 1. Unification of similar DTDs

Figure 1 shows an outline of an algorithm that unifies the similar DTDs represented by tree and finite automata. In the first step of DTD unification, the different names of element having same meaning are resolved to the same name. And

then, DTDs are transformed into tree and finite automata, and tree and finite automata are merged to create a single unified DTD represented by tree and finite automata. Finally, a unified DTD is transformed to a DTD represented by text, and is verified by a DTD parser whether it is suitable to DTD grammar rules or not.

4.2 Preprocessing

Preprocessing performs general entity expansion, parameter entity expansion, element name resolution, nested grouping and repetition solving, and comment and processing instruction removal.

(1) General entity expansion and Parameter entity expansion
Because general entities are not usually used to represent structure of document but they are used for the abbreviation or for the external data utilization, they can be expanded into the original phrase in order to reduce the complexity of DTD unification.

Parameter entities are used in element's names, element's contents, attribute's names, and attribute's contents. Because they are used as the abbreviation for content model that is used often, they can be expanded into the original content model.

(2) Element name resolution
The names of element having the same or similar meaning in each DTD are unified using Element Name Resolution Table. Element Name Resolution Table processes synonyms of domain. It is constructed by a domain expert considering the hierarchical structure of element's name and the meaning of element's name. Table 1 shows a part of Element Name Resolution Table.

Table 1. Element Name Resolution Table

DTD-1	DTD-2	Element Name Resolved
Body	Content	Body
Front	InfoPaper	InfoPaper
Title	KorTitle	KorTitle

(3) Nested grouping and repetition solving
The problem of nested grouping and repetition is solved by following rules: <Rule 1> separates repetition operator from content model group using ',' connector. <Rule 1> intend to reduce the complexity of DTD transformation into tree and finite automata.

<Rule 1> ',' group's repetition operator problem solving rule: If a repetition operator is assigned to the element's content model grouped by ',', a temporary tag (“tmp”) is substituted for that group.

On the other hand, because '|' connector represent the element's existence independently of the element's order, a repetition operator is applied to each element composing group and that group is removed by <Rule 2>.

<Rule 2> '|' group's repetition operator problem solving rule: When a repetition operator is assigned to the element's content model grouped by '|' connector, the content model is expanded by the repetition operator assigned and the group is removed.

When any repetition operator is not assigned for the whole group using ',' connector, we can assume that the inner groups are used merely for readability or comprehensibility, and assume that the inner groups does not have a special meaning. <Rule 3> is a rule to remove the inner groups used in such content models.

<Rule 3> Nested grouping problem solving rule: When an element's content model is a nested group and a sequence of elements(',') is used and any repetition operator is not assigned for the whole group, the inner parentheses are removed.

4.3 Uniform DTD Generation

DTD unification step unifies the elements' logical structure by using tree and finite automata. [Algorithm 1] is a similar DTD unification algorithm that is consisted of a step merging trees of each similar DTD and creating uniform DTD (UDTD) tree, and a step merging finite automata (FA) of each similar DTD and creating UDTD FA through the UDTD tree traversal.

First step of [algorithm 1] is composed of four sub steps. In step 1-1 and step 1-2, the sibling nodes and the parent node for common identifiers (cid) being in several similar DTDs are determined, and UDTD tree is constructed. And, in step 1-3, nodes being terminal node and existing in only one DTD are merged with UDTD tree. In step 1-4, a UDTD tree created in step 1-3 is minimized, so the depth of UDTD tree is shortened. In second step of [algorithm 1], finite automata of cids' parent nodes are merged through BFS (Breadth-First Search) on the UDTD tree created in step 1.

[Algorithm 1] Unification of similar DTDs
Input : Tree, FA
Output : UDTD_Tree, UDTD_FA
begin
1. merges trees of similar DTDs.
  1-1. determines sibling nodes.
  1-2. determines parent node.
  1-3. merges all elements not being a cid with UDTD_Tree.
  1-4. minimizes UDTD_Tree
2. merges FAs of cid's parent node.
end.

[Algorithm 2] is an algorithm to decide sibling nodes of cids. The sibling nodes of cids are determined by union of each sibling node set being in each similar DTD.

```
[Algorithm 2] Sibiling nodes Determination
Input : Tree
Output : Sibs_cid(Sets of sibling nodes of each cid)
begin
 for( each cid ) {
   Sibs_cid[i] = Union of sibling nodes of cid in each similar DTD
 }   /* endfor */
 for( each Sibs_cid ) {
   if( Intersection of Sibs_cid[i] and Sibs_cid[j] is not empty ) {
       After comparison of size of two sets, larger set includes smaller set.
   }   /* endif */
 }   /* endfor */
end.
```

[Algorithm 3] is an algorithm to determine the parent nodes for each cids' sibling nodes set. First, if tmp node exists in parent nodes for those sibling nodes set, then tmp node is a parent node. And the parent node of tmp node is used in the following comparison to decide the parent node. If cids' parent node in each similar DTD are the same, then that node is the parent node for cids' sibling node set; in other cases, the parent node for cids' sibling nodes set is the node in one similar DTD that includes more cid's sibling nodes than those in other similar DTDs.

```
[Algorithm 3] Parent node Determination
Input : Tree, Sibs_cid
Output : Uncompacted_UDTD_Tree
begin
 for( each cid ) {
   if( the parent node of cid is tmp node ) {
       tmp node becomes parent node of cid and cid's sibling nodes.
   }   /* endif */
   if( the parent nodes of cid in each DTD tree are same ) {
      that node becomes parent node of cid and its siblings.
   }
   else {
       the parent node of cid and its sibilings is the parent node which includes more
       number of cid's sibiling nodes than any other nodes in DTD Trees.
   }   /* endif */
 }   /* endfor */
 for( each node that their parent node is not determined ) {
    copies a path from that node to root node into UDTD tree.
 }
end.
```

[Algorithm 4] is an algorithm to merge elements which are terminal node and are not cid with UDTD tree. For the elements that are terminal nodes and exist in only one similar DTD tree, a path from that node to root node is copied into UDTD tree.

```
[Algorithm 4] Merging all elements not being a cid with UDTD_Tree
Input : Tree, Uncompact_UDTD_Tree
Output : Uncompact_UDTD_Tree
begin
 for( each element that is not cid and is terminal node ) {
    copies the path from that node to root into UDTD_Tree.
 }   /* endfor */+
end.
```

[Algorithm 5] is an algorithm to compact UDTD tree. That is, tmp nodes and non-terminal nodes having only one child node are deleted from UDTD and their child nodes become child nodes of their parent node. By using [algorithm 5], the depth of UDTD tree and the complexity of UDTD FA merging are reduced.

```
[Algorithm 5] UDTD_Tree  Minimization
Input : Uncompact_UDTD_Tree
Output : UDTD_Tree
cur_NT :  non-terminal node which is being processed
parent_cur_NT :  parent node of cur_NT
children_cur_NT :  child nodes of cur_NT
begin
 for( each non-terminal node of UDTD_Tree ) {
   if( (cur_NT is tmp node) OR
       ((size of children_cur_NT equals 1) AND
         (children_cur_NT is non-terminal node)) ) {
      sets children_cur_NT to child node of parent_cur_NT.
      deletes cur_NT from UDTD_Tree.
   }   /* endif */
 }   /* endfor */
end.
```

[Algorithm 6] is an algorithm to merge finite automata for parent nodes of cid nodes. Through BFS traversal on UDTD tree created by [algorithm 5], finite automata for each non-terminal are merged. The method of merging finite automata is that if state transitions on two successive input symbols (element names) occurred in all similar DTDs, then those transitions are added to UDTD FA. In other cases, finite automata of similar DTDs are merged by the following three cases:

Table 2. FA merging Conditions

case 0	If the state transitions on two successive input symbols(element names) occurred in all FAs
case 1	① If there are finite automata that have input symbols belonging to other branches of UDTD tree and being cid, or ② If there are finite automata that have input symbols being non-terminal node and not being cid
case 2	If there are finite automata that have input symbols being terminal node and not being cid
case 3	In other cases

In case 3, input symbols are grouped by using OR connector('|') and '+' operator is assigned to that group when UDTD of text form is created.

```
[Algorithm 6]  Merging FA of cid's parent node
Input : FA, UDTD_Tree
Output : UDTD_Tree, UDTD_FA
begin
 while( there is no more node to merge ) {
  if( case 0 meets ) {
    merges immediate state for those input symbols to UDTD_FA.
  }
  else {
    if( case 1 meets ) {
        transites states for those input symbols to ignore those input symbols.
    }
    else if( case 2 meets ) {
       adds those input symbols and state transitions for those symbols to UDTD
       FA as options.
    }
    else {   /*  case 3 */
       adds all permutations of all input symbols to UDTD.
    }   /* endif */
  }   /* endif */
 }   /* endwhile */
end.
```

Figures 2 and 3, respectively, show UDTD Tree and UDTD FA of scientific journals.

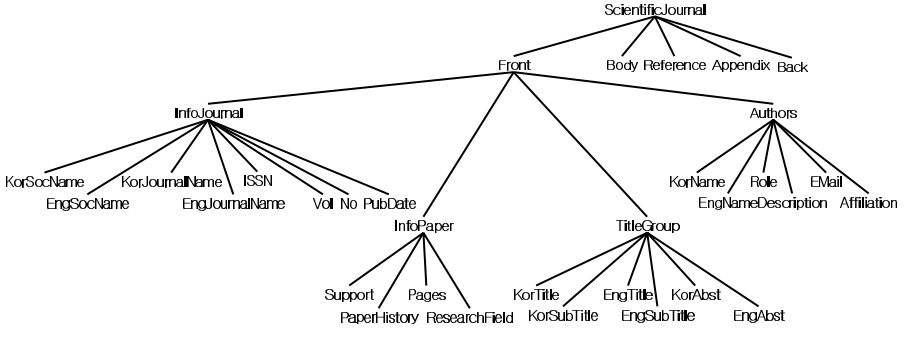


Fig. 2. Final UDTD tree created by [algorithm 5]

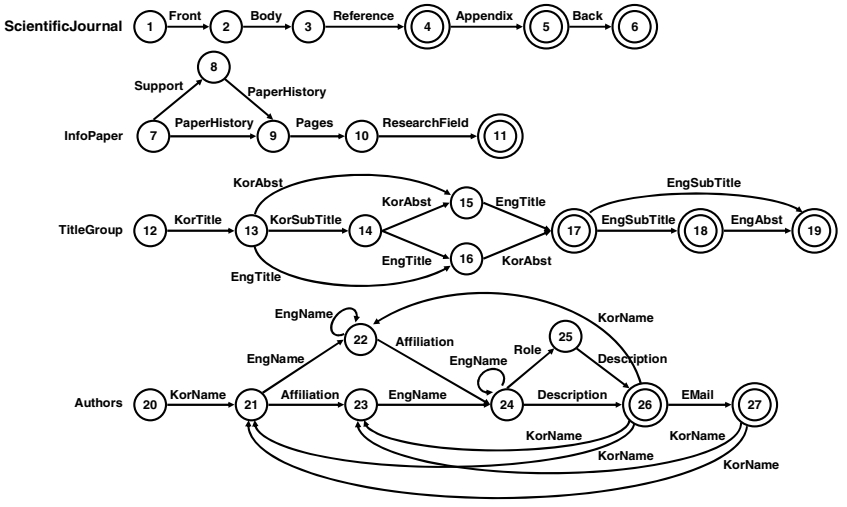


Fig. 3. UDTD FA created by [algorithm 6]

4.4 Postprocessing

Postprocessing performs the verification of UDTD by Microsoft's MSXML parser whether it is suitable to DTD grammar rules or not. Figure 4 shows a uniform DTD of text form after the verification. Looking into a uniform DTD as shown in Figure 4, we can confirm that it includes all contents of original XML documents, that is, all text elements of documents, and preserves each DTD's structure without much loss of meaning.

```
<!DOCTYPE  ScientificJournal [
<!ELEMENT  ScientificJournal (Front, Body, Reference, Appendix?, Back?)>
<!ELEMENT Front (InfoJournal, InfoPaper, TitleGroup, Authors)>
<!ELEMENT InfoJournal (KorScoName, EngSocName, KorJournalName, ISSN, Vol, No,
                       PubDate)>
<!ELEMENT  KorScoName (#PCDATA)>
...
<!ELEMENT  InfoPaper (Support?, PaperHistory, Pages, ResearchField)>
<!ELEMENT  PaperHistory (Submission, Acceptance)>
...
<!ELEMENT  TitleGroup (KorTitle, KorSubTitle?, (KorAbst|EngTitle)+,
                       (EngSubTitle?, EngAbst)?>
<!ELEMENT  KorTitle (#PCDATA)>
...
<!ELEMENT  Authors (KorName, (EngName|Affiliation)+, Role?, Description,
                    EMail?)+>
<!ELEMENT  KorName (#PCDATA)>
 ...  ] >
```

Fig. 4. UDTD transformed to text

5 Conclusion and Future Works

In this paper, we proposed and implemented an algorithm that unifies DTDs of documents having similar structure using tree and finite automata. Using the proposed algorithm, we can get the advantages that the cost of XML document database construction and the number of database access are reduced, and the efficiency of retrieval is improved.

In the future works, we will study methods for refinement the algorithm proposed to apply more various kinds of XML documents and methods for the transformation of the original XML document into document with uniform DTD. And we should research the retrieval technique which uses metadata to retrieve XML documents effectively and consistently. Also, it requires more researches about automatic schema generation and automatic construction of XML document database and metadata.

References

1. Helena Ahonen, Generating Grammars for Structured Documents Using Grammatical Inference Methods. University of Helsinki. Ph. D Thesis (1996)
2. Boris Chidlovskii, Using Regular Automata as XML schemas. 4'th IEEE Advances in Digital Libraries Conference(ADL 2000). Washington, USA (2000) 1-10
3. Euna Jeong, Chun-Nan Hsu, Veiw Inference for Heterogeneous XML Information Integration. Journal of Intelligent Information Systems. Vol. 20. No. 1. (2003) 81-99
4. Ronaldo dos Santos Mello, Silvana Castano, Carlos Alberto Heuser, A method for unification of XML schemata. Information and Software Technology. Vol. 44. No. 4. (2002) 241-249
5. OmniMark, OmniMark: Content Model Algebra. http://www.exoterica.com/white/cma/cma.htm
6. Chantal Reynaud, Jean-Pierre Sirot, Dan Vodislav, Semantic Integration of XML Heterogeneous Data Sources. International Database Engineering & Application Symposium(IDEAS 2001). Grenoble, France (2001) 199-208
7. Patricia Rodriguez-Gianolli, John Mylopoulos, A Semantic Approach to XML-based Data Integration. 20th International Conference on Conceptual Modeling(ER'2001). Yokohama, Japan (2001) 117-132
8. Chun-Sik Yoo, Seon-Mi Woo, Yong-Sung Kim, Automatic Generation Algorithm of Uniform DTD for Structured Documents. Proc. of IEEE Region 10 Conf. TENCON'99. Vol. II. (1999) 1095-1098
9. XML 1.0(Third Edition), W3C Recommendation. http://www.w3.org/TR/2004/REC-xml-20040204
   (2004)

Secure Payment Protocol for Healthcare Using USIM in Ubiquitous

Jang-Mi Baek and In-Sik Hong

Division of Information Technology Engineering Soonchunhyang Univ.,
646, Eupnae-ri, Shinchang-myun, Asan-si, Choongchungnam-do, 336-745, Korea
{bjm1453, ishong}@sch.ac.kr

Abstract. The ubiquitous environment aims to provide user-friendly service, which is the potential area that creates new value by grafting in diverse Internet business models. Healthcare service, which deals with health and the quality of life, is one area that actively features the ubiquitous business models. This paper introduces remote-controlled hospital treatment services and payment protocols for ubiquitous healthcare services. They make use of mobile terminals and USIM under different ubiquitous network environments. Through these new services, a more efficient treatment system will be studied and developed by connecting mobile terminals, USIM, and networks.

1 Introduction

Diverse services on the Internet, including cable networks, have surfaced for wireless networks. As cable and wireless networks have begun to merge, a new term called "ubiquitous environment" is coming to the fore. Ubiquitous environment makes network connection whenever and wherever possible. By connecting objects and people to all invisible networks, interactivity is made possible all the time. The ubiquitous environment aims to provide user-friendly service, which is the potential area that creates new value by grafting in diverse Internet business models. Healthcare service, which deals with health and the quality of life, is one area that actively features the ubiquitous business models. Several applications are being developed to improve healthcare services – from dealing with simple monitoring capabilities to utilizing diverse network with mobile terminals. This paper introduces remote-controlled hospital treatment services and payment protocols for ubiquitous healthcare services. They make use of mobile terminals and the Universal Subscriber Identity Module (USIM) under different ubiquitous network environments. Through these new services, a more efficient treatment system will be studied and developed by connecting mobile terminals, USIM, and networks. To create a hospital treatment system that is compatible with the general ubiquitous environment, technology to transfer and to check personal information frequently is needed. Since authentication is necessary for treatment and the payment system, we propose additional protocols for the safe transfer and authentication of data. In Chapter 2, the general notion of

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 964 – 972, 2005.

healthcare and the healthcare system in the ubiquitous environment will be analyzed from the background of healthcare history. In Chapter 3, the security issues needed in the healthcare system are discussed. In Chapter 4, hospital treatment systems are composed by mobile terminals, based on the analyzed security issues. A proposal of a suitable payment protocol will follow after this analysis. Chapter 5 will summarize the conclusions.

2 Ubiquitous Healthcare

Healthcare is the prevention, treatment, and management of illness and the preservation of mental and physical well-being through the services offered by the medical and allied health professions. Ubiquitous healthcare is the method of enhancing or maintaining health level by making many devices (services) available throughout the physical environment, but making them effectively invisible to the patients. Fig 1 is ubiquitous healthcare.

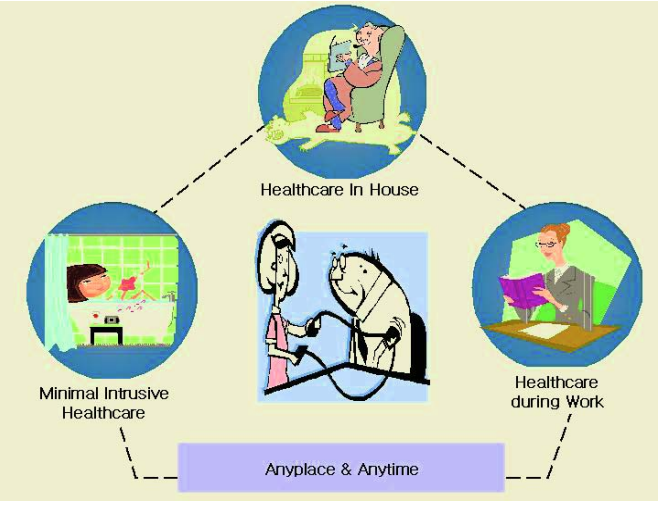


Fig. 1. Ubiquitous Healthcare

3 Security Requirements for Healthcare Service

Mobile healthcare services are value added services that help people cope with emergencies quickly. They facilitate continuous monitoring by providing patient health information to the proper recipient. Because this service involves personal information, this may risk abuse from malicious users without proper security. The following security issues are needed in mobile healthcare services.

3.1 Confidentiality and Authentication

As a basic item for security, confidentiality is needed in the transfer and handling of data in the mobile healthcare system. Only authenticated users should be able to access the data. Hospital treatment service requires patient data especially in processes like booking a hospital, receiving treatment, and prescribing. Since patient data involves personal information and information related to medical treatments, personal privacy should be handled carefully. The scenario suggested in this paper is a healthcare system involving mobile terminals and PDAs. This paper will present ways to authenticate and transfer data safely via SMS services through mobile terminals. To assure the confidentiality of the user's private information, privacy protection should be guaranteed. Additionally, this paper will propose ways not only of providing authentication and confidentiality of transferred data but also of providing authentication service for the users.

3.2 Integrity

Integrity means that data is not forged nor altered during transmission. Hospital treatment services must prescribe based on the patient's accurate information. If data are forged or altered, it will bring about fatal results to the patient. Therefore, the integrity of the hospital treatment system should be guaranteed.

3.3 Availability

All of the business models on ubiquitous environments suggest human centered scenarios. All healthcare scenarios should also be centered on people. It means that authorized users should be able to access data whenever and wherever possible. Therefore, the system should be constructed so that users cannot aware of such weaknesses in the service, even though troubles in the service system exist.

4 Remote Controlled Hospital Healthcare Scenarios

This chapter explains healthcare in remote controlled hospital systems among various business models in an ubiquitous environment. We will especially suggest a protocol for making safe payments after hospital treatment. Users or patients can pay through mobile terminals. We suggest a payment protocol that uses USIM or network money. To pay for treatment, we propose a security protocol of transferring data during the booking process, receiving treatment, and paying after treatment.

4.1 Overall Organization Chart

Under the ubiquitous environment, diverse healthcare models can be composed. This paper focuses on receiving healthcare services using mobile terminals. Assuming diverse business models are possible; this organization chart explains several possible scenarios. Following this is the discussion on the payment protocol used in the hospital paying system. The ubiquitous environment makes it possible to access diverse equipments. This system suggests environments for personal computers, mobile terminals, and sensor networks. That means that hospitals and drugstores will

be constructed under a computer-centered or terminal-centered environment, while users access wireless bio sensor networks and mobile terminals. Among these several diverse scenarios, we propose a protocol where users fall under a wireless network environment. Users or patients can access the hospital using wireless terminals such as PDAs or mobile phones. All terminals will be equipped with USIM for the confidentiality and integrity of personal information. Users can check the patient's status through mobile terminals and sensors. Users can receive healthcare services using USIM built into mobile terminals. In case of emergency, the service system will contact the hospital or doctors by checking the patient's status through sensing network equipment. Hospitals can also monitor the status of patients using diverse terminals. Several kinds of network equipments can facilitate booking services, treatment, and payment between hospitals and patients.

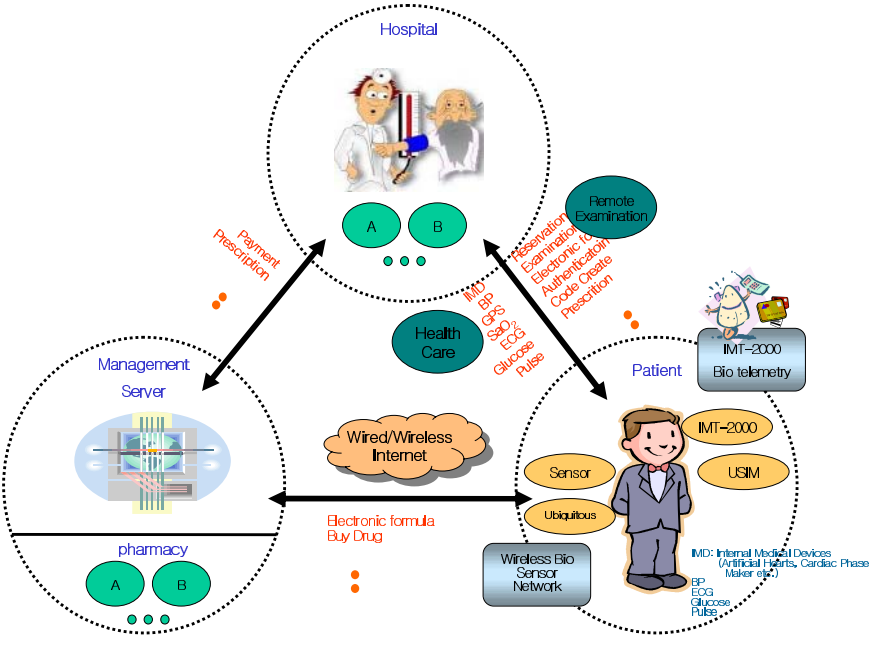


Fig. 2. Overall Organization Chart

As reliable organizations, hospitals can check after-payment, the authentication of treatment and prescription by connecting to medical care insurance management companies. In this paper, USIM connected to wireless terminals provide important roles. By using USIM inside the terminal, users can select hospitals, book them, inquire one's medical information, authenticate the patient in the process of treatment, receive medication prescriptions or shots, and pay hospital fees. To reserve a specific hospital, users must access the medical unification server through wireless Internet using a mobile communication terminal. The medical unification server authenticates access by checking identification numbers and authentication data stored in the USIM. Authenticated users will process booking procedures by selecting one of the

hospitals coded in the medical unification server. In treatment, patient medical history is loaded into each user's database using the authentication data stored in the USIM. Authorized users can confirm data by wireless Internet. After treatment, a user can get correct shots or prescription services by checking the data on a terminal that is connected to USIM. After treatment, if medicine is needed, prescription will be transferred to the appropriate drug store which the user appoints. By connecting to the drugstore's server computer wirelessly, the user transfers authentication information using USIM. The drugstore's server will download the prescription by connecting to the hospital server, using the user's ID through the network. Through unification management in medical issuance, the hospital and drugstore servers can be utilized for treatment and prescription with the patient's information in the USIM. Fig 2 is overall organization chart.

4.2 Elements of System

4.2.1 Hospital

Hospital must provide health services, process management support, activity and process coordination, identification, thorough consistent, integrated, adequate and reliable information dissemination.

4.2.2 Patient(User)

Patient can demand for health service is inherently individualized. Also, He demands that health service is inherently mobile (at work, at home, sports, leisure), mobility, disease information.

4.2.3 Mobile Computing(PDAs, Phones)

It provide potentially high to very high numbers of users. Also, It provide on all levels of public, and private health systems. It provide key factor for customer contact / patient relationship management.

4.3 Payment Protocol for Healthcare

This paper introduces remote-controlled hospital treatment services and payment protocols for ubiquitous healthcare services. They make use of mobile terminals and the USIM under different ubiquitous network environments. Through these new services, a more efficient treatment system will be studied and developed by connecting mobile terminals, USIM, and networks.

4.3.1 Healthcare Payment Protocol Through USIM/Network Type Payment

Fig 3 shows core components for healthcare payment through mobile terminals extracted from the organization chart. We assume that diverse network equipments have been established at the hospitals. These services are provided to doctors and nurses through different kinds of terminals. Payment gateway is a reliable organization and also works as a medical issuance management organization. Banks are virtual organizations that charge money to USIM or pay network-type money. This figure refers to the protocol in selecting hospitals, booking, receiving treatment, and paying fees after treatment. We exclude all other factors regarding drugstores or

other prescriptions. Users will make reservations in hospitals through mobile terminals. To reserve, the user will send user ID, authentication code, and his/ her signature. After checking the received data, the hospital will send back the data including the reservation time, the ID of the doctor-in-charge and the authorized signature from the hospital. The user performs payment through payment gateway by initiating the paying process with payment gateway. Then, the authentication for payment initiation will occur between hospitals and the payment gateway. The hospital will confirm reservation to the user, and then the treatment procedure will be processed when the user visits the hospital or when the doctor at the hospital locates the user. After treatment, the doctor will require the request list on the treatment and shots. The data will be transferred together with the doctor's signature. After checking the request list, the user will pay the fee. The data might be transferred through the hospital's signature. However, the doctor's signature is used, so that doctors cannot deny treatment in case treatment turns out to be wrong. After the payment is completed, the hospital will transfer a receipt to the user.

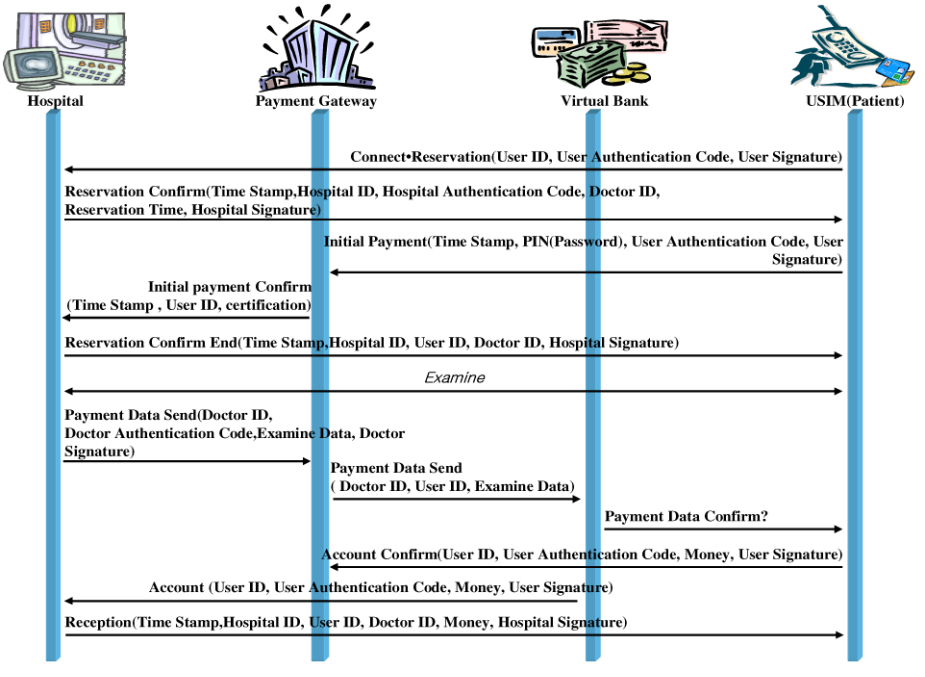


Fig. 3. USIM/Network Type Payment

4.3.2 Healthcare Payment Protocol Through USIM

Fig 4 shows the protocol on payments using USIM. Payment protocol using USIM does not include virtual banks. Because the money in USIM is assumed to be charged through the bank while data has been stored safely through the USIM coding process,

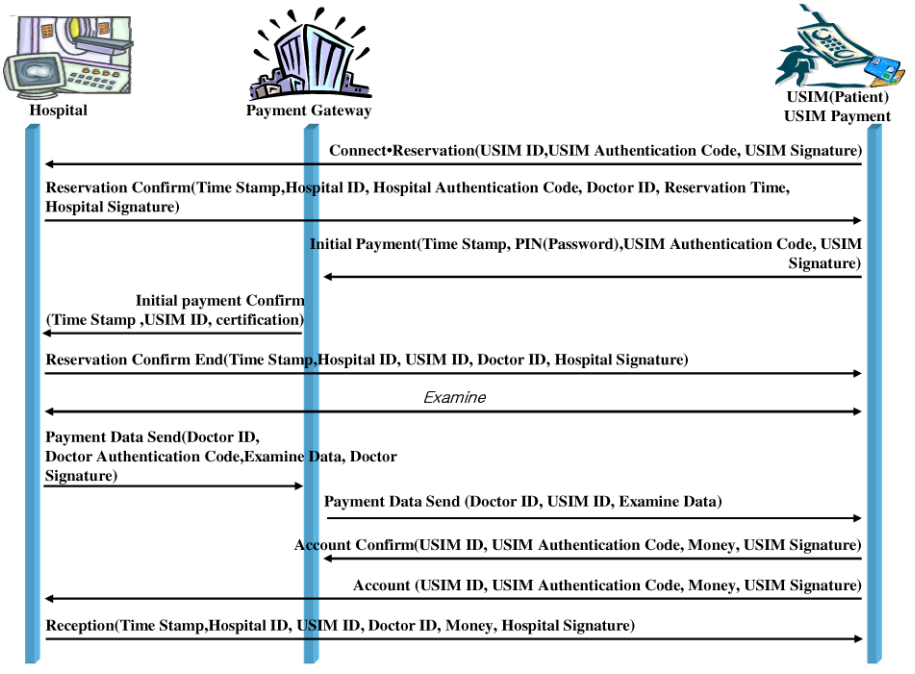


Fig. 4. USIM Payment

banks are not indicated. Like the protocol shown in fig 3, users will process the reservation and payment for the hospital using information in the USIM.

4.3.3 Healthcare Payment Protocol for Network Type Payment

Even though the data built in USIM are used to transfer and authenticate all of the data, the payment data in USIM will not be used in the final process for payment. Instead, the bank's network type electronic money will be used. This case is shown in Fig 5. All of the processes as explained above are almost the same. Because the payment data in the USIM will not be used in payment procedures, the user should register in and pay through the virtual bank, which requires a different procedure from that described above.

5 Conclusion

This paper studied a suitable remote controlled treatment system that will prove appropriate for ubiquitous environments, using USIM in wireless terminals (i.e. mobile terminals or PDAs). Healthcare service in ubiquitous environments provides a broad range of diverse models. This study focuses on payment after healthcare services are completed at the hospital. Before explaining protocols, we analyzed the

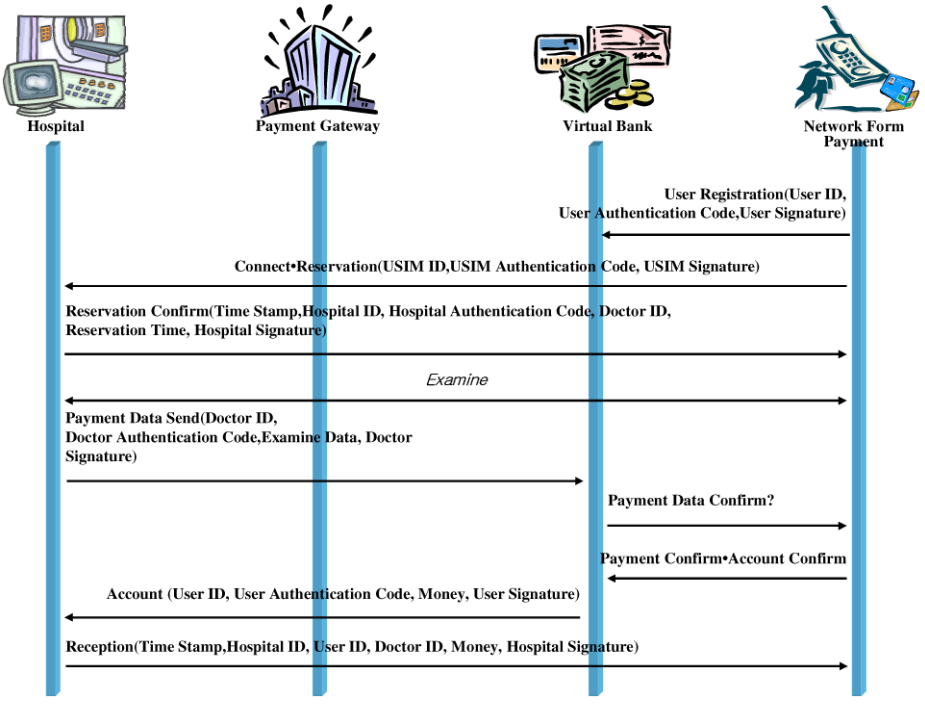


Fig. 5. Network Type Payment

security issues and suggested a general organization chart which can be composed in diverse ways. We focused on payment protocols by selecting payment related factors. Under the basic security provided by USIM, we suggested a system that allows users to receive healthcare information and make payments safely. Because the ubiquitous environment is still in its initial stage, various kinds of research should be accompanied to construct complete systems. While there are several kinds of scenarios suitable for ubiquitous environments, no cases have been actually realized. Thus, this paper can be used as a reference in creating application services for ubiquitous environments. Since the method suggested in this paper will impact not only healthcare but also the fields of economics, society, and culture, the study should be continued to suggest more business models, scenarios, and to create systems.

References

[1] Maritin Reichenbach, Hervert Damker, Hannes Federrath and Kai Rannenberg., "Individual Management of Personal Reachability in Mobile Communication", Information Security in Reserch and Business; 13th international conference on Information Security, Chapman&Hall, May 1997.p. 164-174.

[2] Toshitada NAGUMO, "Innovative Business Models in the Era of Ubiquitous Networks", NRI Papers NO.49, June 1, 2002.

[3] Hideaki NAKAMOTO, Nakoto SHIROTA, "System Integration Technology in the Ubiquitous Network Era", NRI Papers NO.64, May 1, 2003.
[4] Anderson, N., Janson, P., Waidner, M., "The State of the art in Electronic Payment System", Advances in Computers, 53, 2000.
[5] Salla Kalaja, "Security in Mobile Health Care Work", Tik-110.501 Seminar on Network Security, 2000.
[6] Dieter Gollmann, "Computer Security", John Wiley&Sons, 1999.
[7] Dimitri Konstantas, Val Jones, "MobiHealth-innovative 2.5/3G mobile services and application for healthcare", NEC Research, 2000.
[8] Ramon Marti, Jaime Delgado, "Security in a Wireless Mobile Health Care System", NEC Research, 2000.
[9] Hendry, M., "Smart Card Security and Application", Artech House, Boston/London, 1997.
[10] Zhiqun Chen, "Java Card Technology for Smart Cards", Addison-Wesley, 2000.
[11] MobiHealth(http://www.mobihealth.org)
[12] Healthmate(http://www.Healthmate-Project.org)
[13] Gemplus(http://www.gemplus.com)
[14] Java Card(http://www.java.sun.com/products/javacard)
[15] Mobile Java(http://www.mobilejava.co.kr)

Verification of UML-Based Security Policy Model

Sachoun Park and Gihwon Kwon

Department of Computer Science, Kyonggi University,
San 94-6, Yiui-Dong, Youngtong-Gu, Suwon-Si, Kyonggi-Do, Korea
{sachem, khkwon}@kyonggi.ac.kr

Abstract. Since the security policy model plays an important role in any secure information system, its specification has been studied extensively. In particular, UML-based specification has widely used because of its visual characteristics. Although visual specifications are good to write, they are difficult to verify whether some desired properties are hold in a given specification. This paper shows our techniques to verify UML-based specification with a running example.

Keywords: Security policy model, Role-based access control, Visual specification, Formal verification.

1 Introduction

Today, in accordance with the bigger complexity of the system, the concerns of the security increase day by day. One of the main concerns of the high security information system is the security evaluation to define the security function. Security evaluation of information security system is broadly used with respect to Common Criteria (CC) as ISO standards (ISO/IEC 15408:1999). As CC presents common requirements about the security function and assurance means of the system, it makes enable mutual comparison among the evaluation results of the security system that performs independently. CC Evaluation Assurance Levels (EALs) have seven levels, from EAL1 to EAL7. The higher EAL levels (5, 6, and 7) are sometimes referred to as the high-assurance levels. These levels require some applications of formal methods to demonstrate that the appropriate level of assurance has been met.

One of important things to get the high level assurance required to the confidence system is Security Policy Model (SPM). In CC, evaluation of the secure system, so called Target of Evaluation (TOE), is concerned primarily with ensuring that a defined TOE Security Policy (TSP) is enforced over the TOE resources. The TSP defines the rules by which the TOE governs access to its resources, and thus all information and services controlled by the TOE. The TSP is, in turn, made up of multiple Security Function Policies (SFPs). And then the SFP is implemented by a Security Function (SF),whose mechanisms enforce the policy and provide necessary capabilities. Therefore, the system developers also develop a security policy model that is based on a subset of the policies of the TSP, and establish a correspondence between the functional specification, the security policy model, and these policies of the TSP[1].

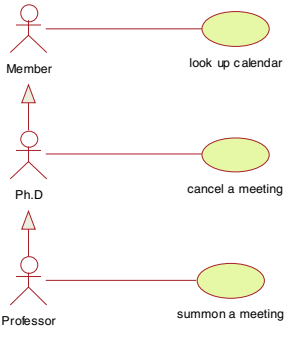
O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 973 – 982, 2005.

As SPM is basically the criteria of security functions when the information security system is implemented, to guarantee the confidence degree of SPM does an important role to the confidence degree of the information security system. Besides High assurance level over EAL5 requires the formal SPM. In this paper, we propose method and process for specifying and verifying the semi-formal and formal SPM of given TOE. Example TOE is the scheduling system and has access control policy that is presented by Role Based Access Control (RBAC). Section 2 introduces RBAC policy and shows SPM specification process from the TOE policy to its formal model. The method for verifying semi-formal which is constructed by UML and formal SPM model by Alloy [2] developed by MIT is described in section 3. Finally, the section 4 gives conclusions and future works.

2 Specification of SPM

2.1 TOE Description

TOE is a scheduling system which supports confirming meeting in a research group where professors, Ph.D. students and other members frequently meet together. Each user (subject), who takes part in the system, is enforced by role-based access control policy. Role-based access control (RBAC) is an alternative to traditional discretionary(DAC) and mandatory access control (MAC) policies that is interest in g in comercial applications. Under RBAC, users are granted membership into roles in system, where a role can be defined as a set of actions and responsibilities. Because of allowing the role hierarchies, some users can have several roles. Permissions can be represented by sets of relations between roles and system resources (objects), and the user and the role with assigned by session are activated. Thus RBAC consists of the user-role assignment, the role-permission assignment, and the permission-object assignment[3]. Figure 1 shows the use case diagram of scheduleing system, where actors in the use case are corresponding to roles in the organization.



In the laboratory room, there are professor, Ph.D. student and members doing research. They use a scheduling system for team meeting.

1. Professor books a meeting in the calendar.
2. Professor requests the meeting to members.
3. Requested members check the date in the calendar.
4. All Members agree to the request.
   4a. There exists a member opposing the meeting.

Fig. 1. Use case diagram

Therefore roles in this organization can be defined by *Role* = {*Professor*, *Ph.D*, *Member*}. Also, through operations presented in sequence diagrams, we can get

corresponding permissions. Figure 2 gives two scenarios, cancel a meeting and summon ameeting. Because assets to be protected are meeting schedules in this system, permissions is built up by whether a method is allowed to an actor. Accordingly, we define the*Permission* as {*show*, *invite*, *agree*, *disagree*, *cancelMeeting*, *addMeeting*}, where *show* is a permission to read the meeting schedules, *invite* is to create a meeting, agree isto consent to the reserved meeting, *disagree* is to oppose join, *cancelMeeting* is tocancel a meeting, and *addMeeting* is a permission to confirm the meeting. Figure 3presents class diagram of example system.

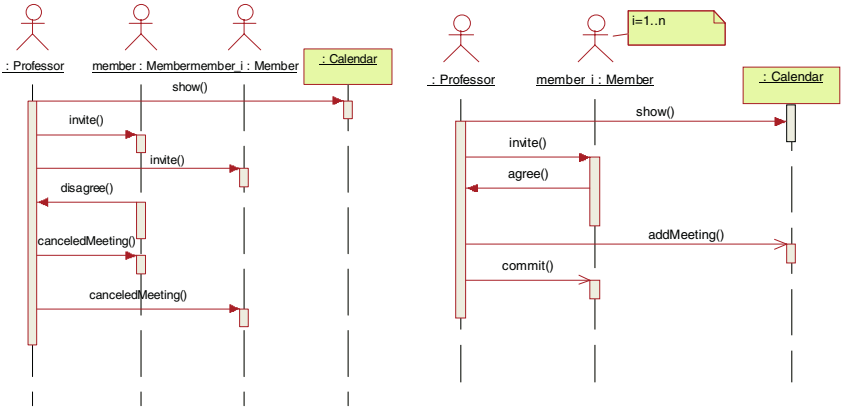


Fig. 2. Sequence diagrams for Cancel a meeting and Summon a meeting

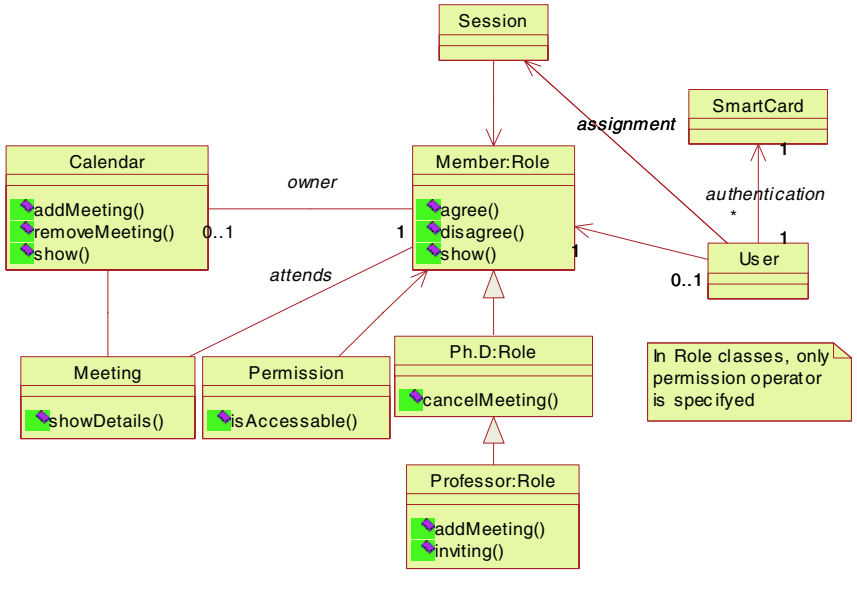


Fig. 3. Class diagram of scheduling system

2.2 Process for Specifying Formal SPM

In the previous section, we describe the TOE and its access control policy by mapping elements of UML diagrams to elements of the RBAC policy. In this section, we will explain a process to drive from system security policies to semi-formal and formal SPM models. Followed policies are only related to access control in the TOE:

P1. All activities accessing to scheduling system are controlled by the system.
P2. RBAC security policy is used.
P3. A *Member* has permissions such as *show*, *agree*, *disagree.*
P4. *Ph.D* inherits all *Member*'s functions, and posses the *cancelMeeting* permission.
P5. Permissions of the *Ph.D*'s and additionally *inviting*, *addMeeting* assign to a *Professor.*
P6. Only one user assigns to the *Professor.*
P7. Each user has a role.

From above security policies, we are able to write an informal SPM model with respect to the concept of RBAC. Security policy model of RBAC has three components, (i) type information of system elements that are enforced by security policy, (ii)policyrules that generate the states of the system accepted by the policy, (iii) constraints thatshall not be contained in any system states (negative constrains) and must beexpli-citly constructed as parts of system states (positive constrains).

(i) Type information
T1. A user is assigned to a role.
T2. A role has another role hierarchically.
T3. A user can establish a session.
T4. A role is assigned to a session.

(ii) Policy rules
R1. A user can be assigned to a role by an administrator.
R2. A user can be revoked from a role by an administrator.
R3. A user can be assigned to a session.
R4. A user is removed from a session.
R5. A user, establishing a session, can activate a subset of authorized role.
R6. A user, using a session, can deactivate the role in the session.
R7. A *Member* has permissions such as *show, agree, disagree.*
R8. *Ph.D* inherits all *Member*'s functions, and posses the *cancelMeeting* permission.
R9. Permissions of the *Ph.D*'s and additionally *inviting*, *addMeeting* assign to a *Professor.*

(iii) Constraints
C1.Separation of duty (sod) – negative constraint : If for relation between two roles those roles cannot simultaneously be assigned to the same user, to prevent the user from being a member of two roles, those roles are established by the sod-constraint.
C2.Cardinalities – positive constraint : This kind of constraint restricts the numbe rof user-role assignment. In this scheduling system, only one user can be assigned to the role *Professor.*

C3. Another additional specific constraint is that any user can be assigned to one role in this system.

In the above description, type information T1, T2, T3, and T4 present basic RBAC type information and represent P2 security policy. Policy rules R1, R2, R3, R4, R5, and R6 also refine the P2 policy. Security polices P3, P4, and P5 are corresponding to R7, R8, and R9, respectively. C1 and C2 constraints represent P2. Constraints C2 and C3 reflect security polices P6 and P7. Thus all TSP is represented on informal SPM model because P1 presents organizational security policy.

Now we are about to drive the informal SPM to semi-formal SPM by means of UM L object diagram. Firstly, type information about system entities is modeled by association relation between objects to be labeled by a type. Each type means smart card, user, role, permission, and session. Basically user access to the system with smart card,but in this paper we do not deal with this authentication and identification policy. The loop of the object typed Role presents role hierarchy. Authenticated user can be assigned to a role and access to system through a session. Figure 4 includes type information from T1 to T4.

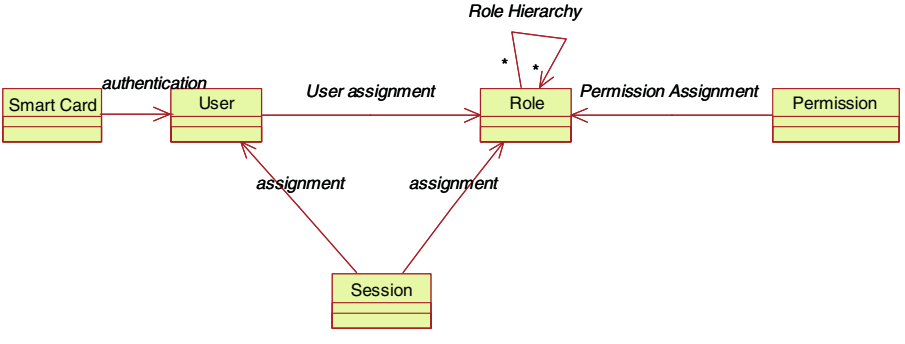


Fig. 4. UML model of Type information

Secondly, figure 5 gives model of policy rules and the number in front of each object diagram models 1) creating a new user, 2) deleting a user and the assigned session of the user, 3) establishing a session to user (R3), 4) destroying session (R4), 5) assigning a role to a user (R1), 6) destroying the role (R2), 7) activating the role authorized (R6), and 8) deactivating the role (R7). Therefore all policy rules are represented in UML model where the star (*) means role hierarchy.

Finally, constraints also model by means of UML model and OCL. Left side in figure 6 models the sod-constraint (C1), and in the right side, upper constraint represents that only one user can be assigned to the *Professor* role (C2). Lower constraint is the constraint C3. We use the OCL language to specify these constraints more precisely.

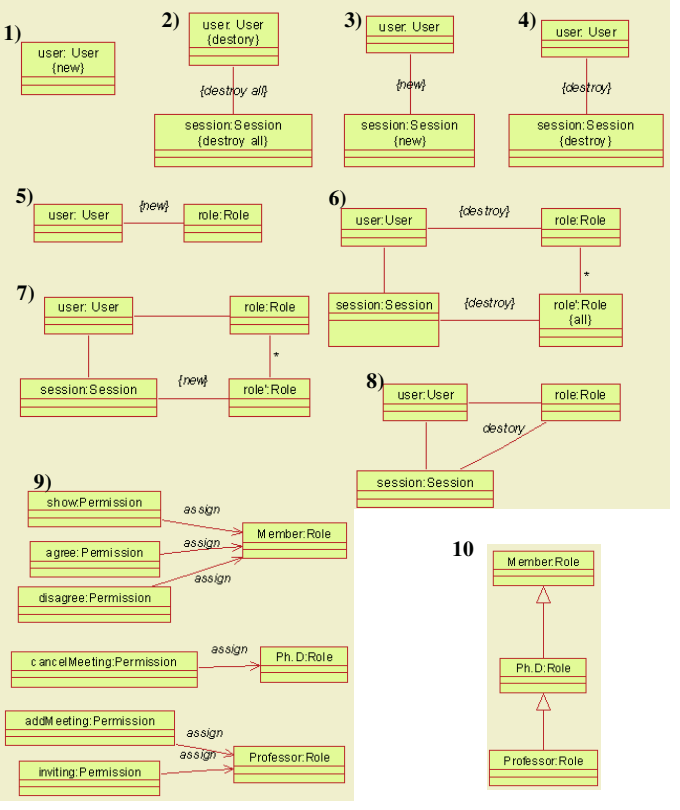


Fig. 5. UML model about policy rules

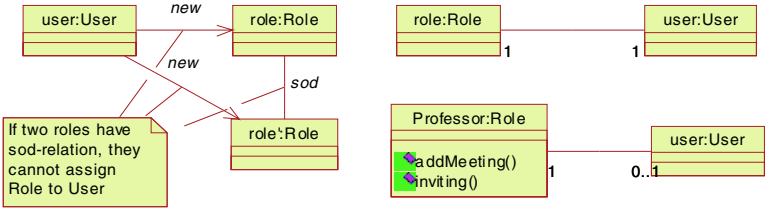


Fig. 6. Constraints presented by UML

Table 1. Constraints presented by OCL

constraint	OCL specification
(C1)	`context User inv self.is_in->forAll( r1,r2 | r1.sod->excludes(r2))`
(C2)	`context Professor inv OneAssignmentProfessor: user -> size < 2`
(C3)	`context User inv OneAssignmentUserToRole: role1 -> size < 2`

So far we explained the process to construct an informal and semi-formal model from given policies. SPM formal model will be constructed by based on these models. To specify the formal model, we borrow a formal language of the Alloy which is a popular lightweight formal method.

```
module schedule
sig User { is_in : Role }                                          // T1
sig Session { has: User, belongs : User, is_activated:Role}        // T2, T3
sig Role { superrole, subrole, sod : Role  }                       // T4
fun NewUser(user, User':User) { User' = User + user } // creating a uer.
fun RemoveUser(user, User':User, Session' : Session) {
  Session' = Session - user.~has
  User' = User - user
}                    // deleting a user and the assigned session of the user
fun NewSession(user, User':User, session, Session':Session) {      // R3
  Session' = Session + session
  User' = session.belongs + user
}
fun RemoveSession(user:User, session, Session':Session) {          // R4
  Session' = user.~has - session
  Session' = Session - session
}
fun AddToRole( user:User, Role', role:Role ) { Role' = user.is_in + role }
                                                                   // R1
fun RemoveFromRole(user:User, Role', role:Role) {                  // R2
  Role' = user.is_in - role
  Role'=user.~has.is_activated-role.*subrole
}
fun ActivateRole(user:User,session:Session, Role':Role, role:Role,role2:Role) {
  role in user.is_in role2 in role.*subrole
  Role' = session.is_activated + role2                             // R5
  session.is_activated' = session.is_activated + role2
}
fun DeactivateRole(user:User, session:Session, Role',role, role2:Role) {
  Role' = session.is_activated – role
  Role' = session.is_activated - role2                             // R6
  session.is_activated' = session.is_activated - role2
}
fact { all u: User, r1: Role, r2: Role |
                     (r1+r2) in u.is_in => r1 !in r2.sod }         // C1

assert SeperationOfDuty                                            // C1 property
{ all r1,r2: Role, u:User |(r1 + r2) in u.is_in -> r1 != r2.sod }
assert OneAssignmentUserToRole                                     // C3 property
{ all r: Role | all u1, u2:User | u1.is_in = r => u2.is_in != r }
```

Fig. 7. Alloy model of scheduling system

3 Verification of SPM

In the previous section, we have modeled semi-formal SPM and formal SPM with three components, type information, policy rules, and constraints. To verify the SPM models, we use the USE[4] for analyzing a UML object diagram and Alloy[2] for the formal model where type information and policy rules are a model to be verified and properties is specified by constraints. Through our analysis we show the way that specified type information and policy rules are consistent with constraints, that is, inter-consistency of SPM model. Figure 8 presents USE model and a snap shot to be analyzed. Results of checking this model are consistent.

```
model schedule
-- type model about USER, Role,
Permission, Session
class User
end
class Role
end
class Permission
end
class Session
end
--
class show < Permission
end
class agree < Permission
end
class disagree < Permission
end
class cancelMeeting < Permission
end
class addMeeting < Permission
end
class inviting < Permission
end
--
class Member < Role
end
class PhD < Member
end
class Professor < PhD
end
class Jungrim < User
end
class Taehoon < User
end
class Sachem < PhD
end
class Gihyun < Professor
end
```

```
-- R1
association UserAssignment between
 User[*] role assign;
 Role[1..*] role new;
end

-- role hierarchy
association RoleHierachy between
 Role[*] role sub_role;
 Role[*] role super_role;
end

--permission assigned to role
association PermissionAssignment between
 Permission[*] role assignment;
 Role[*] role role1;
end

-- R3
association SUassignment between
 Session[*] role user_assignment;
 User[*] role user;
end

-- R5
association SRassignment between
 Session[*] role role_assignment;
 Role[*] role role1;
end

constraints
-- C3
  context User
    inv OneAssignmentUserToRole:
        new -> size < 2
-- C2
  context Professor
    inv OneAssignmentProfessor:
        assign -> size < 2
```

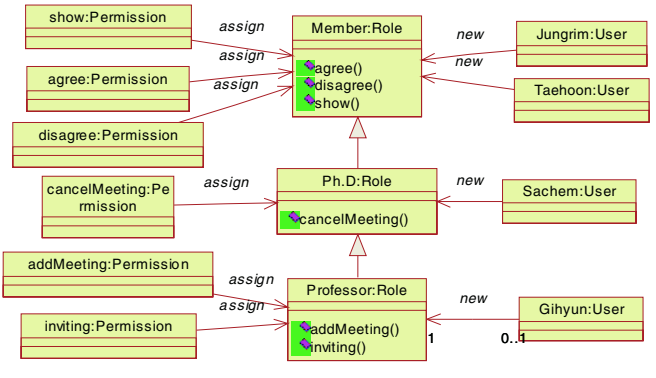


Fig. 8. USE model and one snap shop

If we add illegal user-role assignments like below insert-statements, then the USE analyzer show how the constraints doesn't satisfied in given snap shot. Actually, these user-role assignments violate C3 constraint and figure 9 shows the violation.

```
"!insert (Jungrim, Member) into UserAssignment
 !insert (Jungrim, PhD) into UserAssignment"
```

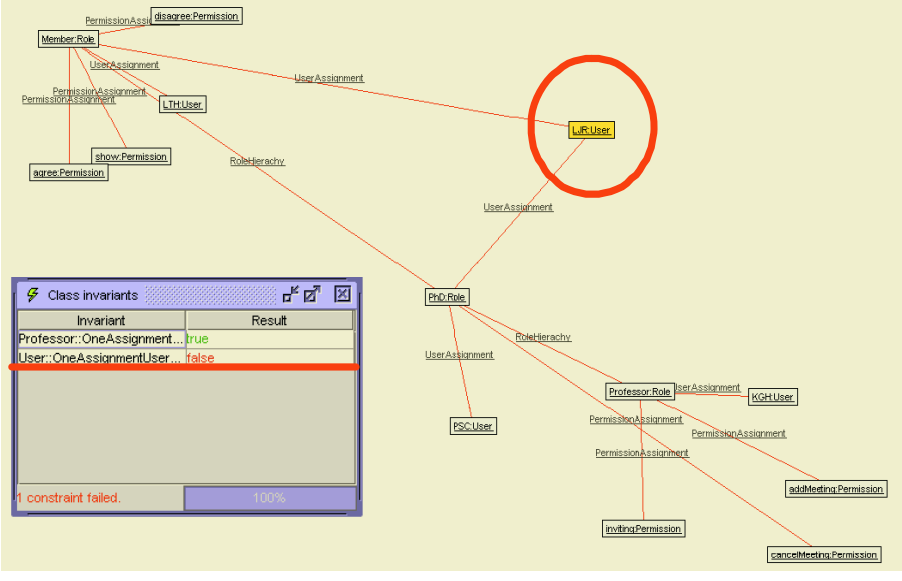


Fig. 9. Violation of C3 constraint

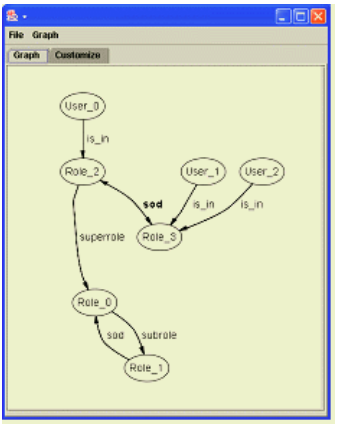


Fig. 10. Counterexample of violation

In the previous section, we intentionally induce the formal SPM model to violate C3 constraint. Of course, the omission of one statement "`fact {all r: Role | all u1, u2:User | u1.is_in = r => u2.is_in !in r}`" causes this violation and in this case Alloy produces one of counterexamples.

4 Conclusions

It is the objective of SPM to provide additional assurance that the security functions in the functional specification enforce the policies in the TSP. This is accomplished via the development of a SPM that is based on a subset of the policies of the TSP, and establishing a correspondence between the functional specification, the SPM, and these policies of the TSP. The formal SPM is also essential component for the high-assurance levels [1]. However, it is very difficult to construct a formal model from security policies of an information system. Recently, several studies about specifying and verifying a formal model for a security policy, particularly RBAC policy, are in progress [5,6,7,8]. But for developers, preparing the CC evaluation, how to specify and verify their SPM couldn't be guided well.

In this paper, we presented specifying the semi-formal and formal SPM of scheduling system and verifying the SPM model with graphical analysis tools, USE, Alloy. Models to be verified were driven from security policies of TOE, step by step. With these SPM models, we performed consistency check. We obtained two benefits through proceeding step by step, which are correct models and the consistency between policies and models. But establishing a correspondence between the functional specification and the SPM, to satisfy the full requirement of CC, is remained works.

References

1. Common Criteria for Information Technology Security Evaluation, Version 2.1, August 1999, CCIMB-99-031, http://csrc.nist.gov/cc/CC-v2.1.html
2. http://alloy.mit.edu/
3. D. F. Ferraiolo, D. R. Kuhn, R. Chandramouli, "Role-Based Access Control," Artech House, 2003.
4. http://www.db.informatik.uni-bremen.de/projects/USE/index.html
5. M. Koch, L. V. Mancini, and F. Parisi-Presicce, "A graph-based formalism for RBAC," ACM, Trans. Inf. Syst. Secur, 5(3), pp.332-365, 2002.
6. M. Koch and F. Parisi-Presicce, "Formal Access Control Analysis in the Software Development Process," in Proc. of Formal Methods in Security Engineering: From Specifications to Code (FMSE'03), 2003.
7. M. Koch and F. Parisi-Presicce, "Visual Specifications of Policies and Their Verification," FASE 2003, pp. 278-293, 2003.
8. J. Zao, H. Wee, J. Chu, D. Jackson, "RBAC Schema Verification Using Lightweight Formal Model and Constraint Analysis," SACMAT, 2003.

From a Small Formula to Cyberworlds

Alexei Sourin

School of Computer Engineering, Nanyang Technological University,
Nanyang Avenue, 639798, Singapore
assourin@ntu.edu.sg
http://www.ntu.edu.sg/home/assourin

Abstract. Cyberworlds created on the web allow for providing personal mentoring of the students with different cultural and educational backgrounds. Virtual Campus of Nanyang Technological University is designed to be such a cyberworld. This is a place for research and education, fun and immersion in campus life. Besides standard VRML, its hybrid function-based extension is used in the design of Virtual Campus. In place of thousands of polygons, small formulas are used for constituting complex geometric shapes and appearances. Collaborative Shape Modeling Laboratory, which is a part of the Virtual Campus, is based on this extension. It is developed to help students with their computer graphics assignments.

1 Online Experiences Versus Online Courses

Cyber-learning has already become an important and vital part of university education. Conventional ways of teaching, when only lectures and practical exercises in a class are used, no longer satisfy the growing demands and challenges of modern education. This becomes especially essential when large or distributed classes are being taught. The common way of electronic education nowadays assumes putting on the web course materials, which can be further fortified with elements of interaction such as discussion groups and quizzes. However, the temptation to provide more and more information electronically, thus far acceding the previously used pace of teaching, may often result in disorientation and exhaust of the students whose abilities and expectations from electronic education may significantly vary. Therefore, the ultimate goal of the electronic education should be to provide personal mentoring to each student through smart virtual instructors, media rich course content, and online experiences rather than just online course materials published on the web.

Cyber-learning is very essential for teaching computer graphics. Traditional project-based ways of teaching computer graphics may appear boring and not really motivating for the students, since many of them do not necessarily see their future professional career in designing and developing sophisticated graphics tools. Also, quite often the course-work offered to them assumes using laboratory-based equipment and software, which are not available in other labs and at home. This also contributes toward a negative attitude and strain due to the possible problems with accessing these facilities. Distant students located outside the institution offering the course, even in other countries, may have no way at all of working with other students on the projects and thus will feel left behind.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 983 – 992, 2005.

In this connection, using and creating cyberworlds appears to be an excellent educational tool for achieving personal mentoring as well as creating a feeling of immersion into real collaborative work when teaching large and distributed classes. Cyberworlds are information worlds built on the web [12]. These shared cyberspaces may be designed both with visual appearances and without it. Cyberworlds are successfully used in cyber-learning for science [17], children [7] and archeological education [8]. Creating cyberworlds as an educational tool for teaching computer science is described in [9].

In this paper, three projects contributing to developing cyberworlds with educational applications are described. These are Virtual Campus of Nanyang Technological University, Function-based Web Visualization, and Interactive Function-based Shape Modeling. The names of the project students can be found in the project webpages [19, 6, 10].The main idea of using these tools in teaching computer graphics is that the students are able to use simple mathematical formulas for defining complex geometric shapes, appearances and transformations. Using any Internet connected personal computer, they are able to connect to the cyberworld, which allows them to design sophisticated shapes as well as discuss and make available to other students the created models. Other ways of teaching including online lectures and interactive consultations with the robot and real lecturers are available to the students as well. This approach, in fact combining fun and education, allowed for creating more motivation and achieving better results when teaching computer graphics to some 450 students per term, who have very different educational and cultural backgrounds.

2 Cyber-Learning on the Virtual Campus of NTU

Nanyang in Chinese means “south seas”—a reference to the Southeast Asian region. Back in the 1940s and 50s, many Chinese from mainland China ventured south to seek their fortunes in new lands. Malaya—now Singapore and Malaysia—was then known as Nanyang to the Chinese. After World War II, a university was founded in Singapore that would provide tertiary, comprehensive education in Chinese. Expanded in 1953, it was known as Nan Tah in Chinese. The modern Nanyang Technological University [15] originated from Nan Tah. NTU occupies a large, beautiful campus with hilly terrain in Jurong, located in the western part of Singapore. Many of the campus buildings have sophisticated, futuristic architecture, some designed by the famous Japanese architect, Kenzo Tange [11].

Virtual Campus of NTU is a shared virtual world built with Virtual Reality Modeling Language and Blaxxun Contact communication platform [4]. It is a virtual model of the real campus of Nanyang Technological University. The whole Virtual Campus including VRML models of the land, buildings, interiors, avatars, and texture images is stored in only about 15 Mb of files and can be accessed from any Internet connected personal computer [19] (Figure 1).

Many visitors to the Virtual Campus are computer graphics students, who either play virtual “hide-and-seek” with their professor, or come to study concepts of virtual reality and shape modeling. There are also visitors from around the world meeting together on this hospitable land. In this cyberspace, visitors can turn themselves into virtually anything. Some choose to look like fancy-dressed people, some turn them

Fig. 1. A snapshot of the real and Virtual Campus

selves into sports cars, and some appear as sparkling clouds or fire-balls. Local students easily navigate the familiar 3D environment, go to their favorite places, or meet with friends in the hostel rooms. Foreign guests usually just wander around and chat, astonished by the size of what is probably the biggest shared cyberspace of this kind.

There are dusks and dawns in this cyberspace, which follow Singapore time, but the Virtual Campus never sleeps. Many bots (robots) populate it. These are avatars of students and professors who walk back and fourth between lecture theatres, libraries, and student hostels. There are also birds hovering in the sky and cars riding by the roads. The bots are programmed to behave realistically for the visitors. Some of these activities are stochastic, and some follow the real class time-tables.

Virtual Campus is not only for walking-through and seeing other avatars or bots. The visitors can talk to them. Blaxxun Contact provides the communication platform for it. It also allows for text-to-voice synthesis so that the visitors can hear their computer-simulated voices. These chats may involve all the visitors or can be organized into private chat groups. The first bot, which the visitors meet, will greet them immediately upon arrival (Figure 2). This bot is an avatar of one of the project students who contributed a lot to the Virtual Campus. Its "brain" is developed using AIML language, ALICE files [1], and computer graphics terms from the computer graphics text book [18], which this bot is incidentally selling to the current students. There are also a few other agents wandering around and inside the academic buildings. They are also "clones" of the former project students. In fact, each of the project students has a personal avatar copy in the Virtual Campus.

Virtual Campus is a place for research on crowd simulation and shared cyberspaces. Its content changes frequently. You can come across an avatar, which is in fact a bot, and it will take time before you understand it. Sometimes it may be a real person disguised as a bot to test human reaction on some avatar activities to be programmed.

Fig. 2. Robots and avatars

Electronic education is one of the priority directions at NTU. The University's e-learning platform edveNTUre [5] is based on the BlackBoard software [4], and several other companion software tools. It is extensively used by the NTU professors to enhance their lectures. Since its introduction in 2001, edveNTUre has developed from a rather exotic way of publishing lecture materials and occasional visits by the

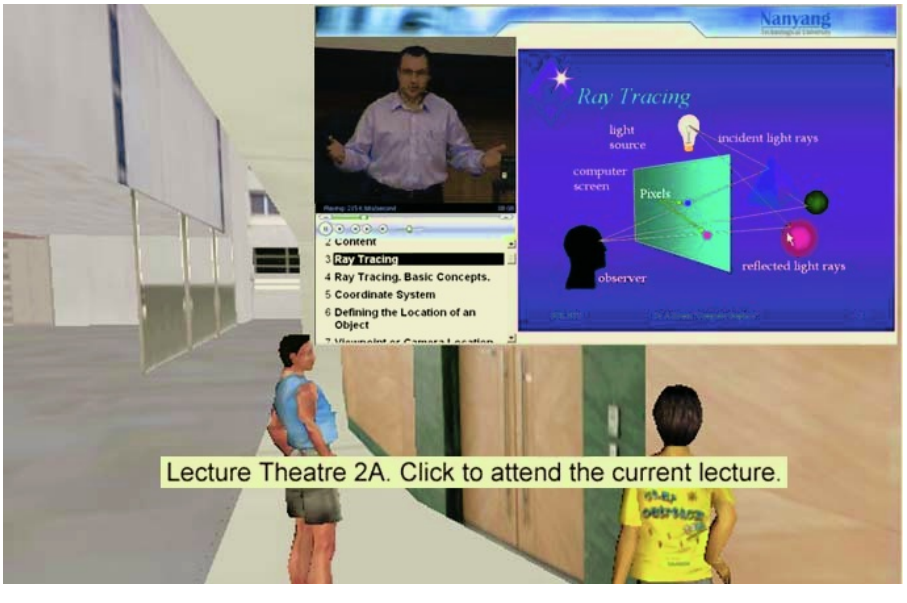


Fig. 3. Combining VRML with streaming video recordings of lectures and events

students to the present time, when it has become a compulsory and very important part of each course with thousands of visits each day. Besides teaching materials such as lecture notes, slides, streaming audio/video presentations, and extra materials, it can be used for setting up on-line quizzes, discussion groups, and uploading assignments. However, being based on html web-pages edveNTUre rather gives a "two-dimensional look" of the teaching process. In contrast and in addition to it, on the Virtual Campus NTU professors are able to meet with their students in virtual 3D classrooms, "see" and communicate with each other, and so add more immersion and fun to education. Besides that, distant overseas students get a feeling of really being on campus. Many features available in edveNTUre are also available on the Virtual Campus. Thus, some of the virtual lecture theatres and other places are linked to streaming multimedia presentations of current and pre-recorded lectures and events (Figure 3).

Of course Virtual Campus is a learning tool for computer graphics students illustrating to them theoretical concepts of virtual reality, real time rendering, and shape modeling. It is used during lectures, as well as after classes for consultations.

3 Collaborative Shape Modeling Hands-on Experience

One of the student assignments is "Implicit Fantasies", which is to design sophisticated shapes using implicit functions, and to make them available in their virtual homes on the Virtual Campus. A part of this assignment is to be done in the Collaborative Shape Modeling Laboratory, which is a part of the Virtual Campus. This virtual laboratory can be entered either from the lobby of the School of Computer Engineering of the Virtual Campus or by a direct link.

Before going there, the visitors have to install a small software plug-in. This plug-in is an extension of VRML, which allows for defining geometric shapes with analytical formulas. By "formulas" analytical definitions with parametric, implicit, and explicit FRep [2, 16] functions are understood:

Parametric representation:	FRep representation:
$x = f_1(u, v, w)$	$g = f(x, y, z)$
$y = f_1(u, v, w)$	inside the shape: $g > 0$; on the surface: $g = 0$;
$z = f_1(u, v, w)$	outside the shape: $g < 0$

All these formulas are functions of three coordinates, which are either parametric or Cartesian coordinates of 3D shapes. In our hybrid model, these different representations can be used concurrently for defining geometry and appearance of shapes. The shape's geometry can be defined by a basic geometric shape and its geometric texture, each defined with either parametric, implicit, or explicit functions. The appearance of shapes can be defined by either function-defined or fixed colors. Similarly to the shape's geometry, parametric, implicit or explicit functions can be used for defining the shape's color on its surface and inside it. When using this hybrid function-based model, geometric textures and appearances are often thought of as other function-defined geometric objects in their coordinate spaces, which are merged with the function-defined underlying geometry of a shape to create a new geometric object. This

approach, though unusual from the first glance, in fact helps the students to better understand the concepts of function based shape modeling. Also, the synergy of using the three different representations results in the advance quality and efficient solutions, which are impossible to achieve when these representations are used on their own. The theoretical foundations and further details of this approach can be found in [13, 14] as well as in two project web pages on Function-based Web Visualization [6], and Interactive Function-based Shape Modeling [10].

After the plug-in is installed, besides the regular VRML objects, function-defined shapes will become visible as well. There will be one big shape hovering in the middle of the room, as well as a few smaller fancy shapes displayed in different parts of the room (Figure 4).

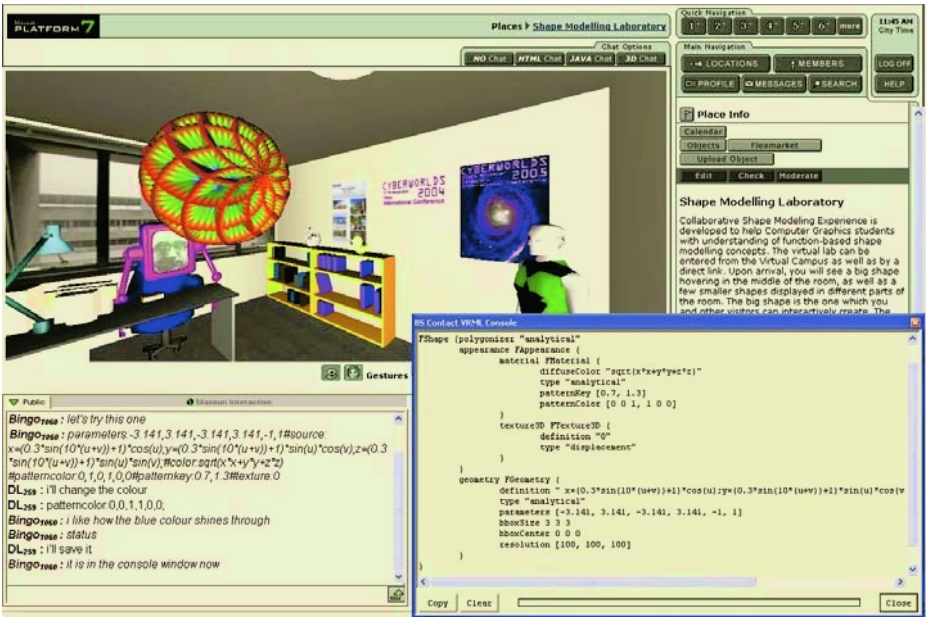


Fig. 4. Collaborative shape Modeling Laboratory

The bigger shape is the one, which the visitors can interactively modify. The smaller function-defined shapes are examples of the best works created by the students. The function-defined shapes can be placed to other part of the Virtual Campus as well, e.g. to private virtual homes of the members. Several visitors may discuss the design in the chat box, type individual shape modeling commands or command scripts, and immediately see how the shape changes accordingly. The VRML description of the shape being modeled can be displayed at any time and saved for future use. Figure 5 illustrates different ways of defining the shape displayed in Figure 4 with individual command, command script, and VRML file.

Interactive commands:

```
parameters:-3.141,3.141,-3.141,3.141,-1,1
source:x=(0.3*sin(10*(u+v))+1)*cos(u);
y=(0.3*sin(10*(u+v))+1)*sin(u)*cos(v);
z=(0.3*sin(10*(u+v))+1)*sin(u)*sin(v);
color:sqrt(x*x+y*y+z*z)
patterncolor:0,1,0,1,0,0
patternkey:0.7,1.3
```

Command script:

```
parameters:-3.141,3.141,-3.141,3.141,-1,1#source:
x=(0.3*sin(10*(u+v))+1)*cos(u);y=(0.3*sin(10*(u+v))+1)*
sin(u)*cos(v);z=(0.3*sin(10*(u+v))+1)*sin(u)*sin(v);
#patterncolor: 0,1,0,1,0,0#patternkey:0.7,1.3#color:
sqrt(x*x+y*y+z*z)
```

VRML code using the function-based extension:

```
FShape {
appearance FAppearance {
  material FMaterial {
   diffuseColor "sqrt(x*x+y*y+z*z)"
   patternColor [0 1 0 1 0 0]
   patternKey [0.7 1.3]
   type "analytical"  }
   }

geometry FGeometry {
  resolution [200 200]  parameters [-1 1 -1 1]
  definition "
   x=(0.3*sin(10*(u*pi+v*pi))+1)*cos(u*pi);
   y=(0.3*sin(10*(u*pi+v*pi))+1)*sin(u*pi)*cos(v*pi);
   z=(0.3*sin(10*(u*pi+v*pi))+1)*sin(u*pi)*sin(v*pi);"
  type "analytical"
  }
}
```

Fig. 5. Individual commands, scripts and VRML code defining the shape in Figure 4

Besides this method of modeling by concurrent typing analytical formulas with an immediate visual feedback, more complex function-based shape modeling can be done with the interactive shape modeling program developed for this project. The

Fig. 6. Virtual embossing Tool. The shape's geometry and appearance have been interactively defined. The resulting model consists of 8,000 explicit functions

program offers an advanced set of interactive operations such as cutting, sculpting, embossing, engraving, and carving. It also allows for interactive painting both on the surface and inside the object. As a result, the program allows for making realistically looking shapes, which are defined with very small function-defined models while can be rendered with any desired precision (Figure 6). The initial basic shape for modeling can be either defined analytically (Figure 7), or created interactively with simple basic shapes. The initial shape is then gradually modified by applying different interactive shape modeling and/or painting operations. All the interactive modifications as well as the colors become an integral part of the function-based model of the shape defined with explicit functions of three Cartesian coordinates. The result of the modeling can be either saved in the proprietary function-based data format or in the func-

tion-based VRML code for further use in the Collaborative Shape Modeling Hands-On Experience or in other shared virtual worlds.

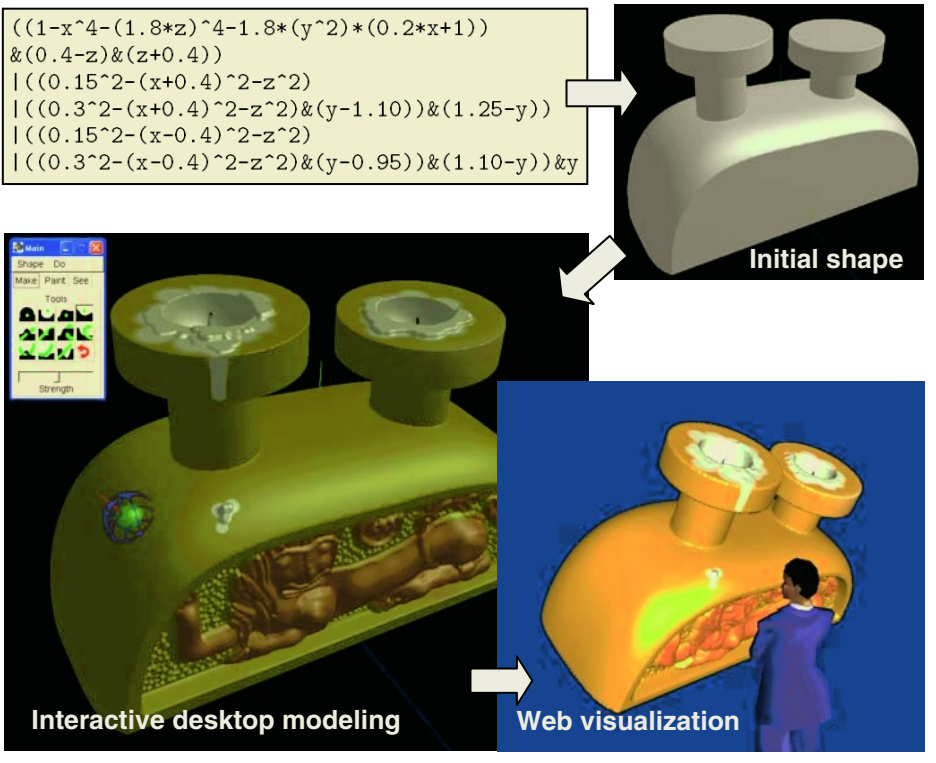


Fig. 7. The basic shape is defined analytically, then it is interactively carved and painted, and finally imported to function-based extension of VRML for web visualization and further modeling

4 Conclusion

Our experience of teaching computer graphics to large and distributed classes of students with different educational and cultural backgrounds proved that shared cyberworlds are efficient tools, which both illustrate how computer graphics, shape modeling, and virtual reality work and provide the students hands-on experiences by allowing them to contribute to these worlds.

When designing cyberworlds, we extensively use hybrid function-based shape modeling, which proved to be an efficient way of expanding shared VRML worlds. Many large VRML models, which required a big number of polygons, were replaced with compact function-based models where shapes and their appearances are defined with small parametric, implicit and explicit formulas.

Designing shapes and their appearances with concurrent using implicit, analytical and parametric functions helped the students to better understand the concepts of shape modeling.

References

1. ALICE files: http://www.alicebot.org
2. Bloomenthal, J. (ed.): Introduction to Implicit Surfaces. Morgan Kaufmann (1997)
3. BlackBoard software: http://www.blackboard.com
4. Blaxxun communication platform: http://www.blaxxun.com
5. edveNTUre: http://edventure.ntu.edu.sg
6. Function-based Web Visualization: http://www.ntu.edu.sg/home/assourin/FVRML.htm
7. Gerval, J.-P., Popovici, D.-M., and Tisseau, J.: Educative Distributed Virtual Environments for Children. 2003 International Conference on Cyberworlds, IEEE CS Press (2003) 382-387
8. Green, D., Cosmas, J., Degeest R., and Waelkens, M. A.: Distributed Universal 3D Cyberworld for Archaeological Research and Education. 2003 International Conference on Cyberworlds, IEEE CS Press (2003) 458-465
9. Gutiérrez, M., Thalmann, D., Frédéric, V.: Creating Cyberworlds: Experiences in computer Science Education. International Conference on Cyberworlds, IEEE CS Press (2004) 401-408
10. Interactive Function-based Shape Modeling: http://www.ntu.edu.sg/home/assourin/Intshape.html.
11. Kenzo Tange Associates: http://www.ktaweb.com/works/en_nti.html
12. Kuni, T.L., Luciani, A. (eds.): Cyberworlds. Springer Verlag (1998)
13. Levinski, K., Sourin, A.: Interactive Function-Based Shape Modeling for Cyberworlds, 2004 International Conference on Cyberworlds, IEEE CS Press (2004) 54-61
14. Liu, Q., Sourin, A.: Analytically-defined Collaborative Shape Modeling in VRML. 2004 International Conference on Cyberworlds, IEEE CS Press (2004) 70-77
15. Nanyang Technological University: http://www.ntu.edu.sg
16. Pasko, A.A, Adzhiev, V.D., Sourin, A.I., Savchenko, V.V.: Function Representation in Geometric Modeling: concepts, implementations and applications, The Visual Computer, 11(8), (1995) 429-446
17. Shin, Y.-S.: Virtual experiment environments design for science education. 2003 International Conference on Cyberworlds, IEEE CS Press (2003) 388–395
18. Sourin, A.: Computer Graphics. From a Small Formula to Virtual Worlds. Singapore: Prentice Hall (2004)
19. Virtual Campus of NTU: http://www.ntu.edu.sg/home/assourin/vircampus.html

Visualization and Analysis of Protein Structures Using Euclidean Voronoi Diagram of Atoms

Deok-Soo Kim[1,2], Donguk Kim[2], Youngsong Cho[2], Joonghyun Ryu[2], Cheol-Hyung Cho[2], Joon Young Park[3], and Hyun Chan Lee[4]

[1] Dept. of Industrial Engineering, Hanyang University, Seoul, Korea
dskim@hanyang.ac.kr
[2] Voronoi Diagram Research Center, Hanyang University, Seoul, Korea
{donguk, ycho, jhryu, murick}@voronoi.hanyang.ac.kr
[3] Dept. of Industrial and Systems Engineering, Dongguk University, Seoul, Korea
jypark@dgu.edu
[4] Dept. of Industrial Engineering, Hongik University, Seoul, Korea
hclee@wow.hongik.ac.kr

Abstract. Protein consists of amino acids, and an amino acid consists of atoms. Given a protein, understanding its functions is critical for various reasons for designing new drugs, treating diseases, and so on. Due to recent researches, it is now known that the structure of protein directly influences its functions. Hence, there have been strong research trends towards understanding the geometric structure of proteins. In this paper, we present a Euclidean Voronoi diagram of atoms constituting a protein and show how this computational tool can effectively and efficiently contribute to various important problems in biology. Some examples, among others, are the computations for molecular surface, solvent accessible surface, extraction of pockets, interaction interface, convex hull, etc.

1 Introduction

A molecule consists of atoms. For example, a protein consists of amino acids, and an amino acid consists of atoms. Suppose that we have a biological molecule such as a protein, DNA or RNA. Given an atomic biology system such as a protein, understanding its functions is very important in the design of new drugs and treating diseases, and so on [19, 20].

While scientific efforts since the early days of science have been mainly focused on the understanding physicochemical aspect of a molecular system, there have been efforts to understand the geometric structure of the system as well [21, 24, 25, 26]. One important example is the discovery of double helix structure of DNA by Watson and Crick [31].

Ever since Watson and Crick, similar efforts have been made extensively for proteins as well. Especially in proteins, it is now known that the structure of a protein directly determines its functions. Therefore, there have been recently very strong research activities towards understanding the geometric structure

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 993–1002, 2005.

of proteins [3, 5, 7, 13, 18, 20, 25, 26, 29, 32]. Nowadays, every discovery on protein structure regarding on the atom coordinate is publicized in Protein Data Bank(PDB) [33] and can be freely accessed.

To understand the structure of proteins, researchers have developed various computational concepts and tools. The work by Lee and Richards on the definition of *solvent accessible surface*(SAS) is the first effort on this avenue [20]. A SAS is the set of centers of a spherical probe rolling around the protein. A probe is used for the computational convenience of small molecule which interacts with the protein. Hence, the SAS provides information on the free space that a small molecule can move without penetrating the protein [20].

Connolly defined a *molecular surface*(MS) and showed its computation in a computer graphics display [7]. A MS consists of the most inward points on the probe toward the interior of protein when the probe is in contact with the protein. A MS plays the role of basis for the frequently asked questions such as the volume and surface area of the protein, the volume of space where a particular atom has more influence than others, etc.

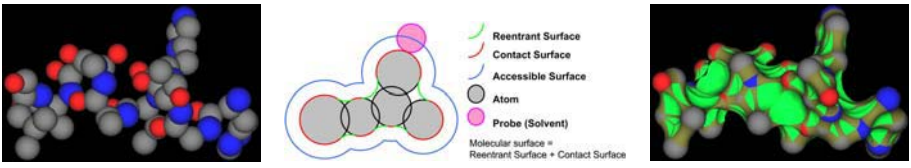
Many queries regarding on the geometric structure of proteins need to be answered efficiently and this requirement usually involve the computations on both geometry and topology. For example, intersection between spherical triangles is a necessary geometric computation for MS while the decision for which pair of spherical triangles needs to be tested is an example of topological computation for MS. As the technology progresses, questions of higher level abstractions in a molecule are also often asked such as the existence of pockets on the surface of protein [1, 8, 18, 19, 21, 24]. It is quite well-known that atoms located exterior part of a protein determines the function of the protein. Hence, a MS is important in the study of functions of proteins since the surface has direct relation with exterior atoms. To facilitate efficient computations for such important concepts on proteins, researchers tried to use various kinds of computational tools.

Discussed in this paper is the computational tool, Voronoi diagram of atoms, convenient for the computation of neighborhood search and its application for various problems for the study of structure of proteins. We believe that this relatively new application area will provide new opportunities for the CAD/CAM and geometric modelling community in the future.

2 Geometric Model of Molecule and Related Works

Shown in Fig. 1(a) is a part of molecule downloaded from PDB. This model consists of 64 atoms and forms an α-helix. The spheres are van der Waals surfaces of associated atoms in the protein. There are usually two kinds of surfaces associated with a protein model: *Solvent Accessable Surface(SAS)* and *Molecular Surface(MS)*.

To better explain both SAS and MS, we introduce a geometric model of a protein as shown in Fig. 1(b). SAS consists of a set points on the space where the center of probe is located when the probe is in contact with the protein. Hence, SAS consists of a set of points where a point corresponds to the center of



(a) (b) (c)

Fig. 1. Protein data and related geometric structures. (a) an α-helix consisting of 67 atoms (a part of protein data downloaded from PDB), (b) a geometric model for atomic structure, and (c) a molecular surface

probe when the probe is in contact with some atoms of the protein. The innermost possible trajectories of points on the probe surface, then, define MS. SAS usually defines a *free-space* that a small molecule can move around without interfering the protein and therefore plays the fundamental role for folding and/or docking [20]. On the other hand, MS, often called by another name *Connolly surface* after the name of the first researcher defined the surface, conveniently defines the boundary between interior and exterior volume of a protein so that the volume or the density of protein can be calculated [7].

MS, sometimes called a *solvent excluded surface*(SES) again consists of two parts: *Contact surface* and *reentrant surface*. *Contact surface* consists of points on the van der Waals surface of atoms which can be contacted by the probe without interfering any atom in the protein, and *reentrant surface* consists of points in the free-space touched by the probe when the probe is tangential to nearby atoms in the protein. Note that atoms contributing the contact surface define the boundary of the protein. Hence, points on the MS are always accessible by the probe as it rolls over the protein. Fig. 1(c) shows the MS of the previous α-helix.

In the geometric modelling community, the reentrant surface has been called a *blending surface* and its computation has been studied quite extensively in a rather general setting. The *SAS* has been also known in the geometric modelling community as an *offset surface* of a protein using the probe radius as an offset distance. Note that the definitions of all surfaces depends on the probe. Note that the notion of SAS is similar to the concept of *configuration space*, *C-space*, in robotics community [22]. Note that a water molecule is frequently used as a solvent around a protein and the corresponding probe is approximated by a sphere with a radius of 1.4 Å. Note that the work by Lee and Richards on the definition of SAS is the first effort on this avenue [20].

3 Euclidean Voronoi Diagram of Spheres: EVD(S)

In the processing of various queries regarding on geometric structure of proteins, it is often necessary to make decisions regarding on the topology among associated atoms. To facilitate the decision making efficient, transforming the atom set into an organized data structure is also necessary.

To effectively and efficiently answer to all of these and perhaps more important questions for a given molecule, the efficient representation of spatial structure for the given molecule is necessary. Upon this requirement, people in biology have been trying to use the ordinary Voronoi diagram of points for such a representation [1, 12, 23, 25, 30]. There has been also other nice approaches such as power diagram and α-hull for the representation of the topological structure of atoms [21]. However, the differences of radii among atoms and the Euclidean nature of the problem have not been appropriately incorporated in these approaches.

While an ordinary Voronoi diagram for point set has been extensively studied, Euclidean Voronoi diagram for spheres in 3D has not been sufficiently explored. Even the efficient and robust construction of Euclidean Voronoi diagram for circles in a plane has been available only very recently [14, 15]. Similarly, there are only a few previous works on the Euclidean Voronoi diagram of spheres [2, 4, 10, 11, 16, 17, 18, 32].

Let $B = \{b_1, b_2, \ldots, b_n\}$ be a set of generators for a Voronoi diagram where b_i is a 3-dimensional spherical ball. Hence, $b_i = (c_i, r_i)$ where $c_i = (x_i, y_i, z_i)$ and r_i denote the center and the radius of a ball b_i, respectively. We assume that no ball is completely contained inside another ball even though intersections are allowed between balls.

Associated with each ball b_i, there is a corresponding *Voronoi region* VR_i for b_i, where $\mathrm{VR}_i = \{p \mid dist(p, c_i) - r_i \leq dist(p, c_j) - r_j, i \neq j\}$. Then, $\mathrm{EVD}(B) = \{\mathrm{VR}_1, \mathrm{VR}_2, \ldots, \mathrm{VR}_n\}$ is called a *Euclidean Voronoi diagram* for B. The topology of the Voronoi diagram in this paper is represented as an *edge-graph* $G = (V, E)$ where $V = \{v_1, v_2, \ldots\}$ and $E = \{e_1, e_2, \ldots\}$ are sets of Voronoi vertices and edges, respectively. Once G is available, the face set $F = \{f_1, f_2, \ldots\}$ of Voronoi diagram can be easily constructed. In this paper, the ordinary L_2-metric is used to define a Euclidean Voronoi diagram. In other words, $dist(c_i, c_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$.

As in ordinary point set Voronoi diagrams, Voronoi regions corresponding to balls on the boundary of the convex hull of B are unbounded. Other regions are bounded by a set of faces, called *Voronoi faces*, where a Voronoi face is defined by two immediately neighboring balls. Note that a Voronoi face is always a connected-subset of hyperboloid of two sheets. A Voronoi face intersects another Voronoi face to form a *Voronoi edge* which is a conic section. When Voronoi edges intersect, a *Voronoi vertex* is defined.

4 Computational Primitives for EVD(S)

There are three computational constituents for the geometry of EVD(S) in 3D: vertices, edges, and faces.

4.1 Voronoi Vertices

A Voronoi vertex is defined among four appropriate neighboring atoms, and it corresponds to the center of a sphere tangent to the atoms. An elegant algorithm to compute the tangent spheres is presented by Gavrilova [11]. She has shown

that the tangent spheres can be computed by solving a system of quadratic equations where the system is obtained by an explicit formulation of equi-distant points from surfaces of four balls.

4.2 Voronoi Edges

Unless it is a degenerate case, an edge is always defined as a locus of points equi-distant from three surrounding balls. Then, it can be also easily shown that a Voronoi edge is the solution of a set of simple non-linear equations, in a similar way for the Voronoi vertex computation. It turns out that a Voronoi edge is always a conic curve. Therefore, the precise geometry of Voronoi edges can be conveniently represented by rational quadratic Bézier curves if five parameters are known [9]. In the middle of this computation, we have to solve the well-known Apollonius 10th problem as well and its solution process is well-described in [15, 27].

4.3 Voronoi Faces

Two topologically neighboring balls define a Voronoi face as a locus of points equi-distant from the two balls. It turns out that Voronoi face is a part of a hyperboloid and its implicit equation can be easily obtained [13].

Having the face equation, it may seem better to represent faces in a parametric form. Since the Voronoi face in our problem is always a hyperboloid, the natural choice for the parametric representation seems to be a rational Bézier surface with an appropriate degrees and topology.

However, the principal uses of Voronoi faces are the computations of volumes and boundary areas of Voronoi regions, and the visualization of Voronoi regions. Hence, we use the following approach for the representation of Voronoi faces.

Suppose that two balls are transformed so that the center of larger ball is located at the origin of the coordinate system and the line defined by the centers of both balls coincides to Z-axis. Then, the Voronoi face between the balls becomes a hyperboloid which is a single-valued function w.r.t. x and y. Voronoi edges, already in a rational quadratic Bézier curve form, can be also similarly transformed. Then, the projection of these transformed edges in xy-plane defines the boundary of domain on the plane for the Voronoi face. If Voronoi faces are represented in this way, both evaluation of a point on the face and testing if a point is on the Voronoi face become much simpler.

5 Topology Construction Algorithms for EVD(S)

While there can be different ways to compute EVD(S), we briefly describe two algorithms, an *edge-tracing algorithm* and a *region-expansion algorithm*, to construct the topologies of the diagram.

5.1 Edge-Tracing Algorithm

The basic idea of the edge-tracing algorithm is quite simple yet powerful. Based on this idea, we recently reported a full implementation of the edge-tracing algorithm with discussions on various applications [17, 18].

Our edge-tracing algorithm is as follows. The algorithm first locates a true Voronoi vertex by computing an empty tangent sphere defined by four appropriate nearby balls. Then, four edges emanating from the initial vertex is identified and pushed into a stack called Edge-stack. Note that those edges have the initial vertex as their starting vertices. After popping an edge from the stack, the algorithm computes the end vertex of the popped edge. Note that a vertex can be found by computing a tangent sphere from each of $n-3$ balls plus three balls defining the popped edge and testing if the corresponding tangent sphere is empty. If an empty tangent sphere is found, the center of the sphere becomes the candidate for the end vertex of the popped edge. In the case that there are several such empty tangent spheres, one closest to the starting vertex of the edge is chosen.

Once the end vertex of currently popped edge is found, it is also possible to define three more edges emanating from this new vertex. Hence, three edges are created and the new vertex is used as the starting vertex of three new-born edges. Note that these edges are also pushed into Edge-stack. By following this process until Edge-stack is empty, the computation of Voronoi diagram is completed. Even though the idea is simple, designing a correct and efficient algorithm is not so easy at all.

The worst-case time complexity for the whole Voronoi diagram takes $\mathrm{O}(nm)$ where n is the number of atoms and m is the number of edges in the diagram. Note that m can be $O(n^2)$ in the worst-case. However, the running time can be reduced as low as $O(n)$ by employing various acceleration techniques.

5.2 Region-Expansion Algorithm

Suppose that we are given with an ordinary Voronoi diagram for the centers of atoms constituting a protein. Then, each atom is associated with a Voronoi region even though this is not correct for the atom itself.

Consider that an atom is being expanded starting from a point, located at the center of the atom, to its van der Waals sphere. While a region is expanding, we want to keep the topology and geometry of the vertices, edges, and faces of the region correct. Repeating this region expansion process for all atoms eventually constructs the valid Voronoi diagram of all atoms if each process can perform the specified task correctly.

It is obvious that growing the size of an atom always increases the volume of the corresponding Voronoi region. It can be easily shown that each vertex on a Voronoi region, during the region expansion, moves away from the initial region by following an edge associated with the vertex. Similarly, each edge on a region moves away from the initial region by following a face. Note that a sufficiently small growth of expanding generator leaves the combinatorial structure of the diagram unchanged but causes changes only in the geometries of vertices, edges, and faces related to the region boundary.

However, certain changes may occur in topological structure at some point of time in the expanding process. For example, suppose that a vertex moves along

an edge to meet a corresponding end vertex of the edge. Then, the edge shrinks and degenerates to a point and disappears afterwards. We call an *event* for such a situation causing changes in the combinatorial structure.

It can be shown that events are classified in four different cases: i) *one-end-event*, ii) *two-end-event*, iii) *mid-event*, and iv) *split-event*. An emphend-event denotes the case when an edge disappears at the end of the edge after the process of edge-shrinking. In 3D, it can be shown that one or two edges are incident to the end vertex of a shrinking edge and these cases corresponds to different events. Hence, we call the situations *one-end-event* or *two-end-event*. A mid-event denotes the case that an edge disappears in the middle of the edge. This case occurs when both end vertices of an edge move toward interior of the edge to meet at a point. Similarly, a split-event denotes the case that a new vertex is created at a point in the middle of an edge so that the edge splits into two edges. According to an event type, associated operations on the topology structure can be also clearly defined. While the event type is decided, the time for the event can be also computed by considering the geometric configurations in the neighborhood.

The computation necessary to expand a Voronoi region, starting from a center point to a complete sphere, takes $O(n^2 \log n)$ time in the worst-case since there can be $O(n^2)$ number of edges to be considered and sorting events is required according to the event time. We believe, however, that the expected time complexity to expand a Voronoi region can be as lower as $O(n)$ in the region expansion algorithm.

6 Protein Structure Analysis Using EVD(S)

Voronoi diagram can be efficiently used for the analysis of protein structure as illustrated in the following examples. Fig. 2(a) shows a part of protein data 1BH8 downloaded from PDB [33]. 1BH8 consists of two groups of atoms interacting each other and shown in Fig. 2(a) is the group A. Note that 1BH8 consists of 1,074 atoms(680 C's, 181 N's, 203 O's, and 10 S's) and the group A of 1BH8 consists of 366 atoms(233 C's, 54 N's, 74 O's, and 5 S's).

Fig. 2(b) shows the convex hull of the model computed from the topology data of Voronoi diagram of the protein. Note that the convex hull, in this example, consists of triangular faces connecting the centers of atoms can take only linear time w.r.t. the number of Voronoi faces. Fig. 2(c) illustrates the molecular surface of the protein using a probe of size 1.4 Åwhich denotes the size of a water molecule. Details will be elaborated in another paper in [28]. Fig. 2(d) shows pockets on the protein where ligands can dock into so that a new drug discovery can be tried on . How to detect such pockets is beyond the scope of this article and therefore will be reported in another paper [6]. While Fig.'s 2(a), (b), (c) and (d) illustrate examples for only group A of protein 1BH8, Fig. 2(e) shows the interaction interface between two groups. Note that points on the surface is a mid-point in between two closest atoms. Usually the shape of such an interaction surface is of importance. Fig. 2(f) shows only the interaction surface ignoring the

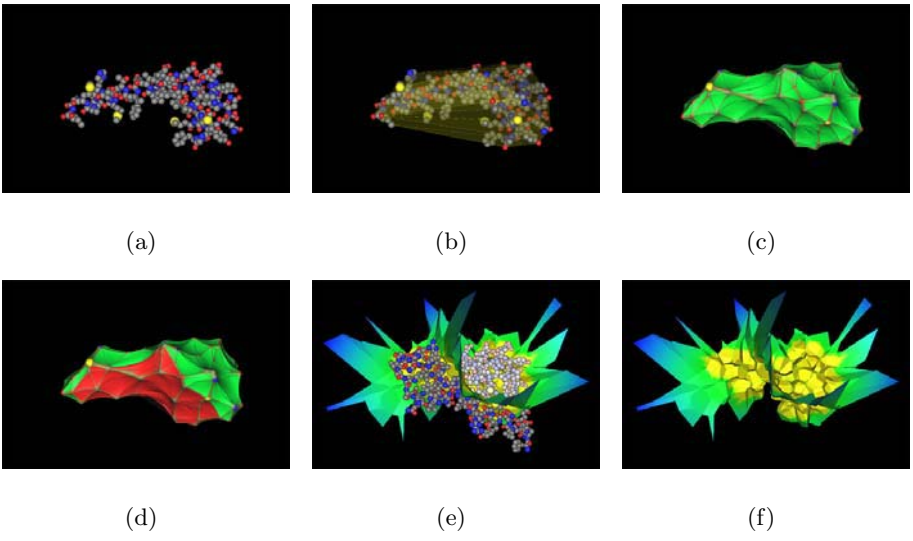


Fig. 2. Various geometric structures related to the protein 1BH8. (a) the van der Waals atoms of group A in the protein 1BH8, (b) the convex hull of group A, (c) the molecular surface of group A, (d) the pocket on the molecular surface of group A, (e) the interaction interface of the protein 1BH8, with atoms, (f) the interaction interface of the protein 1BH8, without atoms

atoms. The intensity on the surface actually illustrates the Euclidean distance between two closest atoms in both groups.

7 Conclusions

Protein consists of amino acids, and an amino acid consists of atoms. Given a protein, understanding its functions is critical for various reasons for designing new drugs, treating diseases, and so on. Since the structure of protein primarily determines its functions, there have been strong needs for the research to understand the geometric structure of proteins in both detailed and abstract levels.

In this paper, we have shown that a Euclidean Voronoi diagram of atoms constituting a protein can facilitate the computations necessary for the analysis of protein structure in various applications. The convex hull, molecular surface, extracting pockets, isolating interaction interface, and so on can be easily computed once the Voronoi diagram of atoms is available. In addition, we have briefly reviewed two important and practical approaches to compute the Voronoi diagrams.

Acknowledgements

This research was supported by Creative Research Initiatives from the Ministry of Science and Technology, Korea.

References

1. Angelov, B., Sadoc, J.-F., Jullien, R., Soyer, A., Mornon, J.-P., Chomilier, J.: Nonatomic solvent-driven Voronoi tessellation of proteins: an open tool to analyze protein folds. Proteins: Structure, Function, and Genetics **49** (2002) 446–456.
2. Aurenhammer, F.: Power diagrams: properties, algorithms and applications. SIAM Journal of Computing **16** (1987) 78–96.
3. Bajaj, C. L., Lee, H. Y., Merkert, R., Pascucci, V.: NURBS based B-rep models for macromolecules and their properties, In Proc. 4th Symposium on Solid Modeling and Applications, (1997) 217-228.
4. Boissonnat, J.D., Karavelas, M.I.: On the combinatorial complexity of Euclidean Voronoi cells and convex hulls of d-dimensional spheres. in Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (2003) 305-312.
5. Cheng, H.-L., Dey, T. K., Edelsbrunner, H., Sullivan, J.: Dynamic skin triangulation, Discrete & Computational Geometry **25** (2001) 525–568.
6. Kim, D.-S., Cho, C.-H., Cho, Y., Won, C.I., Kim, D.: Pocket recognition on a protein using Euclidean Voronoi diagram of atoms, ICCSA2005 conference, submitted.
7. Connolly, M, L.: Solvent-accessible surfaces of proteins and nucleic acids, Science **221** (1983) 709–713.
8. Edelsbrunner, H., Facello, M., Liang, J.: On the definition and the construction of pockets in macromolecules. Discrete Applied Mathematics. **88** (1998) 83–102.
9. Farin, G.: Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide, 4th edn, Academic Press, San Diego (1996).
10. Gavrilova, M.: Proximity and Applications in General Metrics. Ph.D. thesis: The University of Calgary, Dept. of Computer Science, Calgary, AB, Canada (1998).
11. Gavrilova, M., Rokne, J.: Updating the topology of the dynamic Voronoi diagram for spheres in Euclidean d-dimensional space. Computer Aided Geometric Design **20** (2003) 231–242.
12. Gerstein, M., Tsai, J., Levitt, M.: The volume of atoms on the protein surface: calculated from simulation, using Voronoi polyhedra. Journal of Molecular Biology **249** (1995) 955–966.
13. Goede, A., Preissner, R., Frömmel, C.: Voronoi cell: new method for allocation of space among atoms: elimination of avoidable errors in calculation of atomic volume and density. Journal of Computational Chemistry **18** (1997) 1113–1123.
14. Kim, D.-S., Kim, D., Sugihara, K.: Voronoi diagram of a circle set from Voronoi diagram of a point set: I. Topology. Computer Aided Geometric Design **18** (2001) 541–562.
15. Kim, D.-S., Kim, D., Sugihara, K.: Voronoi diagram of a circle set from Voronoi diagram of a point set: II. Geometry. Computer Aided Geometric Design **18** (2001) 563–585.
16. Kim, D.-S., Cho, Y., Kim, D., Cho, C.-H.: Protein structure analysis using Euclidean Voronoi diagram of atoms, Proc. International Workshop on Biometric Technologies (BT 2004), Special Forum on Modeling and Simulation in Biometric Technology (2004) 125–129.
17. Kim, D.-S., Cho, Y., Kim, D.: Edge-tracing algorithm for Euclidean Voronoi diagram of 3D spheres, in Proc. 16th Canadian Conference on Computational Geometry (2004) 176–179.
18. Kim, D.-S., Cho, Y., Kim, D., Kim, S., Bhak, J., Lee, S.-H.: Euclidean Voronoi diagram of 3D spheres and applications to protein structure analysis, In Proc. International Symposium on Voronoi Diagrams in Science and Engineering (2004) 137–144.

19. Kunts, I.D.: Structure-based strategies for drug design and discovery. Science. **257** (1992) 1078–1082.
20. Lee, B., Richards, F.M.: The interpretation of protein structures: estimation of static accessibility, Journal of Molecular Biology **55** (1971) 379–400.
21. Liang, J., Edelsbrunner, H., Woodward, C.: Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. Protein Science. **7** (1998) 1884–1897.
22. Lozano-Perez, T.: Spatial planning: a configuration space approach, IEEE Transactions on Computers **C-32** (1983) 108–120.
23. Montoro, J.C.G., Abascal, J.L.F.: The Voronoi polyhedra as tools for structure determination in simple disordered systems. The Journal of Physical Chemistry **97** (1993) 4211–4215.
24. Peters, K.P., Fauck, J., Frömmel, C.: The automatic search for ligand binding sites in protein of known three-dimensional strucutre using only geometric criteria. Journal of Molecular Biology **256** (1996) 201–213.
25. Richards, F.M.: The interpretation of protein structures: total volume, group volume distributions and packing density. Journal of Molecular Biology **82** (1974) 1–14.
26. Richards, F.M.: Areas, volumes, packing and protein structure, Annu. Rev. Biophys. Bioeng. **6** (1977) 151–176.
27. Rokne, J.: Appolonius's 10th problem. Graphics Gems II (Edited by Arvo, J., Academic Press (1991) 19–24.
28. Ryu, J., Kim, D., Cho, Y., Park, R., Kim, D.-S.: Computing molecular surfaces of proteins, ICCSA2005 conference, submitted.
29. Varshney, A., Brooks, F.P., Jr., Wright, W.V.: Computing smooth molecular surfaces, IEEE Computer Graphics and Applications **14** (1994) 19–25.
30. Voloshin, V.P., Beaufils, S., Medvedev, N.N.: Void space analysis of the structure of liquids. Journal of Molecular Liquids **96-97** (2002) 101–112.
31. Watson, J.D., Crick, F.H.C.: A structure for deoxyribose nucleid acid. Nature **171** (1953) 737–738.
32. Will, H.-M.: Computation of Additively Weighted Voronoi Cells for Applications in Molecular Biology. Ph.D. thesis, ETH, Zurich (1999).
33. RCSB Protein Data Bank Homepage: `http://www.rcsb.org/pdb/` (2004).

C^2 Continuous Spline Surfaces over Catmull-Clark Meshes*

Jin Jin Zheng[1], Jian J. Zhang[2], Hong Jun Zhou[3], and L.G. Shen[1]

[1] Department of PMPI, University of Science and Technology of China,
Hefei, Anhui, P R China 230026
{jjzheng, lgshen}@ustc.edu.cn
[2] NCCA, Bournemouth University,
Poole, Dorset BH12 5BB, UK
jzhang@bournemouth.ac.uk
[3] NSRL, University of Science and Technology of China,
Hefei, Anhui, P R China 230026
hjzhou@ustc.edu.cn

Abstract. An efficient method for generating a C^2 continuous spline surface over a Catmull-Clark mesh is presented in this paper. The spline surface is the same as the Catmull-Clark limit surface except in the immediate neighborhood of the irregular mesh points. The construction process presented in this paper consists of three steps: subdividing the initial mesh at most twice using the Catmull-Clark subdivision rules; generating a bi-cubic Bézier patch for each regular face of the resultant mesh; generating a C^2 Gregory patch around each irregular vertex of the mesh. The union of all patches forms a C^2 spline surface. Differing from the previous methods proposed by Loop, DeRose and Peters, this method achieves an overall C^2 smoothness rather than only a C^1 continuity.

1 Introduction

Although the mainstream parametric surface modeling methods, such as B-splines are powerful and have been adopted into many commercial CAD and computer graphics packages, they suffer from a fundamental limitation that they require the control mesh to form a regular quadrilateral structure.

To overcome this limitation, a number of methods have been proposed for the construction of smooth surfaces of irregular topology. Roughly speaking, these methods are categorized into two groups, the first group describes the final surface with a closed analytical form; the second group produces the final surface as the limit of a refinement procedure.

1.1 Subdivision Surfaces

Recursive subdivision surfaces were developed to tackle the limitation of the topological irregularity by Catmull & Clark [1] and Doo & Sabin [2]. Starting

* This work is supported by "Hundred Talents project" of CAS and NSF of China(60473133).

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1003–1012, 2005.

with a mesh of polygons, these methods employ a refinement procedure which operates on this mesh repetitively. With a proper definition of a refinement procedure, the resultant mesh will converge to a smooth surface in the limit. Due to the nature of the surface generation process, such subdivision surfaces do not admit a straightforward closed analytical expression, although exact evaluation of the limit surface is proposed with complicated formulation by Stam [3].

1.2 Irregular Surfaces with Closed form Expressions

For the purpose of facilitating further operations, such as differentiation, a great deal of effort has been expended to the construction of smooth surfaces with a closed analytical definition. Based on the technology of manifolds, Grimm and Hughes [4] proposed a method to construct smooth surface over irregular topology meshes. Following similar ideas, Cotrina and Pla [5], presented a scheme for constructing C^k-surfaces. They firstly apply Catmull-Clark subdivision a number of times to separate the irregular points. And then a manifold M is defined. A control point is associated with each chart of M. Continuous basis functions are generated with support on each chart. A set of transition functions of parameters between different charts are then constructed. Finally, a surface is obtained from a collection of overlapped pieces. The surface is a tensor product B-spline surface around regular areas and is a rational one around irregular points. Although the method provides C^k smoothness, the process to construct surfaces is not easy to follow. A direct construction approach is proposed by Peters [6]. His spline surfaces are constructed by carrying out a subdivision procedure once or twice, and then forming one surface patch around each vertex. The union of all patches represents a C^1 smooth surface. Using triangular patches, Loop's method [7] also follows similar ideas. His method requires only one refinement step and does not need other pre-processing procedures. Taking an irregular mesh as input, his algorithm takes the following three steps to produce a spline surface:

- Carrying out Doo-Sabin subdivision once,
- Constructing quad-nets,
- Generating four quartic triangular patches for each quad-net (lower degree patches may be used for special cases).

As output, the final surface consists of a collection of bi-quadratic Bézier patches except in the neighborhood of the irregular mesh points where triangular patches are employed.

Combining the knowledge of subdivision mesh and S-patch , Loop and DeRose [15] generate a C^1 spline surface. Zheng & Zhang [8] also developed a C^1 spline surface based on subdivision mesh and Zheng-Ball construction [9] [10]. Their construction process is very simple and easy to be implemented. Based on subdivision mesh, Peters [11] recently developed another C^1 spline surface scheme over a Catmull-Clark mesh. The resultant surface is a collection of bi-cubic Bézier patches. Most patches are joined with C^2 continuity and agree with the Catmull-Clark limit surface except in the neighborhood of the irregular mesh points where only tangent continuity is guaranteed.

1.3 The Proposed Approach

The above straight forward methods, despite their ability to generate smooth surfaces of irregular topology, are only able to achieve an overall C^1 smoothness. This is clearly not adequate for many CAD and computer graphics applications. In this paper we present a method to create C^2 smooth spline surfaces. Our method takes as input an irregular mesh and produces as output a spline surface composed of a collection of C^2 joined patches which will be the same as the Catmull-Clark limit surface except in the neighborhood of the irregular mesh points. The construction process takes three steps. The first step is to carry out Catmull-Clark subdivision at most twice over the initial mesh, resulting in a new refined mesh. This new mesh has a simpler structure, since every face has exactly four edges and the irregular vertices are separated properly. A vertex is said to be irregular if its valence [1] is not 4. In the second step, corresponding to each regular face, a bi-cubic Bézier patch is constructed, which coincides with the Catmull-Clark limit surface. A face is said to be regular if none of its vertices are irregular vertices. In the last step, around each irregular vertex, a Gregory patch [12] is constructed. The union of all these surface patches constitutes a C^2 smooth spline surface.

The remaining sections of this paper are organized as follows: Section 2 introduces some preliminary knowledge of C^2 Gregory patches, which will be used as building blocks. In Section 3, the new technique is introduced to construct smooth spline surfaces over a closed irregular mesh. Section 4 extends this technique to open meshes. Section 5 concludes this paper.

2 Preliminaries

In this section, we briefly review the well known C^2 Gregory patches. For more details, the reader is referred to Ref.[12].

Let $q_i : [0,2] \times [0,1] \longrightarrow \Re^2$, $i = 1, \cdots, n$, form a C^2 parametric continuous rectangular patch complex around an n-sided area, as shown in Fig. 1. It is denoted by

$$\partial_{0,j} q_i(s,1) = \frac{\partial^j}{\partial t^j} q_i(s,t)|_{t=1}, \ j = 0,1,2. \tag{1}$$

A C^2 Gregory patch over an n-sided area can be defined by the following two steps:

(1). The domain. The definition domain is given by a regular n-sided area $\mathcal{D} \subset \Re^2$ with unit length edges, as shown in Fig. 2. The distance from a point $\mathbf{u}\ (u_1, u_2) \in \mathcal{D}$ to the i-th boundary E_i is denoted by $d_i(\mathbf{u})$.
(2). The Gregory surface joined by the surrounding patches with C^2 continuity is defined by

$$\mathbf{r}(\mathbf{u}) = \sum_{i=0}^{n-1} w_i(\mathbf{u}) P_i\left(s_i(\mathbf{u}), t_i(\mathbf{u})\right) \tag{2}$$

[1] The valence of a point is the number of its incident edges.

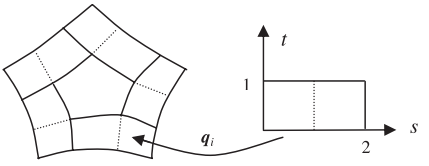


Fig. 1. Patch q_i around an n-sided area

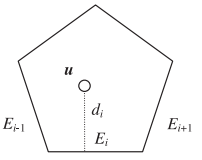


Fig. 2. n-sided domain $\mathcal{D}$. Distant d_i from point u to i-th boundary E_i, $n = 5$

where the weights w_i are given by

$$w_i(\mathbf{u}) = \frac{\prod\limits_{j \neq i, j+1} d_j^3(\mathbf{u})}{\sum\limits_{k=0}^{n-1} \prod\limits_{j \neq k, k+1} d_j^3(\mathbf{u})} \tag{3}$$

and the interpolants are

$$\begin{aligned} \mathbf{P}_i(s,t) = &\sum_{j=0}^{2} \frac{s^j}{j!} \beta^j(t) \partial_{0,j} \mathbf{q}_i(1-t,1) + \sum_{j=0}^{2} \frac{t^j}{j!} \beta^j(s) \partial_{0,j} \mathbf{q}_i(s,1) \\ &- \sum_{j}^{2} \sum_{k=0}^{2} \frac{s^j t^k}{j!k!} \partial_{j,k} \mathbf{q}_i(0,1) \end{aligned} \tag{4}$$

in which

$$\beta(t) = \frac{2t \cos\left(\frac{2\pi}{n}\right) + 1}{2\left(t^4 - 2t^3 + t\right) \cos\left(\frac{2\pi}{n}\right) + 1}$$

$$(s_i(\mathbf{u}), t_i(\mathbf{u})) = \left(\frac{d_{i-1}(\mathbf{u})}{d_{i+1}(\mathbf{u}) + d_{i-1}(\mathbf{u})}, \frac{d_i(\mathbf{u})}{d_i(\mathbf{u}) + d_{i-2}(\mathbf{u})} \right) \tag{5}$$

3 Spline Surface Generation over a Closed Irregular Mesh

The starting point of this method is a user-defined irregular mesh $\mathcal{M}^0$, which is a collection of vertices, edges and faces. Control mesh $\mathcal{M}^0$ may be closed or open. In this section, surface generation over closed meshes is discussed. In the

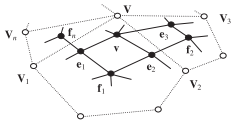


Fig. 3. Applying Catmull-Clark subdivision once around vertex $\mathbf{V}$ whose valence is n

next section, we will deal with open meshes with boundaries. The construction of a spline surface consists of three steps:

(1). Applying Catmull-Clark subdivision twice.
(2). Constructing one bi-cubic Bézier patch for each regular face.
(3). Constructing one Gregory patch around each irregular vertex.

The output is a collection of patches, which are either tensor product bi-cubic Bézier patches or Gregory patches with an overall C^2 continuity.

In Step 2, it is necessary to apply subdivision twice if two irregular points share one face. Twice subdivision can separate irregular points apart to make C^2 construction possible.

3.1 Mesh Subdivision

The purpose of the first step is to sort out the mesh irregularity so that all faces of the mesh have exactly four edges. Given a user-defined irregular mesh $\mathcal{M}^0$, a new refined mesh $\mathcal{M}^1$ can be created by carrying out Catmull-Clark subdivision twice. However, if all its faces are already 4-sided, the user may use the mesh directly. This will create a spline surface closest to the initial mesh $\mathcal{M}^0$.

For a given mesh, three types of points are identified when applying Catmull-Clark subdivision rules. They are the face points $\mathbf{f}$, edge points $\mathbf{e}$ and vertex points $\mathbf{v}$. All these points may be either regular or irregular depending on their valences. Clearly, each vertex $\mathbf{V}$ of valence n of mesh $\mathcal{M}^0$ is incident to n faces and n edges. In the refined mesh $\mathcal{M}^1$, the new vertices associated with vertex $\mathbf{V}$ are computed as follows[1]:

face points: $\mathbf{f}_i$ = averaging of the surrounding vertices of the i-th face incident to vertex $\mathbf{V}$ of mesh $\mathcal{M}^0$, $i = 1, \cdots, n$
edge points: $\mathbf{e}_i = \frac{1}{4}(\mathbf{V} + \mathbf{f}_{i-1} + \mathbf{f}_i + \mathbf{V}_i)$, $i = 1, \cdots, n$
where subscripts are taken modulo the valence n of vertex $\mathbf{V}$, and $\mathbf{V}_i$ is the end point of the i-th edge emanating from $\mathbf{V}$.
vertex point: $\mathbf{v} = \frac{1}{4}\left(2\mathbf{V} + \frac{1}{n}\sum_{i=1}^{n}\mathbf{f}_i + \frac{1}{n}\sum_{i=1}^{n}\mathbf{e}_i\right)$

Note that all faces of the new mesh $\mathcal{M}^1$ are now 4-sided. The valence of the new vertex point $\mathbf{v}$ remains n. The valence of a new edge point is 4. The valence of a new face point is the number of edges of the corresponding face of mesh $\mathcal{M}^0$. Fig. 3 illustrates the subdivision procedure[1].

3.2 Patch Generation over a Regular Area

After subdividing the original mesh $\mathcal{M}^0$ twice, a new mesh $\mathcal{M}^1$ is generated. The second step is to construct one bi-cubic Bézier patch for each regular face of mesh $\mathcal{M}^1$ ensuring all such patches are joined with C^2 continuity.

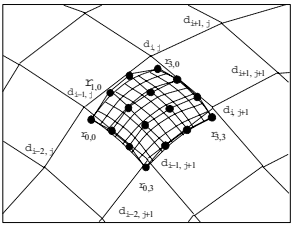


Fig. 4. Bi-cubic Bézier patch corresponding to a regular face F

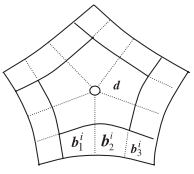


Fig. 5. n-sided patch around an irregular mesh vertex $\mathbf{d}$. $n = 5$

Given a regular face F_i of mesh $\mathcal{M}^1$, suppose its four vertices are $d_{i-1,j}$,$d_{i-1,j+1}$,$d_{i,j}$, and $d_{i,j+1}$, as shown in Fig. 4. The control points $\mathbf{r}_{ij}$ of the corresponding bi-cubic Bézier patch are obtained by the weighted average of the surrounding mesh vertices, which is similar to the 1D conversion given by Böhm[13]. The rules to generate the control points of such a Bézier patch are given in the following:

- Corner control point $\mathbf{r}_{00}$ is a weighted average of the nine mesh vertices surrounding it. The weighting mask is given by $\frac{1}{36}\begin{bmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{bmatrix}$.
- Edge control point $\mathbf{r}_{10}$ is a weighted average of the six mesh vertices surrounding it. The weighting mask is given by $\frac{1}{18}\begin{bmatrix} 1 & 4 & 1 \\ 2 & 8 & 2 \end{bmatrix}$.
- Central control point $\mathbf{r}_{11}$ is a weighted average of the four mesh vertices surrounding it. The weighting mask is $\frac{1}{9}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.
- Other control points are generated similarly by one of the above steps.

3.3 Patch Generation Around an Irregular Vertex

The final step is to construct one Gregory patch around each irregular mesh vertex, which connects the surrounding patches with C^2 continuity.

For an irregular vertex $\mathbf{d}$, the regular faces around it are covered by the bi-cubic Bézier patches generated in the last section, leaving an n-sided area uncovered. Specifically, let the i-th boundary of this n-sided area be formed by three bi-cubic Bézier patches $\mathbf{b}_1^i(s,t)$, $\mathbf{b}_2^i(s,t)$, $\mathbf{b}_3^i(s,t)$, $i = 1, \cdots, n$, as shown in Fig. 5. These bi-cubic Bézier patches $\mathbf{b}_j^i(s,t)$, $i = 1, \cdots, n, j = 1, 2, 3$ are generated by the rules in the last sub-section.

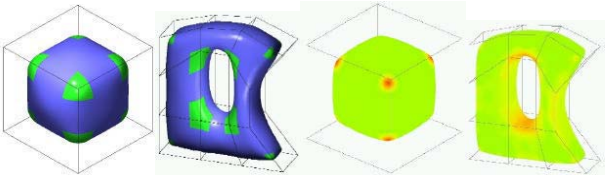


Fig. 6. Models generated from closed meshes

We define

$$\mathbf{q}_i(s,t) = \begin{cases} \mathbf{b}_1^i(\varphi_1(s),t) & s \in [0,\frac{1}{2}] \\ \mathbf{b}_2^i(\varphi_2(s),t) & s \in [\frac{1}{2},1] \\ \mathbf{b}_3^i(s-1,t) & s \in [1,2] \end{cases} \tag{6}$$

where

$$\varphi_1(s) = s + 8s^3 - 8s^4 \tag{7}$$

$$\varphi_2(s) = s + 8(s-1)^3 + 8(s-1)^4 \tag{8}$$

$\varphi_1(s)$ and $\varphi_2(s)$ are so given that $\mathbf{q}_i(s,t) : [0,2]\times[0,1] \rightarrow \mathcal{R}^3$ is C^2 continuous, and all $\mathbf{q}_i(s,t)$, $i = 1,\cdots,n$, form a $C^{2,2}$ patch complex. We then use $\mathbf{q}_i(s,t)$, $i = 1,\cdots,n$, to define a Gregory patch $\mathbf{r}(\mathbf{u})$ in terms of equation (2) to fill the n-sided area. The patch $\mathbf{r}(\mathbf{u})$ is thus joined by the surrounding bi-cubic Bézier patch complex with C^2 continuity.

3.4 Examples

Some models generated from closed meshes are given in Fig. 6. The first two show the shaded models and the corresponding curvature information is plotted on the right. The colors illustrate the distribution of the value $\frac{1}{2}(|\kappa_1| + |\kappa_2|)$, where κ_1 and κ_2 are the principal curvatures.

4 Spline Surface Generation over an Irregular Open Mesh

In this section, we are extending the above rules to cover an open mesh. The main task is to deal with the mesh boundaries.

4.1 Subdivision Rules for Mesh Boundaries

Similar to the treatment of a closed mesh, the process for an open mesh also consists of two steps: subdividing the mesh to make all faces 4-sided; and constructing a surface patch corresponding to each vertex.

Boundary Mesh Subdivision for 2- and 3-Valent Vertices
Given an n-valent vertex $\mathbf{V}$ of the initial mesh $\mathcal{M}_0$ on a boundary, where n=2 or 3, the new vertices of the refined mesh $\mathcal{M}_1$ are generated with the following rules (see Fig. 7).

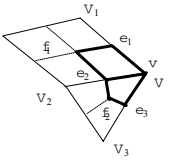


Fig. 7. Subdivision around a boundary vertex $\mathbf{V}$ (n=3)

Face points $\mathbf{f}_i$= average of the surrounding vertices of the i-th face incident to vertex $\mathbf{V}$ of mesh $\mathcal{M}_0$;
edge points $\mathbf{e}_i = \frac{1}{2}(\mathbf{V} + \mathbf{V}_i)$, $i = 1, \cdots, n$;
vertex point $\mathbf{v} = \mathbf{V}$
where $\mathbf{V}_i$ stands for the end point of the i-th edge emanating from vertex $\mathbf{V}$.

Boundary Mesh Subdivision for Vertices of Valence > 3
For a vertex $\mathbf{V}$ of valence $n > 3$, a pre-processing step is carried out which is to insert some extra vertices into the initial mesh. This is to make the valence of all boundary vertices no greater than 3.

For each vertex $\mathbf{V}$ of valence $n > 3$, n new vertices $\mathbf{W}_i$, $i = 1, \cdots, n$, are created

$$\mathbf{W}_i = \frac{1}{2}\mathbf{V} + \frac{1}{2}\mathbf{V}_i, \quad i = 1, \cdots, n. \tag{9}$$

For a convex boundary vertex $\mathbf{V}$, $\mathbf{V}$ and $\mathbf{W}_i$ form an $n+1$-sided face of the mesh $\mathcal{M}^0$, as shown in Fig. 8.

For a concave or flat boundary vertex $\mathbf{V}$, $\mathbf{W}_i$ form an n-sided face of mesh $\mathcal{M}^0$, and $\mathbf{V}$ will be removed from the mesh, as shown in Fig. 9.

After this treatment, all boundary vertices of mesh $\mathcal{M}^0$ have valence of either 2 or 3. Subdivision is then performed twice using the boundary subdivision rules so that all boundary faces are regular.

4.2 Boundary Patches

By subdividing the original mesh twice all boundary faces become regular. One bi-cubic Bézier patch can be generated corresponding to each boundary face in the following.

Suppose a boundary face F has mesh points $\mathbf{d}_{11}$, $\mathbf{d}_{12}$, $\mathbf{d}_{21}$, and $\mathbf{d}_{22}$ as its four vertices. To construct the control points for a bi-cubic Bézier patch with respect to face F, similar to Nasri's method[14], auxiliary mesh points $\mathbf{d}_{11}$, $i = 0$ or $j = 0$ are generated by the following formulas, as shown in Fig. 10.

$$\mathbf{d}_{i0} = 2\mathbf{d}_{i1} - \mathbf{d}_{i2}, \tag{10}$$
$$\mathbf{d}_{0i} = 2\mathbf{d}_{1i} - \mathbf{d}_{2i}. \tag{11}$$

With these auxiliary mesh points, control points for the Bézier patch are generated by the same rules as in Section 3.2.

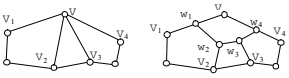


Fig. 8. (Left) Convex boundary vertex $\mathbf{V}$ of valence 4. (Right) New boundary vertices $\mathbf{V}$, $\mathbf{W}_1$ and $\mathbf{W}_4$ of valence 2 or 3

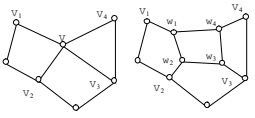


Fig. 9. (Left) Concave inner boundary vertex $\mathbf{V}$ of valence 4. (Right) New boundary vertices $\mathbf{W}_1$ and $\mathbf{W}_4$ of valence 3

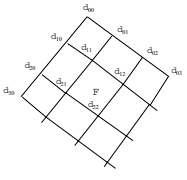


Fig. 10. Boundary face F

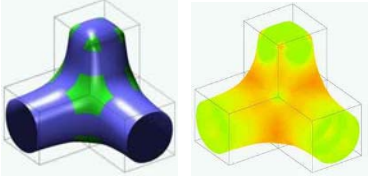


Fig. 11. Model generated from an open mesh

4.3 Example

Fig. 11 shows an example of a model generated from an open mesh. It's boundary patches are constructed by a rule similar to Nasri's method[14].

5 Conclusions

Although several methods have been developed for constructing surfaces over irregular topology meshes with an overall C^1 continuity, achieving an overall C^2 smoothness, however, remains a challenge, especially with simplicity and good efficiency. In this paper, an efficient method has been presented to create a C^2 continuous spline surface over an irregular mesh, which may be either closed or open. The resulting surface is an analytical form of the Catmull-Clark limit surface except in the neighborhood of the irregular mesh points.

The main advantage of this method over the existing methods in the literature is that it ensures an overall C^2 continuity. For an open mesh, the resulting surface possesses zero curvature at its edges which is similar to the surface generated by Nasri[14].

References

1. E Catmull and J Clark. Recursively generated B-spline surfaces on arbitrary topological meshes, *Computer Aided Design* **10** (1978) 350-355.
2. D Doo and M Sabin (1978). Behaviour of recursive division surfaces near extraordinary points, *Computer Aided Design* **10** (1978) 356-360.
3. J. Stam. Exact evaluation of Catmull-Clark subdivision surfaces at arbitrary parameter values, *SIGGRAPH'***98** 395-404.
4. C Grimm and J Hughes. Modeling surfaces of arbitrary topology using manifolds, *SIGGRAPH'***95** 359-368.
5. J Cotrina Navau and N. Pla Garcia. Modeling surfaces from meshes of arbitrary topology, *Computer Aided Geometric Design* **17** (2000) 643-671.
6. J Peters. Constructing C^1 surfaces of arbitrary topology using bi-quadratic and bi-cubic splines, in *Designing Fair Curves and Surfaces*, ed. N Sapidis (SIAM, Philadelphia, 1994) 277-294.
7. C Loop. A G^1 triangular spline surface of arbitrary topological type, *Computer Aided Geometric Design* **11** (1994) 303-330.
8. J J Zheng and J J Zhang. Interactive deformation of irregular surface models, *Lecture Notes in Computer Science* **2330** (2002) 239-248.
9. J J Zheng and A A Ball. Control point surfaces over non-four-sided areas, *Computer Aided Geometric Design* **14** (2001) 807-820.
10. J J Zheng. The n-sided control point surfaces without twist vectors, *Computer Aided Geometric Design* **18** (1997) 129-134.
11. J Peters. Patching Catmull-Clark meshes, *SIGGRAPH* **2000** 255-258.
12. J A Gregory and J M Hahn. A C^2 polygonal surface patch, *Computer Aided Geometric Design* **6** (1989) 69-75.
13. W Böhm. Generating the Bézier points of B-spline curves and surfaces, *Computer Aided Design* **13** (1981) 365-366.
14. A H Nasri. Polyhedral subdivision methods for free-form surfaces, *ACM Transaction on Graphics* **6** (1987), 29-73.
15. C Loop and T DeRose. Generalized B-spline surfaces of arbitrary topology. *SIGGRAPH* **1990**, 347-356.

Constructing Detailed Solid and Smooth Surfaces from Voxel Data for Neurosurgical Simulation

Mayumi Shimizu and Yasuaki Nakamura

Department of Computer and Media Technologies,
Hiroshima City University, 3-4-1, Ozuka-higashi,
Asa-minami-ku, Hiroshima 731-3194, Japan
m_shimizu@toc.cs.hiroshima-cu.ac.jp
nakamura@cs.hiroshima-cu.ac.jp

Abstract. This paper deals with a neurosurgical simulation system with precise volume rendering and smooth tactile sensation. In the system, the Octree based hierarchical representation of volume data with continuous tri-cubic parametric functions, called volumetric implicit functions, and smooth boundary surfaces are introduced to provide detailed solid and smooth tactile sensation in an interactive environment. The volume data represented as voxel data, which are created from CT or MRI images, are divided into sub-volume until volumetric implicit functions can approximate voxel values accurately. An Octree manages the divided volume and parameters of the implicit functions in a hierarchical manner. Furthermore, smooth boundary surfaces are constructed by fitting points on a level surface of the implicit functions. In order to render more detailed solid than voxel precision when objects are zoomed up, sub-sampled voxels are generated by using the implicit functions. As for the tactile sensation, haptic device, PHANToM, is used to actualize a smooth reaction force which is calculated by the surface normal and the distance from a position of an instrument to the nearest surface. Incision with tactile sensation can be executed by making voxels underlying the instrument transparent, when a reaction force is greater than a limit. Several experiments reveal the effectiveness of the proposed methods.

1 Introduction

Medical applications are one of the most important fields of 3D image processing, computer graphics and virtual reality. During past two decades, medical imaging technologies have achieved a rapid progress. 3D images become popular for diagnosing patients and surgical planning. Surgical simulation systems using the 3D images and volume rendering have been proposed [1-10, 12]. Among these systems, a neurosurgical simulation system that supports virtual surgical operation with tactile sensations is desired [7].

Traditionally, surface rendering [3,7-9] is employed to perform volume visualization, because the surface rendering is much faster than the volume rendering.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1013–1022, 2005.

Recently, most of the researches are moving towards volume rendering [5, 6, 10, 11, 13], because real time volume rendering can be performed on an ordinary PC having a special graphic card such as VolumePro1000 [14]. A 3D solid model, especially 3D volume data represented as a collection of voxels, can be constructed from two dimensional images captured by the CT, MRI, Ultrasound, and other imaging methods. A surgical simulation is performed by modifying 3D volume data. Our system must provide the delicate tactile sense when an operation is performed. Haptic devices such as PHANToM [15] require a smooth surface to generate a smooth tactile sense. Several techniques are proposed to construct a surface on the volume data represented by voxels. Marching cubes [16] can construct a high resolution 3D surface on the volume data. However, the constructed surface is a set of a large number of small polygons. Therefore, further processing is required to obtain smooth surfaces.

In our system, volume data represented by voxels are divided into sub volumes such that each sub volume can be approximated by a volumetric implicit function precisely. The Octree is employed to manage the hierarchical division of volume data and the implicit function. Finding the boundary points of volume data, the smooth surfaces are generated by fitting bi-cubic parametric surfaces to the points.

In this paper, a configuration of our simulation system for the neurosurgery is described in section 2. In the following sections, data structures managing volume data, sub-voxel visualization by sub-sampling voxels using volumetric implicit functions, a smooth surface construction for smooth reaction force generation, and surgical simulation are presented.

2 Neurosurgical Simulation System

In the system, a 3D volume model of a head in a virtual space is constructed from CT or MRI 3D images. The Octree representation [17] of the volume data is created, because the spatial operations such as the collision check, distance calculation, and Boolean operations can be efficiently performed by the Octree.

Fig. 1 shows the conceptual configuration of our neurosurgical simulation system. PHANToM [15] controls the position and direction of a virtual surgical instrument displayed in a virtual space with a head model. An intersection of the instrument and the head model is checked whenever the instrument moves. When the instrument contacts the surface of the head model or moves slightly inside the model, the system calculates reaction force PHANToM should present. When an instrument such as a surgical knife intersects with the head model, the voxels under the knife disappear from the virtual space. Practically, the values of those voxels are set to be 0. Then, the Octree is updated to keep the consistency with the voxel model. In the space, a surgeon can touch or see-through the head. Volume rendering board VolumePro1000 [14] generates a new scene by rendering the modified head model.

In surgical planning, since the correspondence of color values and transparency to the voxel values can be changed interactively, surgeons can see through

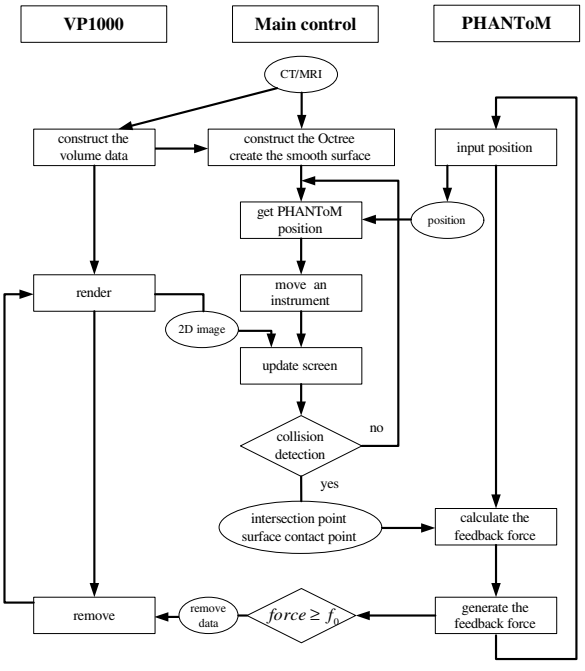


Fig. 1. System configuration

inside a head. Moreover, any cross section of the head model can be displayed in the simulation.

3 Management of Volume Data

The Octree [17] is a hierarchical data structure that can manage the 3D shapes efficiently. Managing the volume data by the Octree, it becomes easy to obtain the distance and direction from an instrument to the nearest point on a surface. The distance and direction are used to calculate a feedback force (reaction force).

The Octree algorithm divides a region into eight sub-regions recursively, until each sub-region becomes uniform. Root of an Octree corresponds to the entire region, and the sub-nodes correspond to sub-regions. The sub-division process is managed by an Octree. In our system, to approximate voxel data and to obtain the smooth boundary surface, a new algorithm using the Octree is developed. In this section, a representation of volume data by tri-cubic parametric function (volumetric implicit function) and our Octree-based algorithm are described.

3.1 Volumetric Implicit Function

Let density value of a voxel at point (x, y, z) be d_{xyz}. In our system, points outside the solid are assigned negative values while points inside the solid have positive val-

ues. Voxel data are given at discrete locations as expressed by equation (1). Equation (2) is a tri-cubic parametric function which approximates voxel data. A three dimensional smooth and continuous solid can be calculated using equation (2).

$$f(x, y, z) = d_{xyz}, \quad x, y, z = 1..N \tag{1}$$

$$\hat{f}(x, y, z) = \sum_{i=0}^{3} \sum_{j=0}^{3} \sum_{k=0}^{3} a_{i,j,k} x^i y^j z^k \tag{2}$$

$\hat{f}(x, y, z)$ is determined such that the square error ϵ in equation (3) between $\hat{f}(x, y, z)$ and $f(x, y, z)$ is minimal.

$$\varepsilon = \sum_{i=0}^{N} \sum_{j=0}^{N} \sum_{k=0}^{N} \left| \hat{f}(x, y, z) - f(x, y, z) \right|^2 \tag{3}$$

3.2 Octree-Based Implicit Volume Representation

The entire volume data are managed and represented in a hierarchical manner by an Octree. Fig. 2 shows an algorithm of our Octree-based volume representation. Firstly, the whole voxel data are approximated such that equation (3) becomes minimal. If ϵ is less than ϵ_0, the sub-division process terminates, and a parameter vector $a_{i,j,k}(i, j, k{=}0,...,3)$ in (2) is assigned to the corresponding leaf node. If ϵ is greater than threshold ϵ_0, the region is divided into eight regions. The voxel values in each sub-divided region are approximated. The same procedure is applied to the rest sub-regions of the Octree until the entire volume data are approximated.

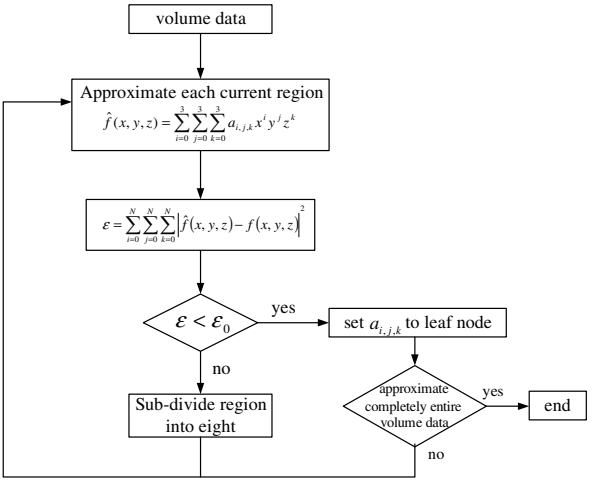


Fig. 2. A hierarchical representation of volume data by an Octree

Surface Fitting. An implicit boundary surface is a level surface represented by equation (4). Boundary surface points satisfying (4) are computed, and a smooth surface is generated by approximating them.

$$\hat{f}(x, y, z) = 0 \tag{4}$$

Mapping the boundary points in a leaf node of an Octree to s-t coordinate, these points can be represented as $Q(s_i, t_j) = (x_{ij}, y_{ij}, z_{ij})$ $(i{=}1,..,n, j{=}1,..,m)$. Bi-cubic parametric surface $\hat{Q}(s_i, t_j) = (x(s_i, t_j), y(s_i, t_j), z(s_i, t_j))$ expressed as equation (5) can be constructed such that the square error ρ in equation (6) becomes minimal. If ρ is greater than a threshold, the region is sub-divided again.

$$\hat{Q}(s,t) = \left(\sum_{k=0}^{3}\sum_{j=0}^{3} a_{k,j} s^k t^j, \sum_{k=0}^{3}\sum_{j=0}^{3} b_{k,j} s^k t^j, \sum_{k=0}^{3}\sum_{j=0}^{3} c_{k,j} s^k t^j \right) \tag{5}$$

$$\rho = \sum_{i=0}^{n}\sum_{j=0}^{m} |\hat{Q}(s_i, t_j) - Q(s_i, t_j)|^2 \tag{6}$$

4 Surgical Operation and Reaction Force Generation

The presentation of the tactile sense is important in the surgical simulation system. Especially, in the interactive simulation of surgery, an actual feeling cannot be obtained without the sense of touch. PHANToM is used for the presentation of a tactile sense. A reaction force is generated when an instrument touches the head model. A huge amount of calculation is required to find the nearest voxel on the surface of an object without any data structures. By managing a 3D model by the Octree, the number of computation can be sharply reduced, and it becomes possible to set up hardness, softness, and friction force by the organization of a portion in an interactive environment.

Surgical operations such as cutting, scraping, and drilling, are realized by updating the voxels under an instrument to be transparent. Surgical operation is performed only when the reaction force to an instrument exceeds a threshold of the instrument, that is, the voxel value corresponding to the cutting portion is set to be transparent. To visualize the voxels, the updated volume must be transferred to VP1000 and rendered. At each transfer cycle, as the minimum bounding volume that encloses the updated voxels is managed by an Octree, only the voxels inside the bounding volume are transferred.

4.1 Presentation of the Tactile Sensation

Fig. 3 illustrates the calculation scheme of the feedback force. PHANToM presents a reaction force from an object. When the cursor of PHANToM collides with a voxel, feedback force F is calculated by equation (7).

$$\boldsymbol{F} = \boldsymbol{F}_v + \boldsymbol{F}_k \tag{7}$$

where $\boldsymbol{F}_v$ and $\boldsymbol{F}_k$ are normal force and friction force, respectively. $\boldsymbol{F}_v$ is calculated by the following spring-damper model. Dynamic friction $\boldsymbol{F}_k$ is proportional to $|\boldsymbol{F}_v|$

$$\boldsymbol{F}_v = -c_v \cdot \boldsymbol{x}_v - c_d \cdot \frac{\partial}{\partial t}\boldsymbol{x}_v \tag{8}$$

$$\boldsymbol{F}_k = c_k \cdot |\boldsymbol{F}_v|, \tag{9}$$

where c_v, c_d and c_k are a spring coefficient, a damper coefficient, and a coefficient of dynamic friction, respectively. The direction of $\boldsymbol{F}_v$ is a surface normal at the nearest point (SCP) on a surface from ACP, and $\boldsymbol{F}_k$ is perpendicular to $\boldsymbol{F}_v$ as shown in Fig. 3. In the case a solid object is represented by voxel data, normal vectors have only 6 directions because voxels are cubic. Therefore, smooth tactile sensation cannot be generated. On the other hand, in the case a solid object is represented by continuous function as equation (2), since our system builds the smooth parametric surface expressed as equation (5), the smooth reaction force can be generated.

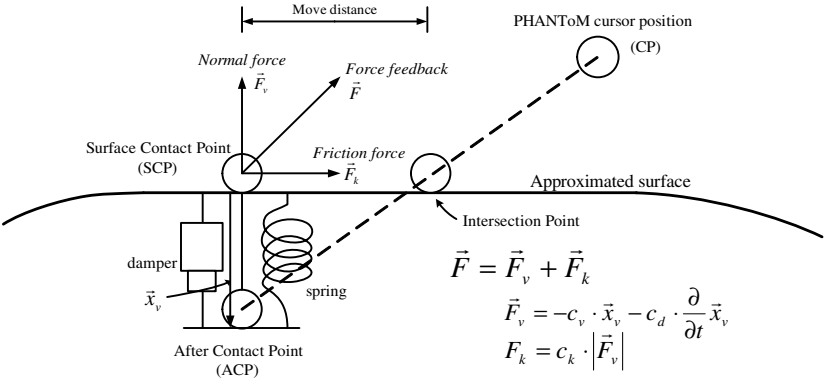


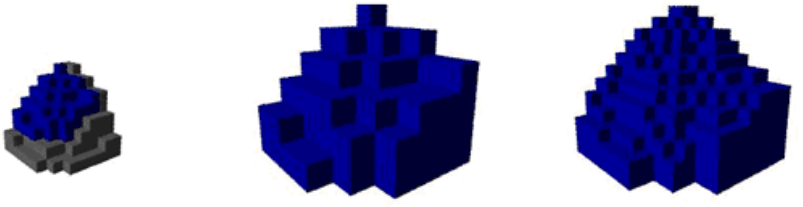
Fig. 3. Calculation of the feedback force

4.2 Smooth Reaction Force Generation

The collision point is an intersection point of boundary surface $\hat{Q}(s,t)$ and the vector from CP to ACP. SCP is determined as the nearest point from ACP to the 3D surface, then the distance from ACP to SCP is calculated. A surface normal vector of the parametric bi-cubic surface is represented as equation (10).

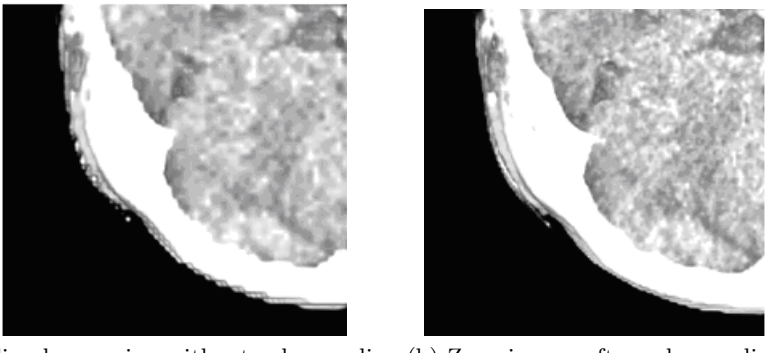
$$normal = \frac{\partial Q}{\partial s} \times \frac{\partial Q}{\partial t} \tag{10}$$

Since smooth surfaces of volume data are built, a smooth tactile sense can be obtained by calculating the feedback force based on equations (7), (8), (9), and (10). As the parameters in equations (8) and (9) are adjusted to represent the hardness, softness, and friction force of the substance (organ) corresponding to a voxel value, the tactile sensation can be changed according to operation instruments, the way of operation and the contact part.



(a) Before zooming up (b) Zooming up with subsampling (c) Zooming up without sub-sampling

Fig. 4. Zooming up images of voxels



(a) Simple zooming without sub-sampling (b) Zooming up after sub-sampling

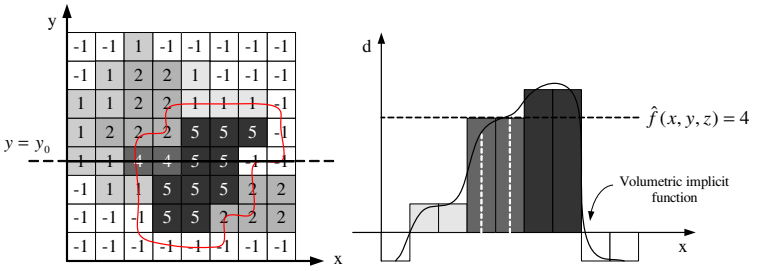
Fig. 5. Effect of sub-voxel construction

4.3 Rendering Detailed Volume Data

When a solid model represented by light blue voxels is zoomed up, as shown in Fig. 4 (a), a jaggy solid in Fig. 4 (b) is displayed, because the voxels themselves are enlarged. To render smoother solid after expansion, the voxels lying on an approximated boundary surface are divided into eight sub-voxels and voxel values of the sub-voxels are calculated by the volumetric implicit function (2). By re-assigning the sub-voxels to the voxels in VolumePro1000, the volume model is reconstructed. Fig. 4 (c) is an example zooming up image after sub-sampling by the reconstruction method. As a result, boundary voxels become smaller and the boundary of the solid can be represented in detail. Fig. 5 (a) and (b) show images without sub-sampling and with sub-sampling of the same part of an actual volume data created from CT images.

4.4 Extracting Organs

Using the volumetric solid function, a certain organ (e.g. a blood vessel, a neural network, or a tumor) can be segmented by the marching cube method that can extract a cluster of voxels with closer values. Fig. 6 illustrates the method of extracting the boundary in two-dimensional case. The marching cube method



(a) Voxel values in two dimension (b) extracting boundary points on level $\hat{f}(x, y, z) = 4$

Fig. 6. Level surface extraction (two dimensional case)

finds voxels lying on an implicit boundary satisfying $\hat{f}(x, y, z) = v$, where v is a given boundary value. Strictly, adjacent voxels which have different signs of $\left(\hat{f}(x, y, z) - v\right)$ are extracted. Voxels inside the boundary are displayed and others are set to be translucent.

4.5 The Procedure of the Simulation

The system always tracks the PHANToM's cursor position and moves the instrument model according to the movement of the cursor. At the same time, the Octree checks whether the instrument intersects with organs. When the instrument moves inside an organ, the system generates a feedback force. The visualization of ongoing surgical operation can be performed by the following procedure.

Simulation Steps.

(1) Select an instrument.
(2) Detect the poison of the cursor of PHANToM.
(3) Move the instrument to the detected position.
(4) Determine whether the instrument moves in an organ by the Octree.
(5) If the instrument is inside an organ,
  (5.1) calculate the distance and direction from the position to SCP.
  (5.2) Feedback the reaction force to PHANToM.
  (5.3) If the reaction force exceeds a threshold, change the voxel value of the position to be transparent.
  (5.4) Update the volume data and render it.
(6) Go to (2)

When some voxels are removed by a surgical operation, the parameters of an approximation function in a leaf node including those voxels are re-calculated. The see-through function can be easily realized by rendering only an organ region. That is, when a surgeon wants to see only a tumor, only the tumor region is extracted by the method described in 4.4 and other voxels are made transparent. Rendering the whole voxels, only the tumor region is displayed clearly.

5 Experimental Results and Discussion

A 3D head model created by volume rendering is shown in Fig. 7 (a). By changing the transparency corresponding to the voxel values, not only the surface of the head but also inside of the head can be visualized. Fig. 7 (b) and (c) show the images before and after performing incision operation, respectively. When incision operation is performed, the voxels with which an instrument touched are made transparent. In this study, by constructing smooth surfaces on the boundary of the 3D model by the volumetric implicit functions, the tactile sense similar to an actually colliding with an object can be realized. The Octree is employed to check the interference and intersection with objects, and to find the nearest voxel from the cursor position to the surfaces of the objects.

It takes about 15 minutes for 512x512x512 voxel data to construct an Octree structure managing the volumetric functions and objects regions. However, as the construction of an Octree can be performed as a pre-processing before starting the simulation, this does not matter for interactive simulation. As for the rendering time, it takes approximately 43.2 seconds to render 300 frames; the rendering speed is about 7 frames per second. Since it is well known that at least 10 frames per second are needed in order to realize the interactive operation, the further improvement of algorithms for rendering and overlaying images generated by VP1000 is necessary. Moreover, as the construction of a detailed voxel data can be carried out in a few seconds, an algorithm must be improved.

(a) A head model (b) Before incision (c) After incision

Fig. 7. Volume rendering and a simulation result

6 Conclusion

We have developed a neurosurgical simulation system on an ordinary PC with VolumePro1000 and PHANToM. VolumePro1000 was very useful to visualize the internal structure of organ in detail and the deformation of the three dimensional model interactively. Managing the volume data and volumetric implicit functions by the Octree, efficient search could be carried out and the computational cost of collision detection was reduced. Furthermore, sub-sampling the voxel values using the implicit functions, more detailed solid than voxels can be

displayed. Tactile sense was produced by PHANToM. In the case of computing the feedback force on the voxel, the tactile sense has unevenness because voxel data are angular. To obtain the smooth reaction force, the smooth surface was created from a volumetric implicit function approximating voxel data. As a result, unevenness was eliminated and the smooth tactile sense can be obtained.

References

1. Quistgarrd, J. U.: Signal Acquisition and Processing in Medical Diagnostic Ultrasound, IEEE Signal Processing Magazine, 1053-5888, (1997) 67-74
2. Levoy, M.: Display of Surface from Volume Data, IEEE Computer Graphics & Applications (1988) 29-37
3. Robinson N.: Biomedical Virtual Environments, Interacting with Virtual Environments, John Wikey & Sons (1994) 119-132
4. Borgefors, G.: Notes on the Multiscale Representation of 2D and 3D shapes, Graphics Models and Image Processing (1999) 44-62
5. Quingson, Z., Keong, K. C., Sing N. W.: Interactive Surgical Planning Using Context Based Volume Visualization Techniques, Proc. MIAR 2001 (2001) 21-25,
6. Galyean, T. A., Hughes, J. F.: Sculpting: A Interactive Volumetric Modeling Technique, Computer Graphics, Vol. 25, No. 4, (1991) 267-274
7. Kockro, R. A., et al.: Planning and Simulation of Neurosurgery in a Virtual Reality Environment, Neurosurgery, Vol. 46, No. 1, (2000) 118-137
8. Tiede, U., et al.: Investigation of Medical 3D Rendering Algorithms, IEEE Computer Graphics & Applications, (1999) 41-53
9. Shi, J.Y., and Yan, L.X.: Deformation and Cutting in Virtual Surgery, Proc. MIAR 2001, (2001) 95-102
10. Shimizu, M., Nakamura, Y. : Virtual Surgical Simulation with Tactile Sensation for Neurosurgery Operation by Volume Graphics, Proc. CCCT 2004, Vol. 1, (2004) 112-117
11. Hua, J., Qin, H. : Haptics-Based Dynamic Implicit Solid Modeling, IEEE trans. Visualization and Computer Graphics, Vol. 10, No. 5,(2004) 574-586 .
12. Halle, M.: Multiple Viewpoint Rendering, Proc. SIGGRAPH 98,(1998) 243-254
13. Kreeger, K. A., Kaufman, A. E. : Mixing Translucent Polygons with Volumes, Proc. Visualization 99, (1999) 191-198
14. VolumePro 1000 Principles of Operation, TERARECON, INC. , 2001.
15. GOHST SDK Programmer's Guide, SensAble Technologies, Inc. , 2000.
16. Lorensen, W. E., Cline, H. E.: Marching Cubes: A High Resolution 3D Surface Construction Algorithm, SIGGRAPH 87, Vol. 21, No. 4, (1987) 163-169
17. Samet, H. : The Design and Analysis of Spatial Data Structures, Addison-wesley.(1990).

Curvature Estimation of Point-Sampled Surfaces and Its Applications

Yongwei Miao[1,2,⋆], Jieqing Feng[1], and Qunsheng Peng[1]

[1] State Key Lab. of CAD&CG, Zhejiang University, Hangzhou 310027, P.R.China
{miaoyw, jqfeng, peng}@cad.zju.edu.cn
[2] College of Science, Zhejiang University of Technology, Hangzhou 310032, P.R.China

Abstract. In this paper, we propose a new approach to estimate curvature information of point-sampled surfaces. We estimate curvatures in terms of the extremal points of a one-dimensional energy function for discrete surfels (points equipped with normals) and a multi-dimensional energy function for discrete unstructured point clouds. Experimental results indicate that our approaches can estimate curvatures faithfully, and reflect the subtle curvature variations. Some applications for curvature information, such as surface simplification and feature extraction for point-sampled surfaces, are given.

1 Introduction

With the improvements in 3D capture technology for the surfaces of real objects and improvements in graphics hardware to handle large numbers of primitives, point-sampled surfaces have emerged as a versatile representation for geometric models in computer graphics, and become the focus of analysis and digital geometry processing techniques [1, 2, 3, 4, 6, 7, 8]. Point-sampled surfaces are used for interpolation, free-form deformation, boolean operation, spectral filtering, surface painting, shading and so on [16, 18, 21].

The point-sampled surface representations can fall into two major categories: implicit and explicit representations. Implicit representations such as MLS surface [1, 5, 13] and radial basis functions [8] have a particularly simple algebraic structure. The MLS procedure projects points close to an anticipated surface approximation onto this surface and the set of fix points of the projection is conjectured to be a MLS surface. This definition is useful but it does not give much insight into the properties of the surface. Explicit representation such as extremal surfaces [6, 7] is defined based on an energy function and a normal vector field. They described a simple iterative procedure for taking a point in the neighborhood of the point cloud onto the surface, and implemented a point set surface for surfels, which input points equipped with normals.

⋆ This is a paper for Computer Graphics and Geometric Modeling(TSCG'2005). Correspondence Author E-mail: mywfjy@hzcnc.com

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1023–1032, 2005.

A point-sampled surface can be considered as a discrete sampling of a continuous surface, resulting in discrete unorganized point clouds equipped with normals or not. It is a lack of local surface differentials such as normal and curvature for every sample point, which leads to some difficulties not only in the context of digital geometry processing applications such as surface modeling and editing, but also in the context of rendering applications such as illumination and non-photorealistic rendering. Normal information for sample point is very important in order to generate realistic images of point-sampled surfaces [2, 12]. On the other hand, curvature information plays an important role in surface modeling and editing [10, 15, 17, 19]. In Pauly et al. [17], they indicated the potential of surface curvature characteristics on controlling the local sampling density and proposed efficient surface simplification methods. In Gumhold et al. [10] and Pauly et al. [19], they presented feature lines extraction method based on curvature information of point-sampled surfaces.

1.1 Related Work

The problems of normal and curvature estimations have been studied by various communities such as computer graphics, image processing, and applied mathematics, but mostly for the cases of 2D manifold [9] and mesh representations of the surfaces [14, 20].

For point-sampled geometry, many researchers adopted a statistical approach — principal component analysis (PCA) for local neighborhoods to estimate normals and local surface properties, such as curvatures. This method is also called as covariance method [10, 11, 17, 19]. Recently, Alexa and Adamson [5] use the gradient of the implicit function of Levin's MLS surface to compute surface normal faithfully, and present efficient orthogonal projection operators for sampling theory. But neither of the approaches can estimate curvatures efficiently and faithfully.

1.2 Our Work

Based on the theory of extremal surfaces, we introduce some energy functionals to evaluate the best fitting sphere to given neighbour points. It is a one-dimensional function for the case of discrete surfels and a multi-dimensional function for the case of discrete point clouds. Then, we propose a new approach to estimate curvatures and normals of point-sampled surfaces faithfully. They can reflect subtle surface curvature variations. We also explore the intrinsic property of the osculating sphere center. Finally, we apply estimated curvatures and normals of point-sampled surfaces to surface simplification and feature extraction.

1.3 Paper Overview

The paper is partitioned into four sections. Section 2 describes the curvature estimation method for surfels, i.e. points equipped with normals. In Section 3, based on sphere fitting, we describe curvature and normal estimation approaches

for discrete point clouds, in the absence of normal information. Applications of estimated differentials are presented in Section 4 before we end with a brief discussion and conclusion in section 5.

2 Curvature Estimation Based on Surfels

We assume that normals are available when converting a mesh or implicit surface to discrete point clouds. Then, we can define a point sampled geometry which takes a set of surfels rather than discrete point clouds as input. The input surfels can be represented as clouds of point-direction pairs $\{(\boldsymbol{p}_i, \boldsymbol{n}_i)\}$.

For each sample point, we employ a sphere with proper size to fit its neighbour points, which centeriod at $\boldsymbol{p}_i^0 = \boldsymbol{p}_i - r \cdot \boldsymbol{n}_i$ and radius r. See Figure 1. For the sake of measuring the quality of sphere fitting to local neighbour points, we introduce the following energy function as in Amenta and Kil [6, 7]:

$$e(r) = \sum_{\boldsymbol{p}_j \in N_i} (d(\boldsymbol{p}_j, \boldsymbol{p}_i^0) - r)^2 \theta(\boldsymbol{p}_j, \boldsymbol{p}_i)$$

where $d(\cdot, \cdot)$ means the Euclidean distance, θ denotes a Gaussian weighting or a normalized Gaussian weighting function.

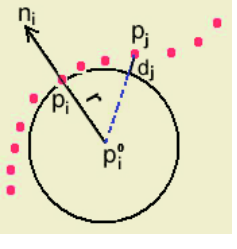


Fig. 1. *Osculating Sphere Fitting*

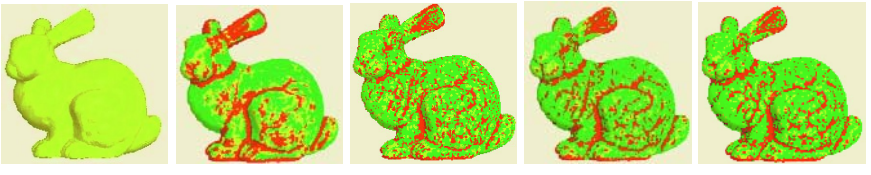
The energy function, which define over $\boldsymbol{R}^+$, occur at a one-dimensional set of r, each corresponding to a fitting sphere. So, in order to best fitting the neighbour points by the means of sphere, we should find the radius r^* for local minimal of the energy function. If we note $arglocalmin_r$ as the set of radius r^* corresponding to local minima of a function of variable r. The radius r^* of best fitting osculating sphere can be represented as:

$$r^* = \underset{r \in \boldsymbol{R}^+}{arglocalmin}\, e(r)$$

We can use an implementation of Brent's method for this one dimensional non-linear optimization.

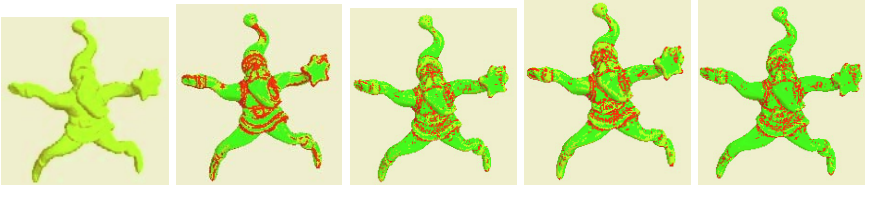
After finding the best fitting sphere, we can estimate curvature K at sampling point $\boldsymbol{p}_i$ as reciprocal of the best radius r^*, such as $K(\boldsymbol{p}_i) = \frac{1}{r^*}$.

Practically, the estimation of local surface differentials is affected significantly by the local neighborhood sizes of sample points. Small neighborhood provides better estimation for clean point cloud, but the estimation is sensitive to noise. With the increase of neighborhood size, the estimation will be less sensitive to noise, however, smoothing problem will occur. A trade-off strategy is to use adaptive neighborhood size for each sample point. It makes the local sampling density $\rho = K/r^2$ constant, where r is the radius of the enclosing sphere of K-nearest neighbors of sample point. Using different neighborhood sizes, we give some curvature estimation examples for point-sampled surfaces and compare them with results by covariance method, which are shown in figure 2 and figure 3. Figure 2(e) and Figure 3(e) show the corresponding curvature estimation results by adaptive neighborhood sizes. For bunny and santa models, our approach can estimate curvatures faithfully, and reflect the subtle curvature variations.



(a) (b) (c) (d) (e)

Fig. 2. *Comparison of curvature estimations on the bunny model. In (b), curvature information is estimated by covariance method. (c) and (d) show curvature estimation by our minimal energy method for different neighborhood sizes σ_{16} and σ_{30}, separately. (e) shows curvature estimation by our minimal energy method for adaptive neighborhood sizes*



(a) (b) (c) (d) (e)

Fig. 3. *Comparison of curvature estimations on the santa model. In (b), curvature information is estimated by covariance method. (c) and (d) show curvature estimation by our minimal energy method for different neighborhood sizes σ_{16} and σ_{30}, separately. (e) shows curvature estimation by our minimal energy method for adaptive neighborhood sizes*

3 Curvature Estimation for Point Clouds

3.1 Curvature Estimate

Only discrete unstructured point clouds without normal information are obtained in general when sampled from a smooth, two-manifold surface S of a 3D object. Our method adopts a proper sphere to fit neighbour points. The sphere depends on normal $\boldsymbol{n}$ and its radius r. Thus, measurement for fitting quality can be represented as the following energy function:

$$e(\boldsymbol{n}, r) = \sum_{\boldsymbol{p}_j \in N_i} (d(\boldsymbol{p}_j, \boldsymbol{p}_i - r \cdot \boldsymbol{n}) - r)^2 \theta(\boldsymbol{p}_j, \boldsymbol{p}_i)$$

The local minimal of the energy function, over $\boldsymbol{S}^2 \times \boldsymbol{R}^+$ ($\boldsymbol{S}^2$ is the space of directions, the ordinary two-sphere), occur at a discrete set of inputs $(\boldsymbol{n}, r)$, each corresponding to a fitting sphere center $\boldsymbol{p} = \boldsymbol{p}_i - r \cdot \boldsymbol{n}$.

To simplify investigation for osculating sphere fitting, we rewrite the energy function. We adopt $\boldsymbol{p}$ as a unique variable of energy function, which gives as:

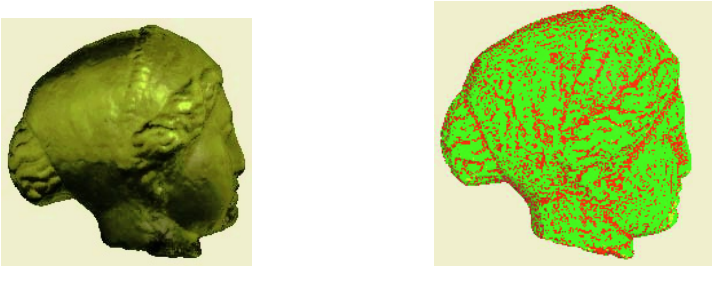
$$e(\boldsymbol{p}) = \sum_{\boldsymbol{p}_j \in N_i} (d(\boldsymbol{p}_j, \boldsymbol{p}) - d(\boldsymbol{p}_i, \boldsymbol{p}))^2 \theta(\boldsymbol{p}_j, \boldsymbol{p}_i)$$

where $d(\boldsymbol{p}_i, \boldsymbol{p})$ is the radius of fitting sphere.

The local minimal of new energy function $e(\boldsymbol{p})$ defines over $\boldsymbol{R}^3$, so we can use Powell's method for multi-dimensional non-linear optimization. After we find the best fitting sphere centroid $\boldsymbol{p}^*$, we can easily estimate the normal $\boldsymbol{n}_i$ and radius r_i as following:

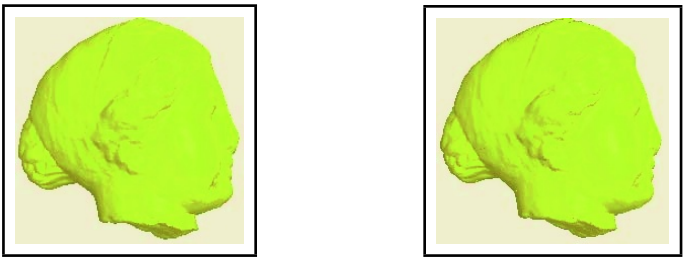
$$r_i = d(\boldsymbol{p}_i, \boldsymbol{p}^*), \boldsymbol{n}_i = \frac{\boldsymbol{p}_i - \boldsymbol{p}^*}{r_i}$$

So, the curvature $K(\boldsymbol{p}_i)$ at point $\boldsymbol{p}_i$ can estimate as the reciprocal of radius, i.e. $K(\boldsymbol{p}_i) = \frac{1}{r_i}$.



(a) Venus model (b) curvature estimation for adaptive neighbor sizes

Fig. 4. *Curvature estimation on the point cloud venus model (a). In (b), curvature information is estimated by our multi-dimensional minimal energy method with adaptive neighborhood sizes*



(a) original normal for Venus (b) estimated normal for Venus

Fig. 5. *Comparison of normal estimation for the venus model. In (a), rendering model with original normal information. (b) rendering model with estimated normal information by our multi-dimensional minimal energy method*

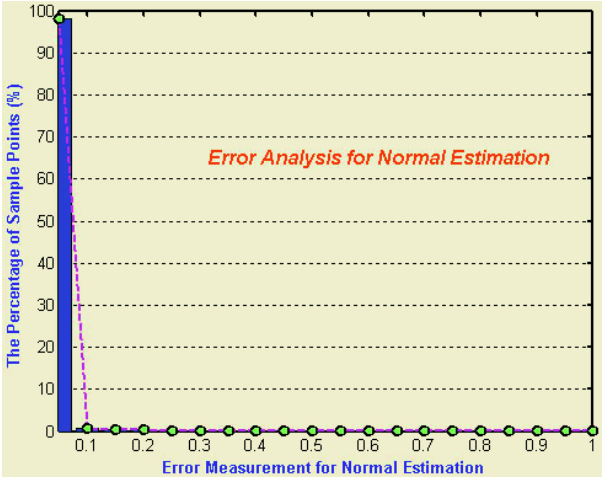


Fig. 6. *Error analysis for normal estimation*

Figure 4 shows the curvature estimation result for the point cloud venus model by our multi-dimensional minimal energy approach with adaptive neighborhood sizes.

Figure 5 shows the normal estimation result by our multi-dimensional minimal energy approach. We compare our estimated normal result with original normal information of the model. In order to measure the accuracy of our method, we use angle between estimated normal and original normal as a measurement of normal error, i.e. $Error = 1.0 - \langle \boldsymbol{n}_{estimated}, \boldsymbol{n}_{original} \rangle$, where $\langle , \rangle$ means vector inner product. Experimental result (see figure 6) indicates that the normal errors for more than 98% of sample points are less than 0.05, which confirms effectiveness of our algorithm for normal estimation.

3.2 The Intrinsic Property of Osculating Sphere Center

The fitting sphere center $\boldsymbol{p}^*$ is a local minimal of multi-dimensional energy function $e(\boldsymbol{p})$, i.e.,

$$e(\boldsymbol{p}^*) = \underset{\boldsymbol{p}\in\boldsymbol{R}^3}{arglocalmin}\, e(\boldsymbol{p})$$

According to the extreme property for multi-dimensional function, we have the directional derivative of $e(\boldsymbol{p})$ is zero, that is:

$$\boldsymbol{n}\cdot\nabla_{\boldsymbol{p}} e(\boldsymbol{p}) = 0$$

We examine the gradient of $e(\boldsymbol{p})$ in the cartesian-normal system $\{\boldsymbol{e}_k\}_{k=1,2,3}$:

$$\nabla_{\boldsymbol{p}} e(\boldsymbol{p}) = \Big(\frac{\partial e(\boldsymbol{p})}{\partial \boldsymbol{e}_1}, \frac{\partial e(\boldsymbol{p})}{\partial \boldsymbol{e}_2}, \frac{\partial e(\boldsymbol{p})}{\partial \boldsymbol{e}_3}\Big)$$

The product rule for differentiating vector fields yields the gradient of $e(\boldsymbol{p})$:

$$\frac{\partial e(\boldsymbol{p})}{\partial \boldsymbol{e}_k} = 2\sum_{\boldsymbol{p}_j\in N_i}\Big(\frac{\boldsymbol{e}_k^T(\boldsymbol{p}-\boldsymbol{p}_j)}{\|\boldsymbol{p}-\boldsymbol{p}_j\|} - \frac{\boldsymbol{e}_k^T(\boldsymbol{p}-\boldsymbol{p}_i)}{\|\boldsymbol{p}-\boldsymbol{p}_i\|}\Big)\big(\|\boldsymbol{p}-\boldsymbol{p}_j\| - \|\boldsymbol{p}-\boldsymbol{p}_i\|\big)\theta(\boldsymbol{p}_j,\boldsymbol{p}_i)$$

where $\|\boldsymbol{p}-\boldsymbol{p}_i\|$ means the Euclidean distance between sample points $\boldsymbol{p}$ and $\boldsymbol{p}_i$.

Then, we get the directional derivative of $e(\boldsymbol{p})$ as:

$$\boldsymbol{n}\cdot\nabla_{\boldsymbol{p}} e(\boldsymbol{p}) = 2\sum_k \frac{\boldsymbol{e}_k^T(\boldsymbol{p}-\boldsymbol{p}_i)}{\|\boldsymbol{p}-\boldsymbol{p}_i\|}\sum_{\boldsymbol{p}_j\in N_i}\Big(\frac{\boldsymbol{e}_k^T(\boldsymbol{p}-\boldsymbol{p}_j)}{\|\boldsymbol{p}-\boldsymbol{p}_j\|} - \frac{\boldsymbol{e}_k^T(\boldsymbol{p}-\boldsymbol{p}_i)}{\|\boldsymbol{p}-\boldsymbol{p}_i\|}\Big)\big(\|\boldsymbol{p}-\boldsymbol{p}_j\| - \|\boldsymbol{p}-\boldsymbol{p}_i\|\big)\theta(\boldsymbol{p}_j,\boldsymbol{p}_i)$$

So, the extreme property becomes:

$$\sum_{\boldsymbol{p}_j\in N_i}\Big(\frac{\boldsymbol{p}-\boldsymbol{p}_i}{\|\boldsymbol{p}-\boldsymbol{p}_i\|}\cdot\frac{\boldsymbol{p}-\boldsymbol{p}_j}{\|\boldsymbol{p}-\boldsymbol{p}_j\|} - 1\Big)\big(\|\boldsymbol{p}-\boldsymbol{p}_j\| - \|\boldsymbol{p}-\boldsymbol{p}_i\|\big)\theta(\boldsymbol{p}_j,\boldsymbol{p}_i) = 0$$

And the osculating sphere center $\boldsymbol{p}^*$ should satisfy the above intrinsic property.

4 Applications

4.1 Surface Simplification

For surface data acquisition, modern 3D scanning devices are capable of producing point clouds that contain over millions of sample points. These sample points is often converted into a continuous surface representation for further processing. Many of these conversion algorithms are computational expensive and require substantial amounts of main memory. Reducing the complexity of such data sets is one of the key preprocessing steps for subsequent applications, such as point-based modeling and rendering, multi-resolution modeling and visualization, etc.

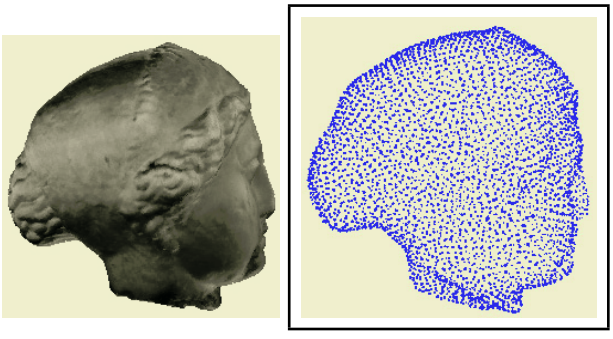
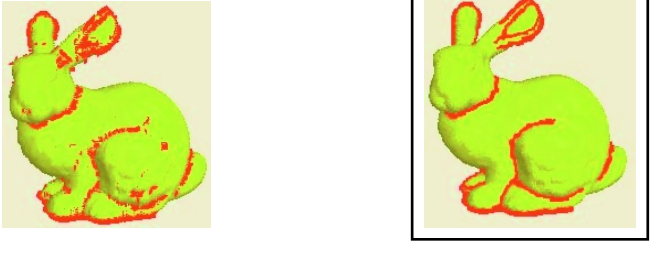


Fig. 7. *Hierarchical clustering simplification for venus model from 134,345 (left) to 3,914 sample points (right)*



(a) minimal spanning graph (b) short branches pruned

Fig. 8. *Feature extraction and reconstruction for bunny model*

We argue that effective surface simplification can be performed directly on the point cloud, similar to other point-based processing and visualization applications. We adopt the hierarchical clustering to simplify input point sampled geometry. It recursively splits the point clouds into the set of clusters using a binary space partition.

The point cloud is split if:

- the size is larger than the user specified maximum cluster size or
- the curvature variation is above a given threshold.

The curvature variation information can be obtained by the methods in Section 3. If the splitting criterion is not fulfilled, the point cloud becomes a cluster. Then, we can build a binary tree, and each leaf node corresponds to a cluster. Each cluster is replaced by a representative sample, typically its centroid, we can create a simplified point cloud. The hierarchical clustering simplification result for venus model is shown in figure 7.

4.2 Feature Extraction

The feature extraction problem is closely related to surface parameterization reconstruction, which has important applications in point-based modeling, computer vision, medical imaging, and laser range scanning. The feature extraction and reconstruction are useful preprocessing steps for subsequent point-based processing and surface parameterization reconstruction. Then the subsequent problems can be divided into several simpler sub-problems on point clouds.

In feature extraction, we first construct a neighbor graph on the point clouds, its edge weight reflects the curvature for two sample points and sampling density. We define the penalty functions at every feature vertices, whose curvature variation is above a user specified maximal threshold, and the penalty weights for the edges of the neighbor graph are computed. After constructing the neighbour graph, we can find a minimal spanning graph (MSG) that minimize the feature penalty weights(see Figure 8(a)). A pruning algorithm should be executed for cutting off short branches in the minimal spanning graph(see Figure 8(b)). The feature extraction and reconstruction results for bunny model is shown in Figure 8.

5 Conclusion

For point-sampled surfaces determined by a set of surfels, we use the extremal point of one-dimensional energy function to estimate curvature information. However, for point-sampled surfaces defined by discrete point clouds, we use the best fitting sphere center, which is the extremal point of multi-dimensional energy function, to estimate curvature and normal information. We also apply the curvature and normal information for two important applications: surface simplification and feature extraction for point-sampled geometry. Experimental results indicate that our approaches can estimate curvatures faithfully, and reflect the subtle curvature variations.

Future research should be focused on two aspects: the error should be analyzed for curvature and normal estimation for noise and local sampling density cases, another one is to use our method for digital geometry processing of point-sampled surfaces.

Acknowledgement

This work is partially supported by Chinese National 973 Fundamental Science Programs (973 program) (2002CB312101) and National Natural Science Foundation of China (NSFC) (60373036).

References

1. Adamson, A., and Alexa, M.: Approximating and intersecting surfaces from points. Eurographics Symposium on Geometry Processing (2003) 245-254

2. Adamson, A., and Alexa, M.: Ray tracing point set surfaces. Proceedings of Shape Modeling International (2003) 272-282
3. Alexa, M., Behr, J., Cohen-or, D., Fleishman, S., Levin, D., Silva, C.T.: Point set surfaces. IEEE Visualization (2001) 21-28
4. Alexa, M., Behr, J., Cohen-or, D., Fleishman, S., Levin, D., Silva, C.T.: Computing and rendering point set surfaces. IEEE Transactions on Visualization and Computer Graphics 9(1) (2003) 3-15
5. Alexa, M., and Adamson, A.: On normals and projection operators for surfaces defined by point sets. Symposium on Point-Based Graphics (2004) 149-155
6. Amenta, N., and Kil, Y.J.: Defining point-set surfaces. ACM Transactions on Graphics 23(3) (2004) 264-270
7. Amenta, N., and Kil, Y.J.: The Domain of a Point Set Surface. Symposium on Point-Based Graphics (2004) 139-147
8. Carr, J., Beatson, R., Cherrie, J., Mitchell, T., Fright, W., Mccallum, B., Evans, T.: Reconstruction and representation of 3d objects with radial basis functions. ACM SIGGRAPH'01 (2001) 67-76
9. do Carmo, M.P.: Differential Geometry of Curves and Surfaces. Prentice-Hall International, Englewood Cliffs, NJ (1976)
10. Gumhold, S., Wang, X., McLeod, R.: Feature Extraction from Point Clouds. Proceedings of 10th International Meshing Roundtable (2001) 293-305
11. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface reconstruction from unorganized points. ACM SIGGRAPH'92 (1992) 71-78
12. Kalaiah, A., Varshney, A.: Differential point rendering. Proceedings of the 12th Eurographics Workshop on Rendering Techniques (2001) 139-150
13. Levin, D.: Mesh-independent surface interpolation. In Geometric Modeling for Scientific Visualization, G. Brunnett, B. Hamann, K. Mueller, and L. Linsen, Eds. Springer-Verlag (2003) 37-49
14. Meyer, M., Desbrun, M., Schröder, M., Barr, A. H.: Discrete differential-geometry operators for triangulated 2-manifolds. Proceedings of VisMath, Berlin (2002)
15. Moenning, C., Dodgson, N.A.: A new point cloud simplification algorithm. Proceedings of 3rd International Conference on Visualization, Imaging and Image Processing (2003) 1027-1033
16. Pauly, M., Gross, M.: Spectral Processing of Point-Sampled Geometry. ACM SIGGRAPH'01 (2001) 379-386
17. Pauly, M., Gross, M., Kobbelt, L.: Efficient Simplification of point-sampled surfaces. IEEE Visualization'02 (2002)
18. Pauly, M., Keiser, R., Kobbelt, L., Gross, M.: Shape modeling with point-sampled geometry. ACM Transactions on Graphics 22(3) (2003) 641-650
19. Pauly, M., Keiser, R., Gross, M.: Multi-scale Feature Extraction on Point-sampled Surfaces. Computer Graphics Forum 22(3) (2003) 281-290
20. Taubin, G.: Estimating the tensor of curvature of a surface from a polyhedral approximation. Proceedings of the Fifth International Conference on Computer Vision (1995) 902-907
21. Zwicker, M., Pauly, M., Knoll, O., Gross, M.: Pointshop 3d: An interactive system for point-based surface editing. ACM Transactions on Graphics 21(3) (2002) 322-329

The Delaunay Triangulation by Grid Subdivision

Si Hyung Park[1], Seoung Soo Lee[2], and Jong Hwa Kim[2]

[1] Mechanical Design and Production Engineering, Konkuk University, 1 Hwayang dong, Gwangjin-gu, Seoul, 143-701, Korea
parksh@konkuk.ac.kr
[2] CAESIT, Konkuk University, 1 Hwayang dong, Gwangjin-gu, Seoul, 143-701, Korea
{sslee, jhkim}@konkuk.ac.kr

Abstract. This study presents an efficient algorithm of Delaunay triangulation by grid subdivision. The proposed algorithm show a superior performance in terms of execution time to the incremental algorithm and uniform grid method mainly due to the efficient way of searching a mate. In the proposed algorithm, uniform grids are divided into sub-grids depending on the density of points and areas with high chance of finding a mate is explored first. Most of the previous researches have focused on theoretical aspects of the triangulation, but this study presents empirical results of computer implementation in 2-dimension and 3-dimension, respectively.

1 Introduction

Delaunay triangulation has been widely applied in a variety of areas as the computer technologies advance. The area of application includes computational geometry, metallurgy, GIS, virtual reality, computer visualization, reverse engineering, FEA, solid modeling, and volume rendering. The execution time of most two-dimensional Delaunay triangulation algorithm is nonlinear to the number of data points. Moreover, 3-dimensional Delaunay triangulation requires more execution time than 2-dimensional algorithm even the number of data points are equal. Therefore, if an inefficient algorithm is used, the execution time of 3-dimensional Delaunay triangulation increases exponentially in proportion to the number of data points, which implies that it is necessary to develop an efficient algorithm even if high-performance computers are available. One major drawback of the existing Delaunay triangulation using uniform grid is that the execution time increases if the data points are not uniformly distributed. Thus, an algorithm is proposed in this study to improve the execution time of Delaunay triangulation when the points are not uniformly distributed.

The proposed algorithm adopts two strategies: first, after applying a uniform grid, a grid with large number of points is divided into smaller grids to reduce the variance in the number of points in a grid. Second, the mate is selected among the most probable region by investigating the density of the search area.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1033–1042, 2005.

2 Types Delaunay Triangulation Algorithm

2.1 Incremental Insertion Algorithm

In this method[3-4] the points in the set are selected one at a time to form a triangle. Starting with a simplex which contains the convex hull of the point set, the points are inserted into the point set, P one at a time. The simplex containing the currently added point is partitioned by inserting it as a new vertex. The circumsphere criterion is tested on all the simplices adjacent to the new ones, recursively, and their faces are flipped if necessary.

2.2 Incremental Construction Algorithm[5-6]

The Delaunay triangulation is constructed by successively building simplices whose circum-hyperspheres contain no points in point set P.

2.3 Plain Sweep Algorithm

The plain sweep algorithm constructs a cell complex in the plane by sweeping the plane with a 'sweepline'.

2.4 Divide and Conquer Algorithm

This algorithm is based on the recursive partition and local triangulation of the point set, and merging phase where the resulting triangulations are joined. The major problem of this algorithm is how to design the merging phase. The merging phase is fairly simple in E2[7], but it is hard to design in Ed[8].

3 Proposed Delaunay Triangulation

3.1 Proposed Strategies

The algorithm proposed in this study is based on the Delaunay triangulation using uniform grid[9] developed by Fang and Piegl[1-2] However, it is different in two aspects.

In most uniform grid methods, the efficiency of Delaunay triangulation degrades if the variance of the number of points in the grid is high. To reduce the variance, the proposed algorithm divide the grids with more than four points into 8 sub-grids in 3 dimensional space and 4 sub-grids in 2 dimensional space. The sub-grids are used in the same manner as the uniform grids when finding mates and checking emptiness of a sphere.

In searching the space, the space is first divided into several regions by analyzing Delaunay triangulation patches. The region with the most number of Delaunay tri-angulation is investigated first instead of sequential investigation of the entire space. We call this approach as "the highest priority search method". This approach is es-pecially effective if the number of data points is large.

3.2 2-Dimensional Algorithm

Finding Mate to Form a Delaunay Triangle. Delaunay triangulation tends to generate an equilateral triangle by definition. Thus, the ratio of the distance between a line segment $\overline{P_1P_2}$ and the mate P_3 to the length of a line segment $\overline{P_1P_2}$ would be $\frac{\sqrt{3}}{2}$ with high probability. In addition, the chance of forming Delaunay triangle is high if the probability that the circle passing P_1, P_2 contains any points is low. That means the area of a circle passing P_1, P_2 must be minimized. Therefore, the location of mate, P_3 satisfying those condition would be on the line passing through the center of $\overline{P_1P_2}$ and perpendicular to $\overline{P_1P_2}$. and the distance from the center of $\overline{P_1P_2}$ is with high probability. To test the above hypothesis, an experiment is $\frac{1}{2}\overline{P_1P_2}$ performed as follows:

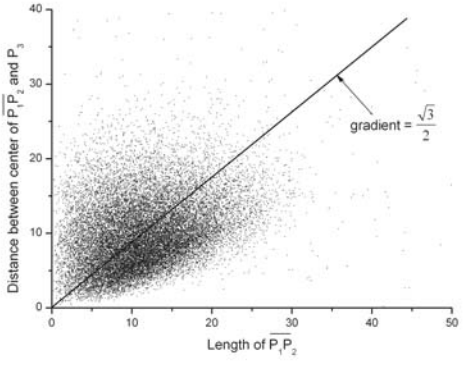


Fig. 1. Distance relation of edge length and mate point (uniform distribution)

After Delaunay triangulation with 50,000 uniformly distributed data points, 10,000 Delaunay triangles are randomly sampled and checked the length of $\overline{P_1P_2}$ and the distance $\overline{P_1P_2}$ between and P_3 are examined as shown in Fig 1. Fig.2 shows the results obtained by normalizing the length of $\overline{P_1P_2}$ and transformed its coordinates into (0, 0), (1, 0). In Fig. 2, the points are most densely distributed in the region with $\frac{1}{4} \sim \frac{3}{4}$ of the height of equilateral triangle. Also, about 80% of the points are within the square of size 2. (The test results with normally distributed points showed similar trends.) The experimental results indicate that those regions must be explored prior to other regions.

Based on the experimental observation, an algorithm to find a mate in 2-dimensional space is developed as follows: 0. Select a point P_1 in the central grid and pick P_2, as the nearest point from P_1.

1. Starting from the center of $\overline{P_1P_2}$, create line segment, $\bar{A}$ which is perpendicular to $\overline{P_1P_2}$ and half the length of $\overline{P_1P_2}$.
2. Let the other end of $\bar{A}$ P_3. Then P_1, P_2 and P_3 form an isosceles triangular.

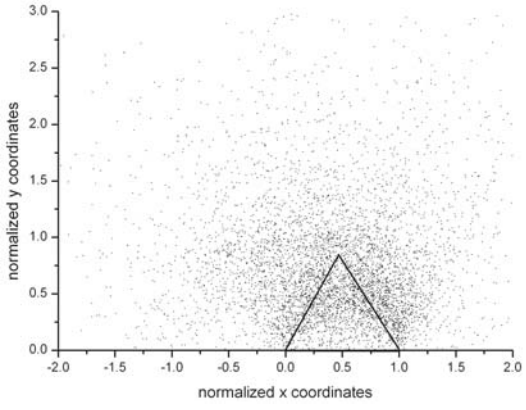


Fig. 2. Position relation of normalized edge($\overline{P_1P_2}$)and mate point(uniform distribution)

3. If the slope of $\bar{A}$ falls in the region 1 or 3 in Fig. 3, extend $\bar{A}$ to the direction of $\pm x$ as shown in Fig. 4 and investigate the grid containing that line segment. In other cases, extend $\bar{A}$ to the direction of $\pm y$ and investigate the grid containing this line segment.

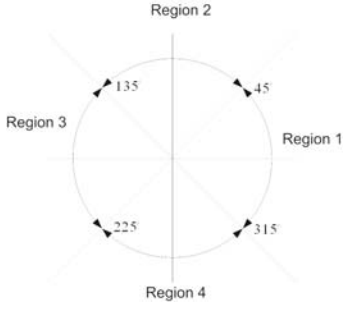


Fig. 3. The search territory selection which it follows to the angle of edge A

4. If there is no point in that grid, investigate the top and bottom (left and right if $\bar{A}$ falls in the region 2 or 4 in the Fig. 3), grids adjacent to that grid. In the next step, investigate outside of this region.
   Stop investigation if boundary of the grid space is reached.
   If x_{mid} (x coordinate of the center of $\overline{P_1P_2}$) is reached, extend $\overline{P_1P_2}$ and investigate the grids between the grid containing $\overline{P_1P_2}$ and x_{mid}.
5. If there are sub-grids, the closest sub-grid from the searching area is investigated first. By performing these steps, the left side of the line segment $\overline{P_1P_2}$ is seldom investigated. 6. If a point is in the searching area, draw a circle passing P_1, P_2 and P_3 and create a square containing the circle. Then, find

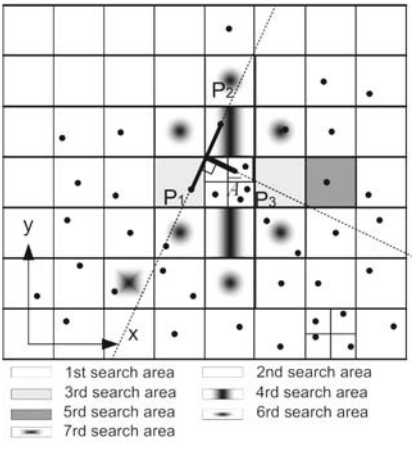


Fig. 4. Process to make a triangle, given an edge in case of domain 1, domain 3

the points other than $P_1P_2P_3$ in the grids which the square covers and check if the points are in the circle. If there is no point in the circle, P_3 is a mate of $\overline{P_1P_2}$ and forms a new Delaunay triangle.

Constructing the Delaunay triangle. In triangulation, creating a perfect triangle without any error is the most important. That is, there must be no holes or bridges. For constructing the Delaunay triangle Fang's method is used with a modification to accomodate sub-grids.

3.3 3-Dimensional Algorithm

Finding mate to form a Delaunay tetrahedron. For 3-dimensional case, we performed similar experiments to 2-dimensional case. Fig. 5 shows the relationship between the average edge length of the triangle $P_1P_2P_3$ and the distance from the center of the triangle to the mate P_4 when forming a Delaunay tetrahedron. The results are the average of 10 experiments with 10,00 randomly selected Delaunay tetrahedron constructed from 100,00 uniformly distributed points in each experiment.(As in 2-dimensional case, the results with normally distributed data points showed similar behavior.)

The results show that the points which has high chance of forming a tetrahedron are distributed in the region with the distance of $\frac{\sqrt{6}}{3}L$ from the center of the triangle, where L is the average edge length of a triangle. (In practice, the region with 80% of the distance has higher density than any other region.) Thus, the grids with the distance of $\frac{4}{5}\frac{\sqrt{6}}{3}L$ are investigated first. Fig. 6 shows the normalized results of the average distance between the Delaunay triangle and the mate. The center coordinate of triangle is transformed to (0, 0, 0).

Search algorithm. As observed in the above experiment, the mates are concentrated in the region $\frac{4}{5}\frac{\sqrt{6}}{3}L$ apart from the center of the triangle and most

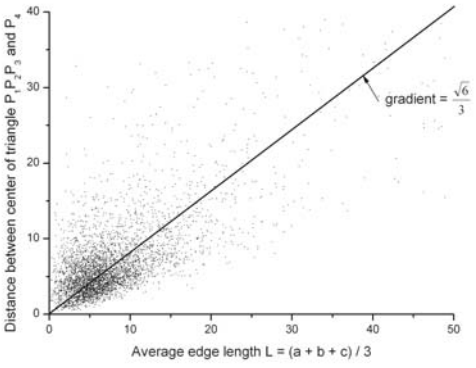


Fig. 5. Distance relation of average edge length and mate point(uniform distribution)

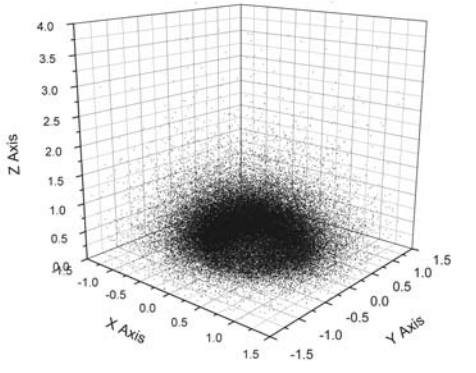


Fig. 6. Position relation of normalized triangle and mate point(uniform distribution)

of the mates are located inside the sphere with radius of and centered in that region. Therefore, it would be more efficient to explore those regions first.

Before starting the search, the first step is to obtain the normal vector of the Delaunay triangle and to find maximum among the absolute coordinate of the normal vector. For example, if the normal vector is (1, -2, 7) then the absolute value of z-coordinate is the maximum. In that case, the direction perpendicular to the z-axis(xy-plane) is investigated first. The search method is similar to 2-dimensional search except that the searching direction is 3-way instead of 2-way and the searching area is surroundings of the axis instead of left and right grids of the axis.

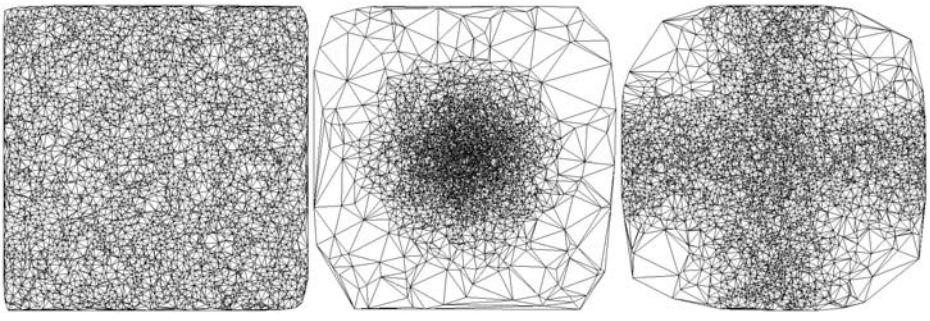
3.4 Constructing the Delaunay Tetrahedron

In 3-dimensional triangulation, creating a perfect tetrahedron without any error is the most important similar to the 2-dimensional triangulation. A good solution to prevent an error is to choose a point inside the 3D grid and create a line segment from that point. Then create a triangle using the line segments, and create tetrahedron with the triangles in a sequential manner. Once a tetrahedron is created, new tetrahedron should be searched and added easily based on the face of the previous tetrahedron without any geometric errors. The pseudo code of constructing tetrahedron is as follows:

```
Find a point, P1 in the central grid
Pick P2, the nearest point from P1
Find P3 which forms an empty meridian sphere with P1 and
P2 and form the first triangle
Find the first tetrahedron(search P4)
Initialize the list of faces as P3P2P1, P1P2P4, P2P3P4, and P3P1P4
while(there is at least a value in the list of faces){
  Set the last entry in the list of faces as current face
  if(P4 exists){
    if(P4 was already used point){
      evaluate contacting condition;
      1. if contacting with one face :
         delete current face from the list
         delete the contacting face from the list
         add two non-contacting faces to the list
      2. if contacting with two faces :
         delete current face from the list
         delete the contacting faces from the list
         add one non-contacting faces to the list
      3. if contacting with three faces :
         delete current face from the list
         delete the contacting faces from the list
      4. if contacting with point or edge
         delete current face from the list
         add three non-contacting faces to the list
    }else{
      delete current face from the list
      add three non-contacting faces to the list
      Flag P4 as used
    }
  }else{// boundary surface is reached
    delete current face from the list
  }
}
```

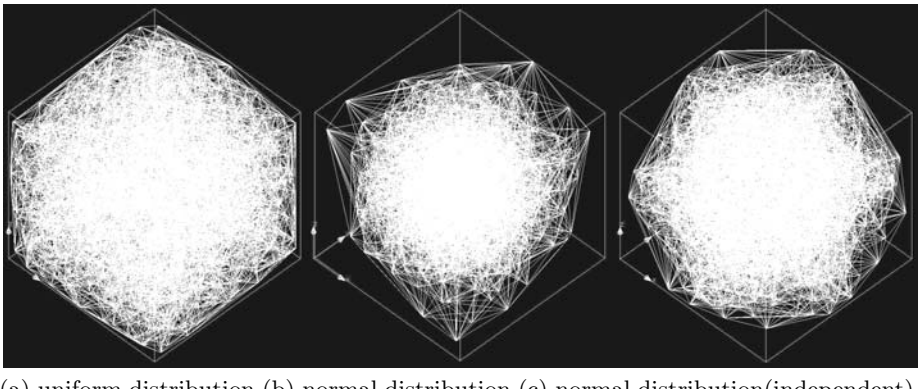
4 Results

Fig. 7 (a) shows the results of Delaunay triangulation with 5,000 uniformly distributed data points. Fig. 7 (b), (c) show the results of Delaunay triangulation with 5,000 normally distributed data points of which x, y, and z coordinates are independent. Fig. 8 shows the results of Delaunay triangulation with 5,000



(a) uniform distribution (b) normal distribution (c) normal distribution(independent)

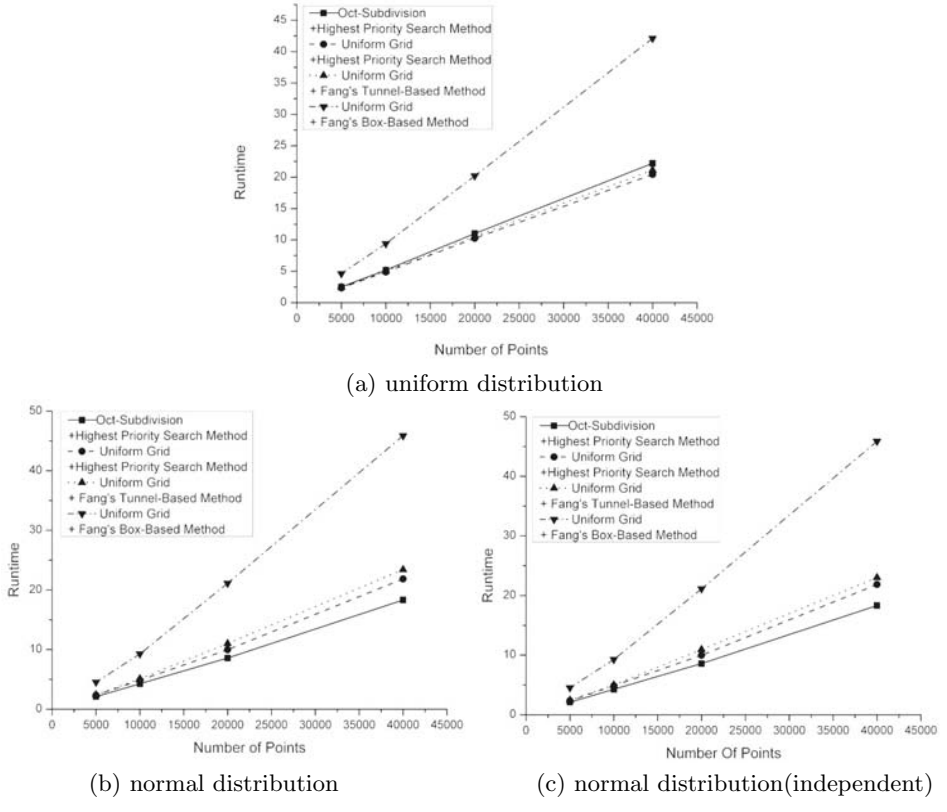
Fig. 7. Delaunay Triangulation in 2D



(a) uniform distribution (b) normal distribution (c) normal distribution(independent)

Fig. 8. Delaunay Triangulation in 3D

data points in 3D. Fig. 9 shows the execution time in 3D space when the number of data points is 5,000, 10,000, 20,000, and 40,000 respectively. The graphs show the normalized execution time of the proposed algorithm (oct-subdivision + the highest priority search method), uniform grid + the highest priority search method, uniform grid + tunnel-based method(This method is based on the experimental observation that while forming a Delaunay tetrahedron, the fourth point would be located above the center of the triangle with high probability.) by Fang[2], and uniform grid + Box based method when the execution time of the proposed algorithm with 5,000 data points is set to 1. If the data points are



(a) uniform distribution

(b) normal distribution

(c) normal distribution(independent)

Fig. 9. Running time of Delaunay triangulation algorithms in 3D

uniformly distributed, uniform grid method performed $3 \sim 4\%$ better than oct-subdivision method(Fig. 9 (a)). That is because the pre-processing time to create oct-subdivision increased the total execution time. However, with non-uniformly distributed data the proposed algorithm outperforms other algorithm more than 25% on the average in terms of execution time. (Fig. 9 (b), (c)) That is because when data points are densely located, oct-subdivision reduces the search area effectively and the highest priority search method increased the efficiency at the same time. On the other hand, the results in Fig. 9 show that the execution time of the proposed algorithm increases linearly in proportion to the number of data points. (We have empirically confirmed with 1,000,000 data points.) That shows the proposed algorithm works effectively with large number of data points.

5 Conclusion

Uniform grid method enables us to find geometric features in a specific area in a short time. However, uniform grid method takes loner execution time if data points are densely distributed. In this study a heuristic with oct-subdivision

and the highest priority search method is proposed to improve the efficiency of uniform grid method. The results show that the proposed algorithm performs on equal terms with the conventional uniform grid method in terms of execution time and outperforms them if the data points are densely distributed. The oct-subdivision reduces searching space effectively and the highest priority search method reduces the searching time by searching most probable area first.

The proposed algorithm can generate Delaunay triangulation faster than Incremental Construction algorithm or uniform grid method especially if the number of data points is large. Also, the proposed algorithm eliminates the possibility of creating non-Delaunay triangle since it performs Delaunay triangulation based on the definition and thus it always satisfy convex hull condition.

References

1. Tsung-Pao Fang, Les A. Piegl : Delaunay Triangulation Using a Uniform Grid. IEEE Computer Graphics and Applications, Vol. 13, No. 3 (1993) 36-47
2. Tsung-Pao Fang, Les A. Piegl : Delaunay Triangulation in Three Dimensions. IEEE Computer Graphics and Applications, Vol. 15, No. 5 (1995) 62-69
3. H. Edelsbrunner, N.R. Shah : Incremental topological flipping works for regular triangulations. In Proceedings of the 8th Annual ACM Symposium on Computational Geometry (1992) 43-52
4. L.J. Guibas, D.E. Knuth, M.Sharir : Randomized incremental construction of Delaunay and Voronoi diagrams. In Automata, Languages and Programming, LNSC N.443, Springer-Verlag (1990) 414-431
5. D.H. McLain : Two dimensional interpolation from random data. The Computer J., 19(2) (1976) 178-181
6. D.P. Dobkin, M.J. Laszlo : Primitives for the manipulation of three-dimensional subdivisions. Algorithmica, 4:3-32 (1989)
7. D.T Lee and B.J. Schachter : Two algorithms for constructing a Delaunay triangulation. Int. J. of Computer and Information Science, 9(3) (1980)
8. P. Cignoni, C. Montani, R. Scopigno : DeWall : A Fast Divide and Conquer Delaunay Triangulation Algorithm in Eb", Computer-Aided Design, Vol. 30, No. 5 (1998) 333-341 9. V.Akman, W.R. Franklin, M. Kankanhalli, C. Narayanaswami : Geometric computing and uniform grid technique. Computer-Aided Design, 21(7) (1989) 410-420
9. V.Akman, W.R. Franklin, M. Kankanhalli, C. Narayanaswami : Geometric computing and uniform grid technique. Computer-Aided Design, 21(7) (1989) 410-420
10. Deok-Soo Kim, Il-Kyu Hwang, Bum-Joo Park : Representing the Voronoi Diagram of a Simple Polygon using Rational Quadratic Bezier Curves. Computer Aided Design Vol. 27, No. 8 (1995) 605-614
11. Michael J. Laszlo : Computational Geometry and Computer Graphics in C++. Prentice Hall. (1996)

Feature-Based Texture Synthesis

Tong-Yee Lee and Chung-Ren Yan

Computer Graphics Group/Visual System Lab(CGVSL),
Department of Computer Science and Information Engineering,
National Cheng-Kung University, Tainan, Taiwan, Republic of China
tonylee@mail.ncku.edu.tw

Abstract. We introduce a new method for texture synthesis on regular and irregular example textures. In this paper, an enhanced patch-based algorithm is proposed to select patches with the best structural similarity and to avoid discontinuity at the boundary of adjacent patches. This new method utilizes a feature-weighted function to measure the structural similarity. A re-synthesis technique is presented to reduce the structural discontinuity. We conduct a comparison study to verify the proposed method. Preliminary experiments show that the proposed method can yield better results than other well-known methods for examples used in this study.

1 Introduction

Texture synthesis is a popular research topic in computer graphics. In the past, many previous methods have been presented. The kernel of these previous works is structural similarity matching. Generally, there are two classes of methods: 1) pixel-based [1, 2, 6, 7, 8] and patch-based [3, 4, 5, 9, 10] methods, respectively. Pixel-based algorithms synthesize only one pixel at a time. Efros et al. [2] synthesize a new pixel by searching the sample image and finding the pixel that has the most similar neighborhood. This algorithm works fine if a larger neighborhood size is chosen to measure similarity. Therefore, it is very slow. Wei et al. [7] utilize a pyramid synthesis and a tree-structured vector quantization method to accelerate [2]. Later, Ashikhmin [1] extends [7] by reducing the searching space of the input image to several candidates. Hertmann [6] presents the other applications, like learning painting styles, based on [7]. Zelinka et al. [8] create a novel jump map, which stores a set of matching input pixels (jumps) for each input pixels. It allows their algorithm to synthesize texture in real time.

In contrast to pixel-based methods, patch-based methods synthesize a patch at a time. Xu et al. [3] synthesize new textures by random patch pasting. The problem with random patch pasting algorithm is that the discontinuity and artifacts at the overlapped region of the adjacent patches. Efros et al. [4] present a minimum-error-boundary-cut within the overlapped regions to reduce the discontinuity and artifacts. Liang et al. [5] apply feathering technique on the overlapped regions and accelerate the search by a quad-tree pyramid data structure. Wu et al. [9] produce a feature map and texture map to guide the patch selection

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1043–1049, 2005.

and align overlapped region by measuring structural similarity. In contrast to most patch-based methods, Nealen et al. [10] adaptively splits patches to suitable sizes while satisfying user specified error for the mismatching in the overlapped regions. A pixel-based re-synthesis is then applied to the mismatched pixels in overlapped regions.

2 Methodology

To synthesize texture, most algorithms consist of two steps: 1) searching most similar neighborhood in the input image and 2) avoiding discontinuity in the generated texture. In this paper, the proposed patched-based method uses a feature-weighted function to guide our searching for the most similar neighborhoods and a re-synthesis approach to reduce the artifacts at the overlapped regions.

2.1 Similar Neighborhood Search

Our approach is a patch-based texture synthesis algorithm and it consists of the following steps:

1. Select an initial patch from input texture randomly and paste it to the left top corner of the output texture.
2. Search a new adjacent patch that has the most similar neighborhood in the input texture constrained by a L-shape.
3. Paste the selected patch to the output image in a scan-line order.
4. Re-synthesize the overlapped region.
5. Repeat step 2 to 4 until we complete output image.

In the past, most methods compare the difference of pixel color when they search the most similar L-shape neighborhood. The importance of each pixel in the L-shaped neighborhood is treated equally. However, in these methods, most artifacts happen to the boundary of adjacent patches. To alleviate this problem, we propose to strengthen the importance of pixels closer to the patch boundaries as we measure the similarity of patches. In our algorithm, a Gaussian distribution function is used to modify the importance of pixels. The distance weight equation W_d is defined as follows.

$$W_d(i,j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{d(i,j)^2}{2\sigma^2}} \tag{1}$$

where $d(i,j)$ is the distance form pixel (i,j) to the closest patch boundary. Figure 1 illustrates the visualization of weight variation for each pixel in L-shape.

To judge whether a synthesis result is good or not depends on if its edge structure is similar to original texture or not. Hence, a structure-weighted function W_s is also considered as the similarity-searching criterion. In eq. 2, W_s is described as follows.

$$W_s(i,j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{[|M_{color} - P(i,j)| - 255]^2}{2\sigma^2}} \tag{2}$$

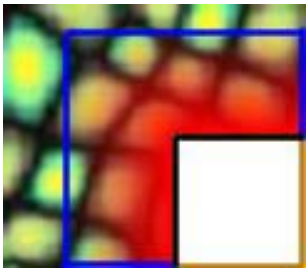


Fig. 1. Visualization of weight variation. Black line represents patch boundary at L-shape. The pixel color in L-shape means the weight of a pixel where red color represents the higher weight

where M_{color} represents the color with the maximum number of pixels among all similar colors in L-shape, $P_{i,j}$ is the color of the pixel at (i, j). In L-shape, most pixels should have colors that are closer to the based color M_{color}, and the structure elements such as edges and lines should have a larger difference to this based color. Therefore, a higher weight is given to pixels that have a larger difference to the based color M_{color}. Finally, both distance and structure weighted function, W_d and W_s, are integrated into our similarity measurement and is called the feature-weighted function below.

$$Sim = \sum_{i,j \in L-shape} P_{diff}(i,j)(W_d + W_s) \tag{3}$$

where $P_{diff}(i, j)$ is the color difference of pixels in the L-shape between input and output images. Searching similar neighborhood using eq. 3 can help us find a patch with less artifacts at the boundary.

2.2 Repair Artifacts in Overlapping Regions

After the most similar patch is selected and is pasted into the output image, some artifacts may still exist. In this section, we will further reduce artifacts. First, we compute the error for each pixel in the overlapped region. Then, we use Efros et. al's method [4] to find a minimum error path through overlapped region. After the minimum error path is found, we treat this path as the central axis and define a five-pixel wide region, called *repairing area*. Figure 2 shows an example of a repairing area.

In the repairing area, we treat pixels as mismatched pixels if their errors exceed a selected threshold δ. Then we re-synthesize these mismatched pixels in a region-growing manner from the boundary of the repairing region to its central axis. The value of the mismatched pixel is repaired by the median value of the 5×5 matched neighboring pixels. After re-synthesizing all mismatched pixels, we finally apply Gaussian smoothing filter to these newly repaired pixels.

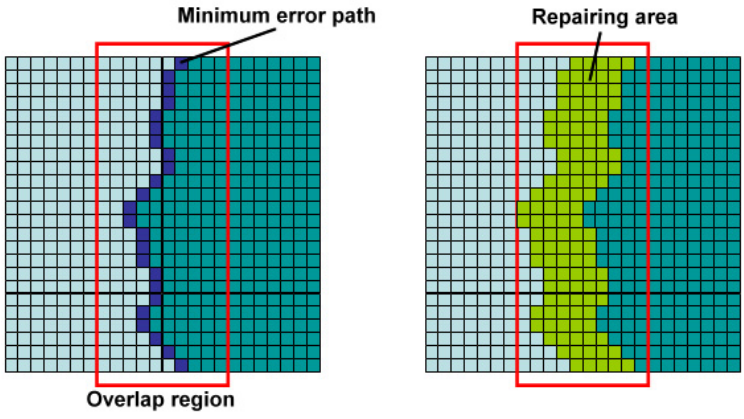


Fig. 2. Left: Computing the error for each pixel in overlapped region and finding a minimum error path. Right: Repairing area

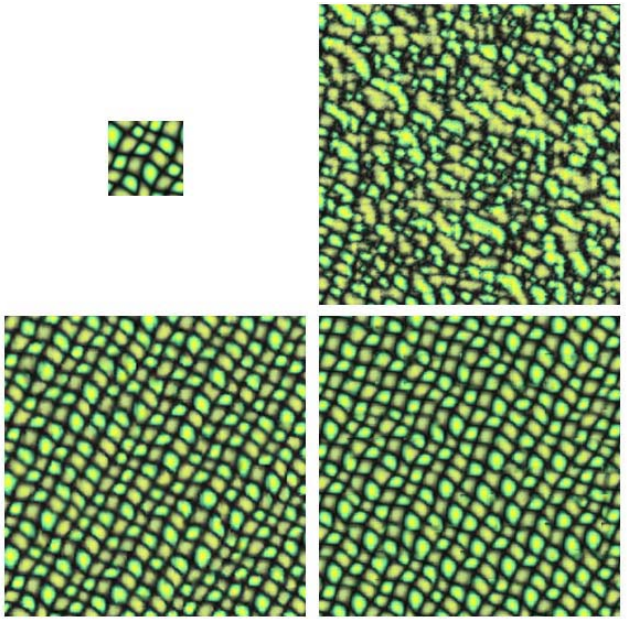


Fig. 3. Experimental Results

3 Experimental Results

In this section, we demonstrate some preliminary results in comparison with two well-know methods: pixel-based [1] and patch-based [5] algorithm. Figure 3 and Figure 4 are near-regular texture. Figure 5 shows an irregular texture example.

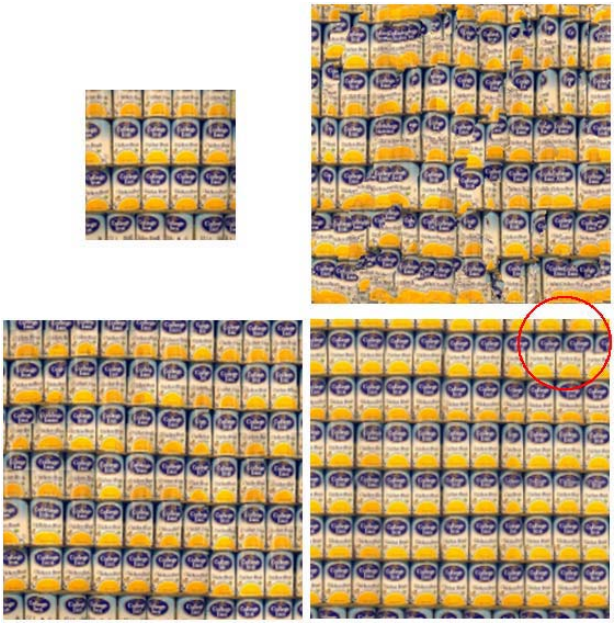


Fig. 4. Experimental Results

Fig. 5. Experimental Results

In these three examples, the left-top image is original input, the right-top image is synthesized by [1], the left-bottom image is synthesized by our method and the right-bottom image is synthesized by [5]. From our preliminary results on these three examples, the pixel-based method [1] does not work well. The results by our method are comparable to those by [5]. For further comparison, we see some difference between [5] and ours. In Figure 3, the structure distribution, grid size and shape by the proposed method are almost similar to those of the original sample texture. Figure 4 shows that our method can yield better result in preserving complete structural elements and continuity at the patch boundary. Please watch the top-right part of the result by [5]. There is some obvious discontinuity on the shapes of cans. In Figure 5, it is hard to compare difference between ours and [5].

4 Conclusion and Future Work

Pixel-based texture synthesis tends to introduce blurring effect and is more suitable for irregular textures, like clouds, ripples, and other natural images. Patch-based texture synthesis can preserve the global structure of the sample texture, but to reduce artifacts and discontinuity at patch boundary is an important work. In this paper, we proposed a new algorithm on both regular and irregular textures. The proposed method is a patch-based algorithm and it consists of two parts: 1) we first apply feature-weighted function to search the most similar neighborhood patches and 2) then, we re-synthesize mismatched pixels to reduce artifacts and to minimize the structural discontinuity. This new scheme can solve these problems well. In the near future, we plan to extend the current work to the perspective texture synthesis. We also like to apply our method to some textures with special structure, like the wing of butterfly and the tail of peafowl fish. In addition, we would like to apply our technique to synthesize texture for the expression change of head animation [11] or for 3D morphing applications [12, 13].

Acknowledgements. This paper is supported by the National Science Council, Taiwan, Republic of China, under contract No. NSC-93-2213-E-006-026 and NSC-93-2213-E-006-060.

References

1. M. Ashikhmin: Synthesizing natural textures. ACM Symposium on Interactive 3D Graphics, (2001) 217–226.
2. A. Efros and T. Leung: Texture synthesis by non-parametric sampling. Intl. Conf. Computer Vision, (1999) 1033–1038.
3. Y. Xu, B. Guo, And H.-Y Shum: Chaos mosaic: Fast and memory efficient texture synthesis. Microsoft Res. Tech. Rep. MSR-TR-2000-32, April (2000).
4. A. Efros, and T. Freeman: Image Quilting for Texture Synthesis and Transfer. Proceedings of SIGGRAPH, (2001) 27–34.

5. L. Liang, C. Liu, Y. Xu, B. Guo, and H.-Y Shum: Real-time texture synthesis using patch-based sampling. ACM Trans. Graphics, Vol. 20, No. 3, (2001) 127–150.
6. A. Hertmann, C. Jacobs, N. Oliver, B. Curless, and D. Salesin: Image analogies. Proceedings of SIGGRAPH, (2001) 327–340.
7. L.-Y. Wei And M. Levoy: Fast texture synthesis using tree-structured vector quantization. Proceedings of SIGGRAPH, (2000) 479–488.
8. S.Zelinka and M. Garland: Towards real-time texture synthesis with the jump map. Proceedings of the Thirteenth Eurographics Workshop on Rendering Techniques, (2002) 99–104.
9. Q. Wu and Y. Yu: Feature Matching and Deformation for Texture Synthesis. ACM Transactions on Graphics (SIGGRAPH 2004), Vol. 23, No. 3, (2004) 362–365.
10. A. Nealen and M. Alexa: Hybrid Texture Synthesis. Proceedings of Eurographics Symposium on Rendering, (2003) 97–105.
11. Tong-Yee Lee, Ping-Hsien Lin, and Tz-Hsien Yang: Photo-realistic 3D Head Modeling Using Multi-view Images. in Lecture Notes on Computer Science (LNCS) 3044, Springer-Verlag, May (2004) 713–720.
12. Tong-Yee Lee, and Po-Hua Huang: Fast and Intuitive Metamorphosis of 3D Polyhedral Models Using SMCC Mesh Merging Scheme. IEEE Transactions on Visualization and Computer Graphics Vol. 9, No. 1, Jan-March (2003) 85–98.
13. Chao-Hung Lin, and Tong-Yee Lee: Metamorphosis of 3D Polyhedral Models Using Progressive Connectivity Transformations. IEEE Transactions on Visualization and Computer Graphics, Vol. 11, No.1, Jan-Feb. (2005) 2–12.

A Fast 2D Shape Interpolation Technique

Ping-Hsien Lin[1] and Tong-Yee Lee[2]

[1] Department of Management and Information Technology,
Southern Taiwan University of Technology, Taiwan, R.O.C.
phlin925@mail.stut.edu.tw
[2] Department of Computer Science and Information Engineering,
National Cheng-Kung University, Taiwan, R.O.C.
tonylee@mail.ncku.edu.tw

Abstract. This paper proposes a computationally inexpensive 2D shape interpolation technique for two compatible triangulations. Each triangle in a triangulation is represented using a stick structure. The intermediate shape of each triangle is interpolated using these sticks. All of these intermediate triangles are then assembled together to obtain the intermediate shape of the triangulation according to a predetermined order. Our approach is inspired by Alexa et al's work [1], but is simpler and more efficient. Even though we ignore the local error, our approach can generate the satisfactory (as-rigid-as-possible) morph sequence like Alexa et al's.

1 Introduction

Morphing, also called metamorphosis or shape interpolation, is a technique that blends two shapes smoothly and reasonably. Its brilliant visual effects have made it appear in wide applications such as movies, games, cartoons and music videos. However, the huge necessary manual intervention always makes it very expensive and time consuming. The three main features emphasized in morphing techniques are the shape's rationality, automaticity and processing time. Shape rationality refers to the process being locally least-distorting or so-called as-rigid-as-possible [1]. Figure 1 gives a morph example of a dancing man. The morph sequence in the first row is generated by linearly interpolating the corresponding vertex between the source and target shapes. We can see that the dancing man's left hand shortens among the morph sequence. The morph sequence in the second row is generated using our method. It looks smoother and can maintain a normal human shape.

Numerous aspects of morphing techniques have been investigated, including the 2D image [2, 3, 4, 5], 3D volume data [6, 7] planar polygons, polylines [1, 8, 9, 10, 11, 12, 13], and polyhedron [1, 14]. The two main research issues in morphing are vertex correspondence [8, 15, 16, 17] and vertex trajectories [9, 10, 11]. Dividing the two shapes into two triangulations with a common topology can solve the vertex correspondence problem. This issue is called compatible triangulation and many previous papers on this problem can be found [1, 13, 18]. In this paper

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1050–1059, 2005.

we assume the source and target triangulations are already compatible and focus on the vertex trajectory problem.

Some researches [9] blend the 2D shapes using just their boundary information to produce vertex trajectory. However, a proper 2D morph sequence should also consider the information inside the boundary. An intuitive and reasonable scheme to accomplish this idea is to utilize the shape's skeleton information [10]. Generating a skeleton from a shape is complicated. For cases in which the source and target shapes are quite different, the compatible skeleton is difficult to define. In [1], the source and target shapes are triangulated into two smooth compatible triangulations. An optimization is then performed that minimizes the overall local deformation. This method can produce a satisfactory (as-rigid-as possible) morph sequence in most cases. However, global optimization is very computationally expensive. Our work is inspired by Alexa et al's work; first we discuss the morph between two triangles and then extend the process to triangulations. A three-stick structure is used to interpolate the shape of each triangle. All of the interpolated triangles are then assembled together to get the final shape according to a predetermined order.

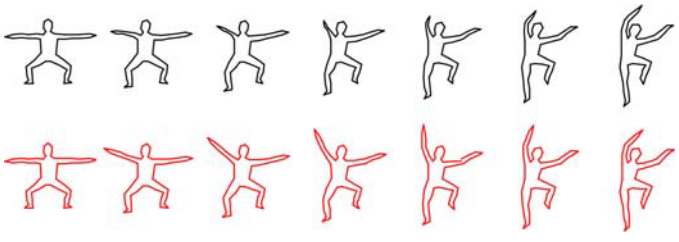


Fig. 1. The morph sequence of a dancing man. The first row was generated using linear interpolation. The second row was generated using our method

The remainder of this paper is organized as follows. In Section 2, we introduce the shape interpolation between two triangles. In Section 3, we propose our shape interpolation method for two triangulations. Section 4 demonstrates our results. Our conclusion is presented in Section 5.

2 Triangle-to-Triangle Interpolation

First we consider the simplest case that both source and target shapes are a single triangle. The corresponding vertices of these two triangles are known in advance. The question is "what is the reasonable intermediate shape for a triangle in a morph sequence"? In [1], the affine transform between the source and target triangles was decomposed into a rotation and a stretching transformation. The free parameters in the factorization were linearly interpolated to achieve the least-distorted morph. Here, we propose a more intuitive method to produce the intermediate triangle shape. We use a stick structure to represent a triangle. This

structure consists of three straight sticks with one side jointed at the centroid of the triangle and the other side connecting the triangles three vertices as shown in Figure 2. We then interpolate the corresponding sticks of the source and target triangles to get the intermediate triangle.

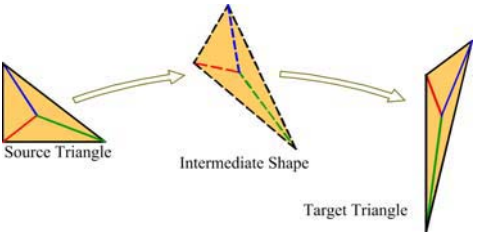


Fig. 2. The single triangle shape interpolation using sticks

2.1 Stick Interpolation

The most intuitive stick interpolation scheme is to linearly interpolate the stick's length (l) and angle (θ) using Equation 1

$$\begin{cases} l(t) = (1-t)l_S + tl_T \\ \theta(t) = (1-t)\theta_S + t\theta_T \end{cases} \tag{1}$$

where the suffix S stands for source and T for target.

The stick linear interpolation is illustrated in Figure 3. Figure 3(b) shows the stick interpolation curve at the same condition shown in Figure 3(a) but with an additional 2π rotation angle. Though the range of time parameter t needed in a morph sequence is the real number within [0,1] , a value beyond this range can be used for extrapolation.

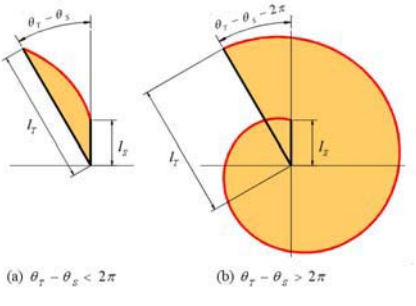


Fig. 3. The stick interpolation curves using Equation 1

A much smoother interpolation curve for sticks is the curve whose included angle between the tangent direction and the position vector at any position P_i

on the curve is fixed as shown in Figure 4, i.e., $\frac{dr}{rd\theta} = a$, where a is a constant, and P_i is represented by (r, θ) in polar coordinates. Integrating both sides of this equality, we can obtain the curve equation in polar coordinates:

$$r = Ce^{a\theta}, \quad \text{where } C \text{ is a constant.} \tag{2}$$

Applying Equation 2 to the stick interpolation in equation 1, we can get:

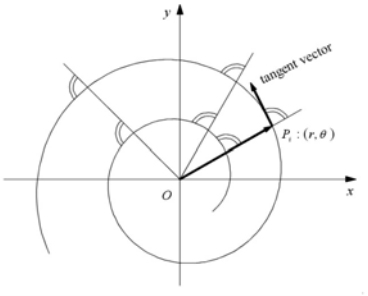


Fig. 4. The plot of $r = e^{a\theta}$

$$\begin{cases} \theta(t) = (1-t)\theta_S + t\theta_T \\ l(t) = l_S e^{a(\theta(t)-\theta_S)} \end{cases} \tag{3}$$

where $a = \frac{1}{\theta_T - \theta_S} ln \frac{l_T}{l_S}$ and $\theta(t)$ is a monotonically increasing or decreasing function that meets the conditions: $\theta(0) = \theta_S$ and $\theta(1) = \theta_T$.

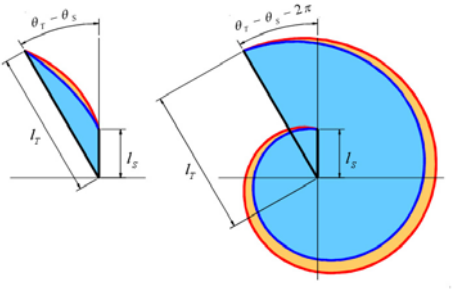


Fig. 5. The interpolation curve for Equation 3 (blue) and that for Equation 1 (red)

Figure 5 shows a comparison of the stick interpolation curves for Equations 1 and 3. We can see that the stick length for the curve in Equation 1 (the red curve) increases faster along the curve near the short stick and slower near the long stick. By contrast, the increasing rate for the stick length for the curve using

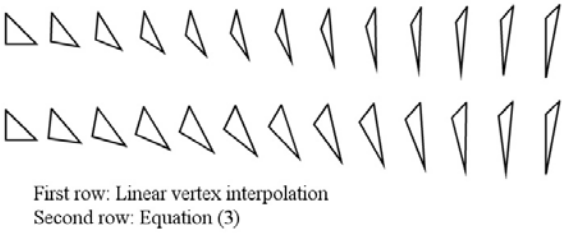


Fig. 6. The morph sequences using different interpolation methods

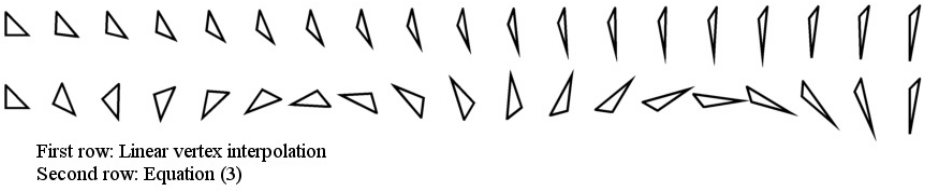


Fig. 7. The morph sequences using different interpolation methods. The triangle rotates additional 2π angle than that in Figure 6

Equation 3 (the blue curve) is constant. It therefore looks smoother. Figure 6 and 7 show the morph sequences generated by linear vertex interpolation and Equation 3. It is obvious that using Equation 3, we can generate more reasonable intermediate shapes. Moreover, linear vertex interpolation cannot deal with the case that the triangle rotate more than π.

3 Triangulation Shape Interpolation

At morph time t, each triangle in the intermediate shape is generated by interpolating the corresponding triangle pair between the source and target shapes using the stick interpolation described in Section 2. All of these interpolated intermediate triangles are then assembled together to form the intermediate triangulation. However, these intermediate triangles cannot be assembled together tightly because each triangle was interpolated individually such that the corresponding edges of two neighboring interpolated triangles may have different lengths and angles. In [1], they used an optimization method that minimizes the overall local deformation to deal with this problem. Though such approach is very reasonable and precise, it requires much computation cost. In this section we introduce an approach that can perform triangulation shape interpolation very fast.

3.1 Assembly Order Decision

All of the intermediate triangles are assembled together one by one according to a predetermined order to generate the intermediate triangle shape. A planar graph [19] is built first from the triangulation according to its position. One triangle is chosen as the root triangle. The root triangle is the foremost in the assembly

order. The remainder assembly order is obtained by traversing the planar graph from the root in breadth-first order.

3.2 Assembly Rule

A mechanism for assembly is needed because the interpolated intermediate triangles cannot be tightly assembled together. Assume that the rotation angles of all of the sticks of all triangles are known (this will be introduced in Section 3.3). At the beginning, set the intermediate triangulation to a null triangulation (a triangulation com-posed of zero triangles). The shape of the root triangle is interpolated using stick interpolation. The interpolated root triangle is added to the intermediate triangulation. According to the assembly order, assemble the triangles one by one to the intermediate triangulation. A principle rule in our method is that once the vertex position has been determined it will never be changed in the process thereafter. Therefore, adding a triangle to the intermediate triangulation involves adding one or no new vertices to the intermediate triangulation. Our approach simply ignores the local errors. For example, as shown in Figure 8, triangle A is a triangle in the current reconstructed intermediate triangulation. Triangle B is the triangle to be added to the current reconstructed triangulation and is neighboring to triangle A. The intermediate shape of triangle A has been decided. The three sticks for triangle B are drawn in dashed lines. The intermediate shape of triangle B, generated by stick interpolation, is drawn in thin solid lines. First, choose one vertex from their shared edges (colored in red) as a base point. Use two sticks (colored in blue) to determine the position of the new vertex (see Section 3.3 for the rotation angle decision of sticks). The remaining stick is useless. The final triangulation is drawn in a thick black solid line. For the case involving adding a new triangle with no new vertex, we just do nothing. This approach may cause error accumulation propagating from the root triangle. However, if both the source and target shapes are triangulated properly (no narrow triangles), according to the results presented in Section 4, a satisfactory morph sequence can always be generated. In the case in which some narrow triangles exist in the source and target compatible triangulations, we can use the method described in [1]; compatible mesh smoothing combined with edge splitting to smooth out the triangulations.

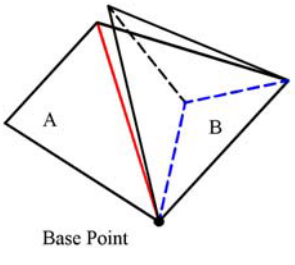


Fig. 8. Assemble the triangle by ignoring the local errors

3.3 Stick Rotation Angle Calculation

To interpolate the intermediate triangle shape using stick interpolation, we need to know the rotation angles of all sticks. The stick rotation angle α and minimum rotation angle (or included angle) α_{min} from the source to the target orientation are related by $\alpha = \alpha_{min} + 2n\pi, n \in Z$. The question is how to quickly find the n value? First of all, the rotation angles of the three sticks of the root triangle are deter-mined using their minimum rotation angle. We can also assign the rotation angle of these sticks of root triangle with an additional $2n\pi$, where $n \in Z$. This will cause the global shape to rotate an additional n cycles around the centroid of the root triangle along the morph sequence. To determine the n value for all the other triangle sticks, we need to find a boundary that is smaller than 2π. As shown in Figure 9(a) (the sticks are drawn with dashed lines), assume that we already know the rotation angle (α) of the red edge on triangle A. The rotation interval of the blue stick on triangle B, relative to this red edge, must be restricted within a π region (shown in semicircular arrows) within the morph sequence. That means the rotation angle (β) of the blue stick must be within the bounds: $\alpha - \pi < \beta < \alpha + \pi$. Thus, n is restricted by

$$\alpha - \pi < \beta_{min} + 2n\pi < \alpha + \pi \Rightarrow \frac{\alpha - \pi - \beta_{min}}{2\pi} < n < \frac{\alpha + \pi - \beta_{min}}{2\pi} \quad (4)$$

From Equation 4, we can get the n value and thus obtain the blue stick's rotation angle (β). In Equation 3, the stick's angle is linearly interpolated. We can therefore conclude that the included angles between the neighboring sticks also vary linearly. Therefore, in Figure 9(a), the rotation angles of the green and purple sticks on triangle B can be obtained easily by adding the corresponding difference included in angles a and b between the source and target stick pairs to the rotation angle of the blue stick. The rotation angle of an edge can be likewise determined quickly. In Figure 9(b), the rotation angle of the red edge is also restricted to a π region against the blue stick (shown in semicircular arrows). Again, we can use Equation 4 to find n for the red edge. These two procedures are performed alternately from the root triangle along the assembly order until all the rotation angles of the edges and sticks of all triangles are found.

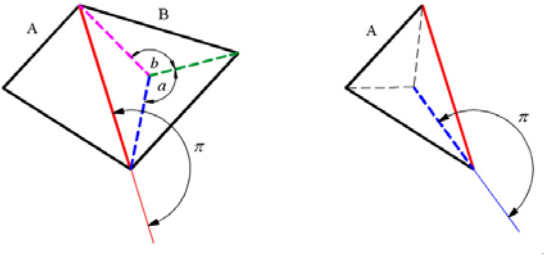


Fig. 9. Rotation angle bound between a connecting stick and edge

4 Experimental Results

Our shape interpolation program was implemented on a Pentium IV 2.4GHz machine. All of the initialization steps from our test data, including the assembly order decision and stick rotation angle calculation, were performed in seconds. All of the morph sequences can be generated in real time. Note that, in our method, the initialization steps only need to be performed once within a morph sequence. At each time t, we only need to interpolate the sticks and assemble them to form the intermediate triangulation. Figure 10 shows four morph sequences. The root triangles are colored in red. We can see that all of the morph sequences are very reasonable (as-rigid-as- possible). Note that, in Figures 10(a), (b), even though there are some narrow triangles in the triangulation, our method can still generate good morph sequences. Figure 10(d) shows that our method can also be applied to a rigid transformation (i.e., rotation and translation). In this case, our method does not counter the error accumulation problem. All of the intermediate shapes are satisfactory.

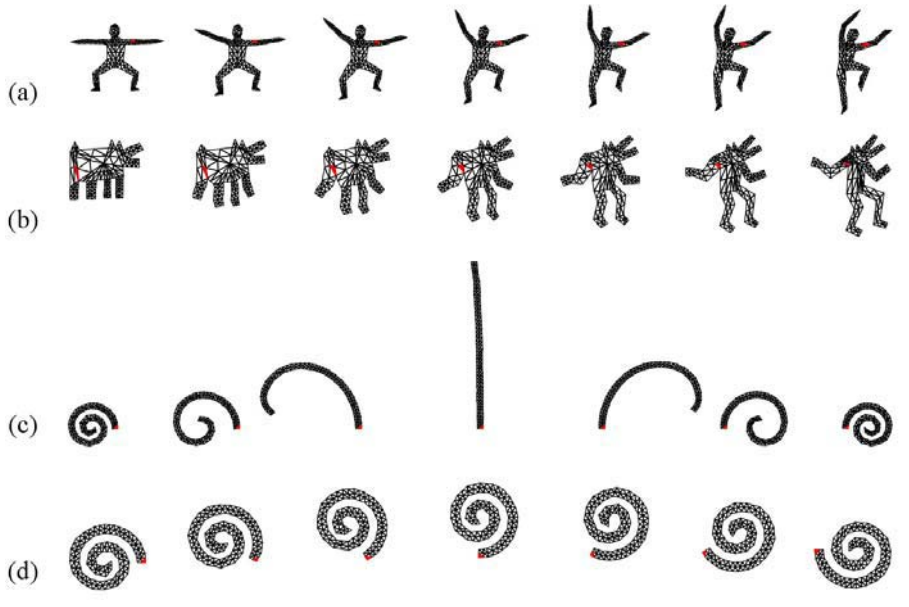


Fig. 10. Four morph sequences using the proposed method

5 Conclusions and Future Work

We proposed a fast 2D shape interpolation technique for compatible triangulations. It is very easy to implement. The morph sequences generated by this method are very reasonable (as-rigid-as-possible). As demonstrated in Figure 10, our method is applicable to most cases. Though our method seems robust,

it is expected that in some cases our method will fail due to the error accumulation. However, such cases occur rarely and can always be solved by compatibly smoothing the source and target triangulations. In addition, we would like to seek the possibility of applying our techniques to improve the triangle quality of 3D morphing applications [20, 21].

Acknowledgements. This paper is supported by the National Science Council, Taiwan, Republic of China, under contract No. NSC-93-2213-E-006-026 and NSC-93-2213-E-006-060.

References

1. M. Alexa, D. Cohen-Or, and D. Levin: As-rigid-as-possible shape interpolation. SIGGRAPH 2000 Proceedings, July (2000) 157–164.
2. T. Beier and S. Neely: Feature-Based Image Metamorphosis. Proceedings of SIGGRAPH '92, Volume 26 (1992) 35–42.
3. S. Y. Lee, K. Y. Chwa, S. T. Shin, and G. Wolberg: Image Metamorphosis Using Snakes and Free-Form Deformations. SIGGRAPH '95 Proceedings, (1995) 439–44.
4. G. Wolberg: Image Morphing Survey. Visual Computer, vol. 14, (1998) 360–372,.
5. A. Tal and G. Elber: Image morphing with feature preserving texture. Computer Graphics Forum (Eurographics '99 Proceedings), 18(3), (1999) 339–348,.
6. J.F. Hughes: Scheduled Fourier Volume Morphing. SIGGRAPH '92 Proceedings, 26, 2, (1992) 43–46.
7. T. He, S. Wang, and A. Kaufman: Wavelet-Based Volume Morphing. Proceedings of Visu-alization'94, (1994) 85–92.
8. T. W. Sederberg and E. Greenwood: A physically based approach to 2D shape blending. SIGGRAPH '92 Proceedings, volume 26 (1992) 25–34.
9. T. W. Sederberg, P. Gao, G. Wang, and H. Mu: 2D shape blending: An intrinsic solution to the vertex path problem. SIGGRAPH '93 Proceedings, volume 27 (1993) 15–18.
10. M. Shapira and A. Rappoport: Shape blending using the star-skeleton representation. IEEE Computer Graphics and Applications, vol. 15, issue 2, Mar. (1995) 44–50.
11. E. Goldstein and C. Gotsman: Polygon Morphing Using a Multiresolution Representation. Proceedings of Graphics Interface, May (1995) 247–254.
12. M. S. Floater and C. Gotsman: How to morph tilings injectively. Journal of Computational and Applied Mathematics, 101:117-129 (1999).
13. V. Surazhsky and C. Gotsman: Morphing stick figures using optimized compatible triangu-lations. Proceedings of Pacific Graphics, Tokyo, October (2001).
14. T. Kanai, H. Suzuki, and F. Kimura: 3D geometric metamorphosis based on harmonic maps. In Proceedings of Pacific Graphics '97, October (1997) 97–104, .
15. J.R. Kent, W.E. Carlson, and R.E. Parent: Shape Transformation for Polyhedral Objects. SIGGRAPH '92 Proceedings, 26, 2, (1992) 47–54.
16. M. Alexa: Merging polyhedral shapes with scattered features. The Visual Computer, 16(1) (2000) 26–37.
17. A.W.F. LEE, D. Dobkin, W. Sweldens, and P. Schroder: Multiresolution Mesh Morphing. SIGGRAPH '99 Proceedings, (1999) 343–350.

18. B. Aronov, R. Seidel, and D. Souvaine: On compatible triangulations of simple polygons. Computational Geometry: Theory and Applications, vol.3, issue 1, (1993) 27–35.
19. T. H. Cormen, C. E. Leiserson, and R. L. Rivest: Introduction to Algorithms. MIT Press, Cambridge, Massachusetts, (1994).
20. Tong-Yee Lee, Po-Hua Huang: Fast and Intuitive Metamorphosis of 3D Polyhedral Models Using SMCC Mesh Merging Scheme. IEEE Transactions on Visualization and Computer Graphics Vol. 9, No. 1, Jan-March (2003) 85–98.
21. Chao-Hung Lin, Tong-Yee Lee: Metamorphosis of 3D Polyhedral Models Using Progres-sive Connectivity Transformations. IEEE Transactions on Visualization and Computer Graphics, Vol. 11, No.1, Jan-Feb. (2005) 2–12.

Triangular Prism Generation Algorithm for Polyhedron Decomposition

Jaeho Lee[1], JoonYoung Park[1] Deok-Soo Kim[2], and HyunChan Lee[3]

[1] Department of Industrial System Engineering, Dongguk University,
26, 3-ga, Pil-dong, Chung-gu, Seoul, Korea
{rapidme, jypark}@dgu.edu
[2] Department of Industrial Engineering , Hanyang University,
17 haengdang-dong, Seong dong-gu, Seoul, Korea
dskim@hanyang.ac.kr
[3] Department of Industrial Engineering , Hongik University,
Sangsu-dong, 72-1. Mapo-gu, Seoul, Korea
hclee@wow.hongik.ac.kr

Abstract. RP(Rapid Prototyping) is often called as Layered Manufacturing because of layer by layer building strategy. Layer building strategy is classified into two methodologies. One is based on the 2D layer and the other is based on the 3D layer. 2D layer is simply created by the intersection between the polyhedron and a slicing plane whereas 3D layer is created with some constraints such as cuttability and manufacturability. Currently, 3D Layer is generated by using the boundary surface information in the native solid modeling format. However, most input data in Rapid Prototyping is the polyhedral surface data. We propose a geometric algorithm that uses the triangular prism to create 3D layers. Examples are shown to show the validity.

1 Introduction

RP(Rapid prototyping) has been developed by many vendors and research groups and supplied to market since 1992. RP has a strong point that can usually manufacture a sample within 24 hours. Especially, 3D layer is being used to overcome the limit of 2D based RP technology and handle the composite material. To achieve these goals, 3D models should be divided properly depending on individual RP equipment[6].

Currently, most RP equipments are based on 2D layer. That is, the side of the layer is vertical and this layer is called a vertical layer. Since the vertical layer is basically a 2.5D block, a stair-step effect happens. On the other hand, there are some efforts to make side walls with a curved or sloped layer. This technique is called a 3D layer technology and can be classified into three approaches as shown in Fig 1. One is a cutting and bonding approach which is to make a large prototype with topologically spherical shape by bonding thick plates. Another is a hybrid approach, which is a combination of 2.5D deposition RP technology and CNC milling operation. The other is double deposition approach which, generates a 3D layer by using the two orthogonal deposition machine.

There are some recognizable researches on the cutting and bonding approach by using thick plate. In 1955, Hope et al. developed TruSurf(True Surface System),

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1060 – 1069, 2005.

which generates a sloping layer by using a 4-axis water jet machine to the thin plate based RP machine[4]. Imre *et al.* from Delft University also developed a system based on the foam cutting[2]. In this system, a tool-path is generated for two planes connected side by side and two layers of foams are semi automatically bonded after the machining[5].

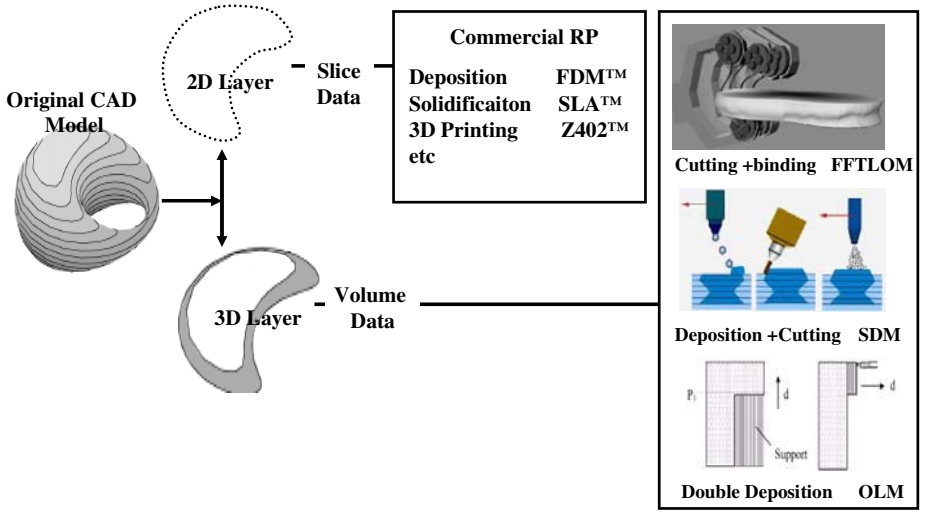


Fig. 1. Original CAD Model, 2D Layer Based Systems and 3D Layer Based Systems

This approach has an advantage of making a large prototype which is bigger than one cubic meter in a very short time. Most RP equipments available today can not handle the model with that size. This approach, however, has difficulties for making models with a cavity between the layers. For example, if a pipe with small diameter is passing through a thick layer, making a prototype without partition and cutting is not possible.

There is another approach called SDM(Shape Deposition Manufacturing) developed in Stanford university. In this approach, when a complicated model is to be made, the cutting and deposition are alternatively used after the model is decomposed into the smallest level. Although this approach is not desirable for a large prototype, it can be applied to a geometrical shape with interior cavities or holes where a traditional machining can not be used. Particularly, it has advantages of depositing high quality materials and making a model with embedded components. It can also be used in a rapid tooling which makes prototypes with a high precision. This system has been developed as a hybrid system combining a 5-axis CNC equipment and a powder based RP machine[3, 10].

Final approach is called double deposition method. This Approach is capable of using two orthogonal deposition in vertical and horizontal[12].

2 Preliminaries

We present the polyhedron based decomposition algorithm for cutting and deposition approach. The cutting and deposition approach is based on successive cutting and deposition operations. For the purpose, the Object is effectively decomposed by small volumes which, can be handled by the cutting or deposition machine.

Ramaswami *et al*. present the solid model decomposition algorithm by using the silhouette edge with ACIS kernel. They generate the silhouette edge by using the ACIS kernel and split the solid by using the infinite sweeping of the silhouette edge. In result, the three categorical volumes are generated as shown Fig. 2. these are respectively the cutting volume(I), the object decomposition volume(II-1, II-2) and the support volume(III).

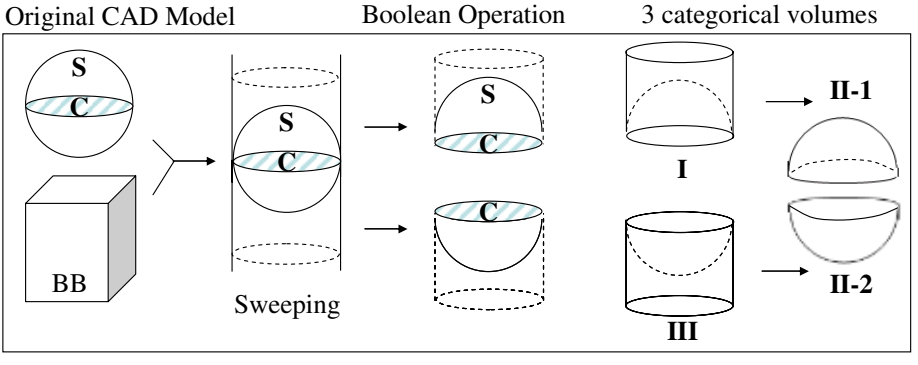


Fig. 2. Solid Decomposition in SDM

The criteria which decompose the object is the silhouette edge of the object. In SDM, they use the silhouette edge in ACIS kernel for native solid modeling system. However, in general, the polyhedron model is impossible to save the brep solid information like shell, cavity, hole etc.

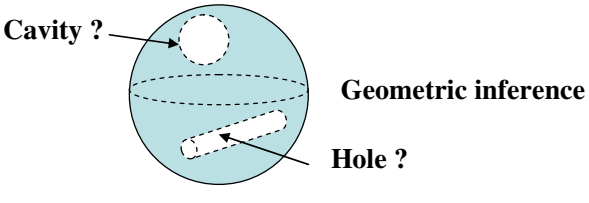


Fig. 3. Geometric Inference in polyhedron model

If we need to use this information, which can not save in polyhedron model, we must extract the features by using the geometric inference as shown in Fig. 3. However, the geometric inference is difficult to make the robust criteria by detecting the valid feature. Thus, we develop the more intuitive and robust approach for the object decomposition.

3 Polyhedron Decomposition

In general, the object decomposition is used to enumerate the space using voxels, quadtree, octree and dexels. Since their shape is box type as shown in Fig. 4, original model is converted to these models approximately and can be lost the original boundary surface information.

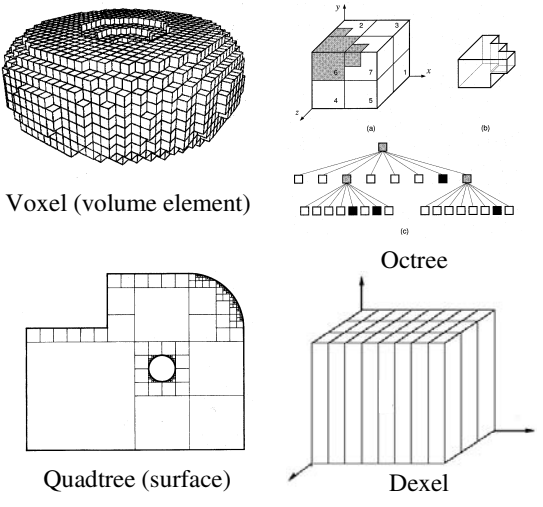


Fig. 4. Various decomposition Models

Our approach is to generate the triangular prism for the polyhedron decomposition. The triangular prism is defined as the swept volume from a triangle to the projection plane as shown in Fig. 5. The triangular prisms have the advantage of conserving the boundary surface information.

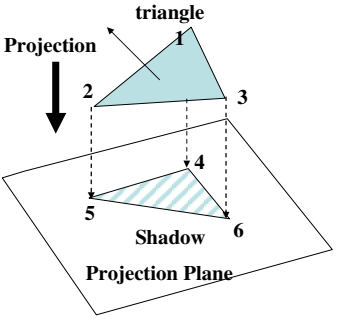


Fig. 5. Triangular prism generation of a triangle

The triangular prism generation algorithm is composed of four steps: (1) visibility test of a triangle (2) triangular prism generation using inner projection with object inside direction (3) triangular prism generation using outer projection with object

inside direction (4) spatial sort and order sequencing. Fig. 6 shows the flow chart of the proposed algorithm.

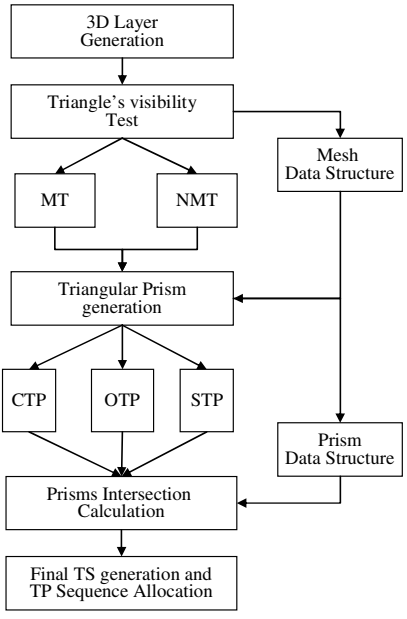


Fig. 6. The flow chart of the proposed algorithm

3.1 Triangular Prism

The triangular prism has the advantage of avoiding the Boolean operations. In SDM, the Boolean operation is needed to split the volumes. For the purpose, the clean and closed silhouette curve must be generated from the model. The silhouette curve is defined as a curve along with $N \bullet R=0$, where N is the normal vector at point(u, v) on Surface S, and R is the selected ray direction. However, in the polyhedral model, the silhouette curve do not represent in the monotonous pattern as shown in Fig 7.

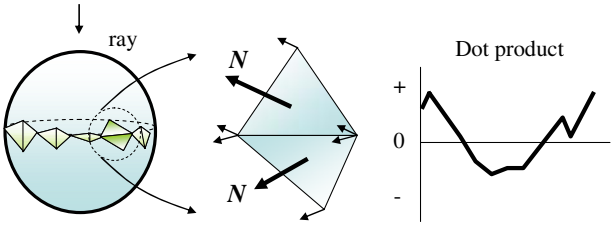


Fig. 7. Dot product result in polyhedron model

Thus, we decompose the polyhedral object by generating the triangular prisms which, are not boundary surface patches but the volumetric elements.

3.2 Visibility Test for Triangles

Triangular prisms are generated from the triangles in STL file. Since the criterion which, is verified to decompose the object is the visibility of triangles from a given ray direction. first, we check the visibility of triangles. The proposed procedure is as follows.

Procedure 1. Visibility test for triangles

```
Initializing the R(ray vector[0,0,1]) and BB(Bounding Box) of object;
 FOR(calculate the inner product(N•R) between R and the  normal vector(N)
    of each triangle Ti)
 {
    IF(0 < N•R < 1), THEN{ triangles are machinable triangles(MT)};
     ELSEIF(-1 < N•R < 0), THEN{triangles are non-machinable triangles(NMT)};
     ELSEIF(N•R = 0), THEN{triangles are the vertical triangles(VT)};
     ELSEIF(N•R = 1), THEN{triangle as the horizontal triangle(HT)};
  Save the list of 4 categories for triangles(MTlist, NMTlist, VTlist, HTlist);
}ra
```

We classified the triangles into the four types by using the procedure 1.

· MT(Machinable Triangle) : MT means a triangle is machinable by a milling machine.
· VT(Vertical Triangle) : VT means a triangle that composes a side wall(Vertical Wall).
· HT(Horizontal Triangle) : HT means a triangle put on support structure.
· NMT(Non-Machinable Triangle) : NMT means a triangle that needs the creation of support structure(Support structure) as triangles put on curved surface (overhang surface) that need support structure. Triangle classed by NMT is combined by Non-machinable Surface.

Each triangle is classified by our decomposition module in the initial stage.

In this research, the triangular prism is automatically generated by triangles properties. Thus we can easily build the object volume, the cutting volume and support volume from the triangular prisms.

The triangular prism is defined as the swept volume of the triangle in both direction, inside of object and outside of object. Each triangular prism is also defined as sub space which, is compose of the sub volume of object and is put around the object. Thus, the features like cavity or hole are automatically captured without geometric inference and are easily handled.

3.3 Triangular Prism Generation

The triangular prism is generated from the projection of the triangle which, is classified in the section 3.2. The generation of each triangular prism is similar to generate the dexel.

Each triangular prism is verified by three types as shown in fig. 8. These are the cutting triangular prism, object triangular prism and support triangular prism respectively. Cutting triangular prism is generated in outside direction in the object which, makes the result of the cutting volume and the support volume as shown in fig. 9.

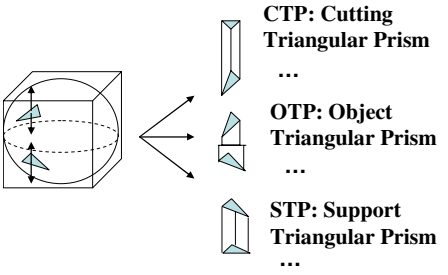


Fig. 8. Three types of triangular prism

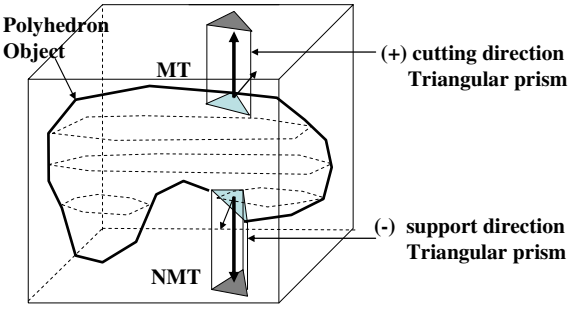


Fig. 9. Triangular Prisms for cutting volume and support volume

If the triangle is MT, this operation makes the cutting volume in cutting direction.. On the other hand, if the triangle is NMT, this operation makes the support volume. If the triangle that projected in inner object direction, the generative triangular prism is intersected with another triangular prism, named triangular prism. Arbitrary two triangular prisms are not allowed to occupy in the same space. The height of the triangular prism is determined to avoid the double occupation between the triangular prisms. This rule is adopted to generate overall triangular prism for the object. If this rule is not satisfied in a certain triangle, the triangular prisms are intersected each other. The proposed procedure is as follows.

Procedure 2 : Triangular prism's height calculation

```
Scan the triangles;
 Sort by z-coordinate of the centroid for each triangle;
 FOR( each triangles in MT/NMT list )
 {
IF ( [i]th triangle is adjacent
          with [i+1] triangle !== eol )
    CASE 1: edge-shared
      Save VSH: Vertex Share Height
    CASE 2: vertex-shared
      Save ESH: Edge Share Height
 };
```

Fig. 10 shows the procedures for the generation of the triangular prisms in inside object. Generally, adjacent two triangles are vertex share or edge share between them.

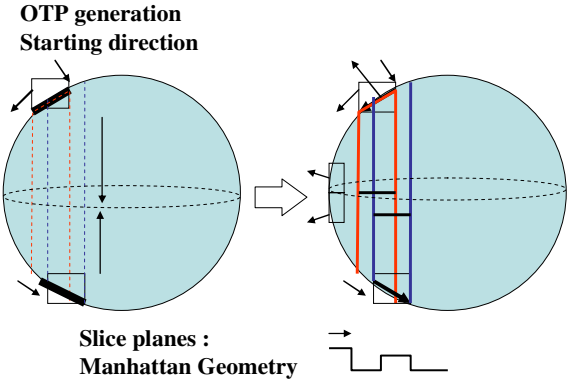


Fig. 10. Triangular Prism Generation.in inside Object

If the adjacent two triangles is the vertex share case as shown in Fig 11, we save the z coordinate of shared vertex(VSH: Vertex Share Height). If the adjacent two triangles is the edge share case, we save z coordinates of two shared vertices(ESH: Edge Share Height). If the hole or cavity is in the polyhedron, it is necessary to split the object. If the overlapping surface exists, the object also has to be splitted.

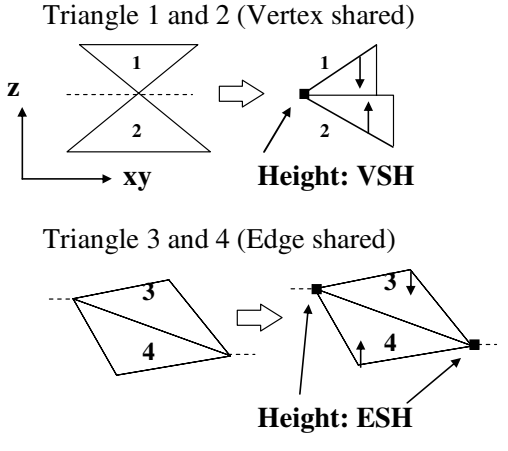


Fig. 11. The calculation of the VSH and ESH

4 Pattern of Triangular Prisms for Object Decomposition

The proposed algorithm makes the triangular prisms for object decomposition. By example, there is a spring model as shown in figure 12. In this model, if the ray hit the upper surface then the below surface is not hit by the ray. This area is verified as self-overlapping surface. If the model is composed of self-overlapping surfaces, the pattern of the model decomposition is more complex.

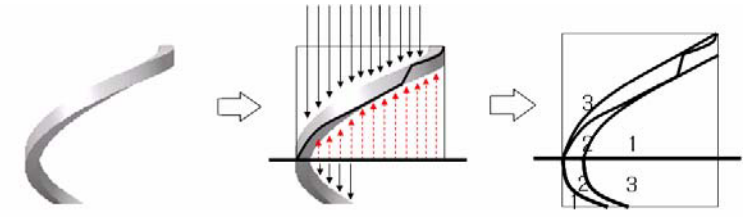


Fig. 12. Self-overlapping Surface Splitting.

The proposed algorithm can easily decompose this shape for additive works. In this case, our algorithm automatically decompose the object with self-overlapping surface into the support volume, cutting volume, object decomposition volume which, is composed of each triangular prisms. Fig. 13 shows the polyhedron and it's triangular prisms and 3D Layers.

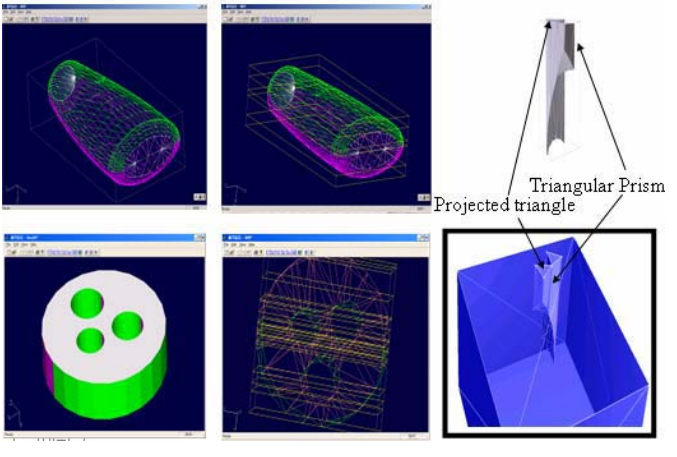


Fig. 13. The polyhedron, triangular prism and 3D layer

5 Conclusion and Future Works

In this research, the triangular prism generation algorithm was done to the polyhedral model. The proposed algorithm is working on not a native solid modeling system but

the STL file format, de facto standard in RP industry. For the purpose, we use the triangular prisms for polyhedron models which, are easily acquired reverse engineering system or the graphics modeler and reconstructed from the CT, MRI in medical field. Since, the triangular prism is based on the spatial decomposition concept, then the data structure is simple and the operations are compact. In contrast to the system based on the solid modeling, our algorithm does not suffer from operational failure in boolean operation. In our approach, we need an additive consideration for the object with many features to build a prototype with smooth surface. However, in this work, we don't consider the additive surface finishing issues. As a result, The triangular prism is regarded as the modified dexel then is useful to visualize for NC simulation.

In the future, we need to study the optimal decomposition direction for the triangular prism generation.

References

1. Andreas K. and Eduard G.: Real Time Simulation and Visualization of NC Milling Processes for Inhomogeneous Materials on Low-End Graphics Hardware, Proc. Computer Graphics International, (1998) 338-349.
2. Broek, J.J., Horvath, I., Smit, B. de, Lennings, A.F., Rusak, Z. and Vergeest, J.S.M.: Free-form thick layer object manufacturing technology for large-sized physical models, Automation in Construction, Vol. 11 No. 3, (2002) 335-347
3. Chang, Y.C., Pinilla, J.M., Kao, J.H., Dong, J., Rawaswami, K. and Prinz, F.B.: Automated Layer Decomposition for Additive/Subreactive Solid Freeform Fabrication, proc. Solid Freeform Fabrication Symposium, (1999) 111-120
4. Hope, R.L., Jacobs, P. A. and Roth, R, N.: Rapid Prototyping with Sloping Surfaces, Rapid Prototyping Journal, Vol. 3, No. 1, (1997) 12-19
5. 5. Horvath, I., Vergeest J.S,M., Broek, J.J., Rusak, Z. and Smit, B. de: Tool profile and tool path calculation for free-form thick-layered fabrication, Computer-Aided Design, Vol. 30, No. 14, (1998) 1097-1110
6. Jacobs, P. F.: Stereolithography and other RP&M Technologies form Rapid Prototyping to Rapid Tooling, ASME Press, 1996
7. Jamieson, R., and Hacker, H.: Direct slicing of CAD models for rapid prototyping, Rapid Prototyping Journal, Vol. 1, No. 2, (1995) 4-12
8. Kumar, V. and Dutta, D.: An assessment of data formats for layered manufacturing, Advances in Engineering Software, Vol. 28, No. 3, (1997) 151-164
9. McMains S. and Sequin, A.: Coherent Sweep Plane Slicer for Layered Manufacturing, Proceedings of the Fifth ACM Symposium on Solid Modeling and Applications, (1999) 285-295
10. Ramaswami, K. Yamaguchi, Y. and Prinz, F.P.: Spatial Partitioning of solids for Solid Freeform Fabrication, Proceedings of the Fourth ACM/SIGGRAPH Symposium on Solid Modeling and Applications, (1997) 346-353
11. Rock, S. J. and Wozny, M. J.: A Flexible File Format for Solid Freeform Fabrication, Solid Freeform Fabrication Symposium Proceedings, (1991) 1-12
12. Yang, Y., Loh, H. T., Fuh, J. Y. H. and Wong, Y. S.: Feature extraction and volume decomposition for orthogonal layered manufacturing, Computer-Aided Design, 35(11), (2003)1119-1128

Tweek: A Framework for Cross-Display Graphical User Interfaces

Patrick Hartling and Carolina Cruz-Neira

Iowa State University, Ames, IA 50011-2274, USA
patrick@vrac.iastate.edu, cruz@iastate.edu

Abstract. Developers of virtual environments (VEs) face an often-difficult problem: users must have some way to interact with the virtual world. The VE application designers must determine how to map available inputs to actions within the virtual world. However, manipulating large amounts of data, entering alphanumeric information, or performing abstract operations may not map well to current VE interaction methods, which are primarily spatial. Furthermore, many VE applications are derived from mature desktop applications that typically have a very rich user interface (UI). This paper presents Tweek, a reusable, extensible framework for UI construction that allows use of the same UI on a desktop system, on a hand-held computer, or in an immersive 3D space. Designers can maintain interaction consistency across conventional visualization settings such as desktop systems and multi-screen immersive systems. This paper covers in detail the design of Tweek and its use as an input device for virtual environments.

1 Introduction

Interaction within a virtual world is perhaps the most limiting factor for the development of effective virtual environment (VE) applications. Spatial interactions such as object selection and navigation can be mapped easily to positional and digital inputs. However, typical interactions found in most scientific and engineering applications (such as manipulating large amounts of data or entering alphanumeric information) do not map well to these types of inputs. These types of interactions can be represented more effectively using two-dimensional (2D) user interfaces found in desktop applications, and VE applications could benefit from utilizing proven 2D interface techniques.

The lack of a common, standard basis[1] for the creation of VE user interfaces presents another problem to VE developers, and current practices tend toward custom, per-application solutions. Such user interfaces usually correspond directly with the capabilities and limitations of the available input device(s), and as such, they often allow only very simple actions. To increase the number of supported actions, developers translate a combination of inputs to perform a single action. Furthermore, per-application VE user interfaces may become tightly coupled with a specific class of display systems or even with a specific immersive system installation. For example, a floating menu positioned

[1] This statement is in the context of VEs lacking something similar to the traditional WIMP paradigm used on 2D desktops.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1070–1079, 2005.

to appear on the left wall of a CAVE™ [1] will not be visible when the application is displayed on a Responsive Workbench™ [2] language.

With the wide range of VE technologies available today, there is an increasing need for applications that can migrate between different display configurations, but the scalability of the interaction methods can prevent this. Interaction methods that are tightly coupled with a specific display technology or VE system configuration tend to have poor scalability. VE applications should be allowed to run on different systems with different input devices to take advantage of the available resources or to use the interaction method that is most appropriate to the task at hand.

Hence, two needs have motivated the work presented in this paper: the need to support 2D interactions in VEs and the need for a flexible user interface that can be utilized in a variety of VE systems. We want to provide a reusable, extensible framework for user interface construction that will allow the same user interface to be employed on a desktop system or on a hand-held computer or in an immersive 3D space. These needs have led to the design of *Tweek* [3][4], a middleware framework for the development of cross-display user interfaces. Tweek combines several technologies to allow a 2D graphical user interface (GUI) to communicate with a remote application such that the two need not be running on the same computer.

We present the Tweek architecture, an implementation of a straightforward extension to a well-known software design pattern that provides the flexibility needed to design cross-display GUIs. We begin with a review of previous work on VE interfaces and our motivation to create Tweek. We follow with an in-depth design and technical detail description. We also present current VE applications based on Tweek and conclude with a discussion of results and future directions for our continued work.

2 Previous Work

JAIVE [5] and other studies [6] helped provide valuable insight into the usability of palmtop computers as input devices in projection-based systems. The JAIVE software used a Java-based GUI and a custom wire protocol for communicating with a remote VE application written in C++. The GUI provided a fixed set of interaction methods for use with an application, and the application could programmatically add and remove components from the GUI. However, the use of a non-standard wire protocol and a fixed set of interaction methods limited the adaptability of the tool. The Tweek implementation makes use of a standard cross-language communication protocol to allow for better scalability and re-usability.

The Virtual Tricorder [7] provides users with a uniform visual representation of a multi-purpose input tool. The visual representation corresponds directly to a physical device with buttons. Providing input in this way allows physical feedback through the button presses and visual feedback through the virtual representation of the tool. In the 3D space, the uses of the tool are unlimited, but without the visual device representation in the virtual world, there is no way to know what actions the Virtual Tricorder can perform. The 3D graphics provide the only information about the device functionality, and the user interface developers may have to provide many custom 3D GUI components in order to create the desired user interface.

Much has been done to provide toolkits for adding 3D GUIs to virtual worlds, both formally for general use [8][9][10] and informally for per-application needs. The common goal shared by these efforts is the introduction of familiar 2D GUI components (menus, scroll panes, sliders, etc.) into the 3D space. Depending on their intended use, however, some 3D GUIs may be tightly coupled with a specific combination of input and display devices. The Tweek design avoids this by separating the GUI from the immersive application, though we will explain how a Tweek-based GUI can be incorporated into the 3D space using the appropriate tools. Finally, each 3D GUI implementation must create the GUI components from scratch using some specific graphics software (OpenGL, Direct3D, etc.), and this can limit the re-usability. The Tweek design, on the other hand, allows GUI programmers to reuse existing 2D GUI software such as the Java Foundation Classes[2], wxWidgets[3], or Qt[4].

An exhibit was presented at the SIGGRAPH 2002 Emerging Technologies exhibition showing the ARS BOX[5] being controlled by the Palmist GUI[6]. The Palmist software runs on a Compaq iPAQ personal digital assistant (PDA) [11], and it is specialized for control of applications executing on projection-based VE devices. It is configured using the Extensible Markup Language (XML), thereby giving users a high-level language for customizing application controls without knowledge of programming interfaces. At a high level, the Palmist software shares some of the goals of Tweek; however, the Palmist focuses on interaction in multi-screen projection systems. A Tweek-based GUI can be used in a wide variety of immersive system configurations.

With Tweek, we are building on ideas from these earlier works to provide a well-defined, general-purpose framework for creating cross-display user interfaces based on 2D interaction techniques. To avoid imposing restrictions on how user interfaces are written, Tweek normally operates at the application level so that it may be used with any VE software tool. We have expanded on the idea of GUIs that can be manipulated by the remote application through the addition of new user- or application-defined GUI components at run time. This allows users to begin with an empty GUI frame that is extended on the fly based on parameters set forth by the virtual world authors.

3 Motivation

Through years of experience writing VR applications, we have identified a need for 2D GUIs to be available within immersive environments. Furthermore, the level of sophisticated interaction available on desktop visualization tools is challenging the acceptance of VEs as scientific and engineering tools. Our work with Tweek is therefore driven by two key issues: GUI migration and GUI technology reuse.

Many VE applications must be capable of executing in a variety of VE system configurations. At the computing level, the configuration can range from low-end PCs

[2] http://java.sun.com/products/jfc/
[3] http://www.wxwidgets.org/
[4] http://www.trolltech.com/products/qt/
[5] http://futurelab.aec.at/arsbox/
[6] http://www.aec.at/en/futurelab/palmist_sdk.asp

to high-end supercomputers. At the display level, we may have a single monoscopic screen, multiple stereoscopic screens, a head-mounted display (HMD), or many other variations. At the interaction level, there are input devices such as gloves and spatial trackers, and there is variation in the virtual space within which the devices operate. For example, in a CAVE™ or an HMD, interaction occurs in the virtual space surrounding the user. On a Responsive Workbench™, the interaction occurs directly in front of the user. If a GUI will be used with a VE application, it should be capable of adapting to different displays along with the application. Thus, cross-display GUIs become an important issue. Some previous work in this area has allowed the GUI to run only in a single computing environment [6][11][12].

Tweek addresses portability limitations by allowing the GUI to migrate from traditional desktop interaction to fully immersive 3D graphics. Moving beyond the desktop display configurations, the Tweek-based GUI can be brought into surround-screen projection systems through a palmtop computer or a PDA. Current wireless technology allows easy use of such computing devices in these settings. When a palmtop computer is not available, the same Tweek-based GUI can be incorporated into the virtual space using 3D windowing tools. This is because the interface between the user and the VE is separated physically from the VE system. Using a Tweek-based GUI in this way also allows it to be utilized with an HMD where a user's hands cannot be seen. In this case, a hand-held visual input device would be very difficult to use, but by projecting the 2D GUI into the 3D space, the same user interface is made accessible.

Current 2D GUI technology is very mature and very well understood. Previous research n incorporating 2D interaction into immersive 3D spaces has shown that some interactions are better suited to 2D GUIs than to 3D GUIs or 3D spatial interaction [13]. The reuse of existing 2D GUI software libraries has allowed us to focus on the interaction methods rather than on implementing a custom GUI library.

4 Design

Traditional GUI design contains three separate but related components: a *model* that contains data, a *view* of the model, and a *controller* that defines how the interface reacts to user input [14]. Formally, this is known as the Model/View/Controller (MVC) system. When the data held by the model changes upon a user request, the view is notified, and it updates to reflect the new state. In Tweek, we have employed this concept with a distinct, physical separation between the model and the view/controller components. We have designed a system where the data and its primary (immersive) view are displayed on one computer while the controller and secondary views are executed and displayed on another. The communication between the data and the controller happens via a network connection, as shown in Fig. 1. The result is a *distributed* MVC concept.

This physical separation between the data and the controller allows for some important capabilities. For example, given an application that maintains some data, the interaction with the data is not tied to the configuration of the display system. Multiple controllers and supporting views can be written, each using different GUI toolkits. Furthermore, a given controller can be rendered on different displays without changing the remote application it controls.

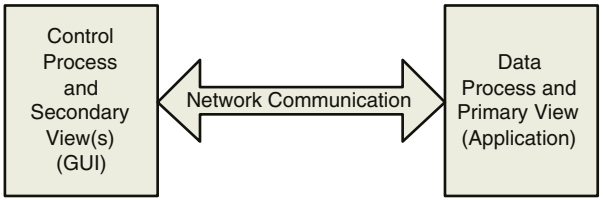


Fig. 1. High-level Tweek design

We have used rudimentary versions of the distributed MVC system in other VE-related projects. Tweek is the result of integrating these previous experiences with our goal of bringing 2D GUIs into 3D spaces. The fundamental design idea is that of the Observer design pattern [14], and it is implemented such that the *subject* is the remote application and the *observer* is the GUI. Thus, the GUI provides one or more views of the data, and widgets within the GUI act as controllers of state information maintained by the remote application. The state information is defined entirely by the application author. For example, in a VE application, the state could contain the user's position and orientation so that a map can be displayed in the GUI as a navigation aid.

Within an immersive projection system, users can bring a handheld computer into the immersive environment and utilize the 2D interaction in the same manner as they use conventional 2D interfaces in a desktop environment. The handheld computer provides a physical mechanism for displaying and interacting with the GUI. State changes made through GUI interaction are transmitted via a (wireless) network connection to the immersive 3D application, and the application responds accordingly. In the same manner, the application can communicate information back to the user through the GUI. All of this enables users to add more information about virtual objects; to see specifications of a selected object; or to make text annotations about objects using the handheld computer. To maintain the usability across multiple VE configurations where a handheld computer would not be available, we can embed the GUI in the immersive 3D space as one more graphical object of the scene.

5 Technical Details

Tweek combines a generalized, component-based framework for GUI creation and a back-end that allows a GUI to connect to remote objects using the Common Object Request Broker (CORBA) [15]. Through the use of CORBA, the application and its user interface can be written in different languages and can be executed on different computing platforms. This allows a physical decoupling of the VE application and its user interface. Java was chosen for the first Tweek-based GUI foundation we developed because we have made repeated use of Java for the creation of cross-platform GUIs that communicate with a variety of remote applications using a computer network. For immersive applications specifically, C++ is our typical choice for performance reasons. CORBA, then, facilitates the cross-language communication and provides an easily understood object-oriented approach to distributed programming. We have an early stage

version of a GUI written in Python that can communicate with the same remote C++ applications as the Java GUI without requiring modifications to the immersive C++ applications.

5.1 The Observer Pattern and CORBA

The Observer pattern defines a separation between the data associated with an object and a view/controller of that data. Within the pattern, there are two objects: the subject and the observer. The subject maintains the data; the observer provides a view of the data and may control the interaction with the data. A single subject may have multiple observers, each of which is notified when the subject state changes. When notified of changes, observers query their subjects to get the latest state information, as shown in Fig. 2. This is implemented using CORBA. As part of the foundational nature of Tweek, we define the two basic interfaces: `tweek::Subject` and `tweek::Observer`, shown in Fig. 3. Using these as the basis, it is up to the programmers of the applications and of the corresponding GUIs to define behavior specific to their applications.

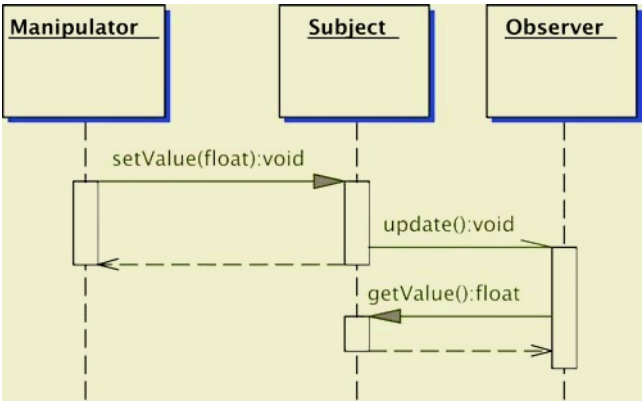


Fig. 2. Observer notification of subject state change

With the current Tweek framework, programmers define their own subject by extending `tweek::Subject` using the CORBA Interface Definition Language (IDL). Then, a C++ implementation of the interface is written that extends the Tweek subject implementation, `tweek::SubjectImpl`. In general, custom subjects define an API made up of accessor methods (so-called "getters" and "setters") that query and manipulate the subject state.

Each subject implementation must have a corresponding observer implementation. There is no need to create custom interfaces that extend `tweek::Observer` because the only method that will ever be invoked by remote code is `tweek::Observer::update()`. Custom observers will implement the `tweek::Observer` interface in the language chosen for the GUI components. The GUI is then built up around observer implementations and references to remote subjects that communicate with the observers.

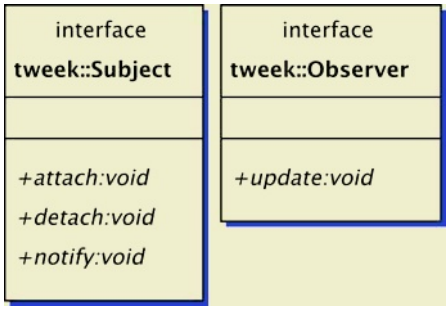


Fig. 3. Interfaces `tweek::Subject` and `tweek::Observer`

5.2 Java GUI

In order to offer users a highly flexible GUI as a starting point, Tweek includes a generic Java-based application framework that loads plug-ins dynamically to extend its functionality. These plug-ins may be discovered when the Tweek Java GUI is initialized; they may be loaded from disk after the GUI has already been activated; or they may be downloaded, or "pushed," from the remote application via CORBA. The plug-ins loaded by Tweek's Java GUI are implemented as JavaBeans™ [16], and they are described using XML files. The Tweek Java GUI design includes a simple class loader that searches for the XML description files, handles errors, and manages the run-time addition of new plug-ins.

5.3 Immersive User Interfaces

Incorporating the 2D GUI into 3D immersive graphics is the final aspect of cross-display GUIs. By providing this capability, users do not have to learn a separate user interface depending on the execution context. Furthermore, the application developer does not have to maintain separate user interfaces for the same VE application. The only difference is that interaction with a 2D GUI using traditional VE input devices requires careful manipulation because there is usually no 2D input surface with inherent force feedback. However, this will be the case for any GUI in the 3D space if no haptics system or interaction surface is available, so application authors may choose to provide visual or aural cues to aid the GUI user.

The Virtual Network Computing[7] (VNC) software, originally developed at AT&T Laboratories Cambridge, creates a 2D desktop that is managed by the VNC server and that is transmitted to VNC clients as a bitmap. We have implemented a VNC client as a stand-alone module that we call "VRJ VNC" that is distributed with the VE application development framework VR Juggler [17]. It is designed so that it can be added to any VR Juggler application object based on OpenGL. VRJ VNC allows any size desktop to be used, and the containing 3D "window" can be resized or moved interactively as necessary. We have also implemented a shade/unshade feature to allow for a form of

[7] http://www.realvnc.com/

window hiding. A snapshot of VRJ VNC in use can be seen in the following section in Fig. 4c. Because we have a full 2D desktop available, we can use VRJ VNC to incorporate *any* 2D application into the 3D space. VRJ VNC was inspired by the Garbo API of 3Dwm[9]. Using this tool, any 2D desktop application can be mapped onto a 3D shape in the virtual world with the basic interaction (point-and-click, drag-and-drop, etc.) working the same as on the desktop. The 3Dwm Garbo API also uses VNC.

6 Discussion

Tweek is Open Source software that can be found at http://www.vrjuggler.org/tweek/ on the World Wide Web. While Tweek can be used as a stand-alone GUI framework, it has been designed to be used in conjunction with VR Juggler. The VR Juggler framework allows applications to be portable across displays and physical devices. Tweek extends this portability by allowing the GUI associated with an application to move between VE system configurations.

One such VR Juggler application developed at our laboratory uses the Tweek Java GUI with a custom JavaBean™ running on a palmtop computer to provide extended input capabilities for a military training application. It includes controls that allow playback of recorded training data; six-degree-of-freedom navigation capabilities; detailed entity data readout for individual military units; and visual options for the units and the environment.

A more interaction-intensive application is our educational tornadic storm application. In the initial stages of the application design, we identified a set of key interactive functionalities expected by users. First, users must be able to request meteorological information such as variance in temperature, pressure, and humidity in the storm simulation. Second, the GUI must use maps to present information in a way that is immediately familiar to meteorologists. Third, the GUI must facilitate user understanding of their location within the storm. Finally, the GUI must allow the current storm data to be stored for later analysis in spreadsheet software. Beyond the specific details of interaction requirements, this application had to be designed to be used in a classroom by meteorology faculty, by students on their home computers, and in guided immersive learning activities. To avoid user confusion, the GUI had to stay the same for all of these

(a) Desktop (b) Tablet computer (c) Embedded 3D GUI

Fig. 4. Tornado simulation with Tweek interface

cases. In this context, Tweek became the ideal foundation for the development of the interactive capabilities of the tornadic storm. The design of Tweek met all of the project needs allowing us to create multiple display variations, ranging from a conventional desktop style display (Fig. 4a) to multi-screen immersive displays (Fig. 4b and Fig. 4c).

7 Conclusions and Future Work

User interaction has been a long-standing issue in VEs. Today, it is widely acknowledged that VEs provide many benefits over conventional desktop visualization. However, the limited capabilities of most VE interaction techniques have thus far restricted the widespread adoption of VEs as a "real" work tool in the daily routine of scientists, engineers, and other professionals. Tweek allows VE application authors to use familiar 2D GUI components in the development of their interaction methods. VE interaction designers can utilize proven 2D interaction techniques, thereby maintaining interaction consistency across different settings ranging from desktop systems to multi-screen immersive systems.

Although we have not performed any formal studies on the effectiveness of Tweek as general-purpose middleware for cross-display GUI creation, we have observed a continued and growing interest and use among the many projects in our research laboratory and in the VR Juggler user community. We feel that users recognize the value of Tweek and how it can provide a uniform interaction method for applications designed for multiple display configurations.

The current Tweek GUI implementation uses the well-known Java Swing libraries The use of Swing does impose some restrictions, however, due to compliant Java runtime availability and JVM memory use. For example, most current PDAs have limited memory (both volatile and persistent), and the mobile editions of Java tend to have a subset of the capabilities of workstation editions. By leveraging CORBA, we are exploring alternatives to Java. Most notably, we have used Python and PyQt[8] to create portable, high-performance, light-weight interfaces that can execute on PDAs with Qtopia[9]. We will pursue this avenue further to develop a second, more "PDA-friendly" foundational GUI for Tweek users.

Acknowledgments

We would like to thank Deere & Co., Procter & Gamble, and the United States Air Force for their on-going support of our research.

References

1. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A.: Surround-screen projection-based virtual reality: the design and implementation of the cave. In: Proceedings of the 20th annual conference on Computer graphics and interactive techniques, ACM Press (1993) 135–142

[8] http://www.riverbankcomputing.co.uk/pyqt/index.php
[9] http://www.trolltech.com/products/qtopia/

2. Krüger, W., Bohn, C.A., Fröhlich, B., Schüth, H., Strauss, W., Wesche, G.: The responsive workbench: A virtual work environment. Computer **28** (1995) 42–48
3. Hartling, P., Bierbaum, A., Cruz-Neira, C.: Tweek: Merging 2d and 3d interaction in immersive environments. In Callaos, N., Pisarchik, A., Ueda, M., eds.: Proceedings of the 6th World Multiconference on Systemics, Cybernetics, and Informatics. Volume VI., Orlando, Florida, United States (2002) 1–5
4. Hartling, P., Bierbaum, A., Cruz-Neira, C.: Virtual reality interfaces using Tweek. In: ACM SIGGRAPH 2002 Conference Abstracts and Applications, San Antonio, Texas, United States, ACM Press (2002) 278
5. Hill, L.C., Cruz-Neira, C.: Palmtop interaction methods for the immersive projection technology vr systems. In: Proceedings of the 4th International Immersive Projection Technology Workshop (IPT 2000), Ames, Iowa, United States (2000)
6. Watsen, K., Darken, R.P., Capps, M.V.: A handheld computer as an interaction device to a virtual environment. In: Proceedings of the 3rd International Immersive Projection Technology Workshop (IPT '99), Stuttgart, Germany (1999) 51–57
7. Wolka, M.M., Greenfield, E.: The virtual tricorder: A uniform interface to virtual reality. In: ACM Symposium on User Interface Software and Technology, Pittsburgh, Pennsylvania, United States (1995) 39–40
8. Bownman, D.A., Wingrave, C.A.: Design and evaluation of menu systems for immersive virtual environments. In: Proceedings of IEEE Virtual Reality 2001, Yokohama, Japan (2001)
9. Elmqvist, N.: 3dwm: Three-dimensional user interfaces using fast constructive solid geometry. Master's thesis, Chalmers University of Technology, Göteborg, Sweden (2001)
10. Heath, D.J.: Virtual user interface (vui): A windowing system for vr. In: Proceedings of the 2nd International Immersive Projection Technology Workshop (IPT '98), Ames, Iowa, United States (1998)
11. Hörtner, H., Lindinger, C., Praxmarer, R., Riedler, A.: Ars box. In: SIGGRAPH 2002 Emerging Technologies, San Antonio, Texas, United States (2002)
12. Tsang, M., Fiztmaurice, G.W., Kurtenbach, G., Khan, A., Buxton, B.: Boom chameleon: Simultaneous capture of 3d viewpoint, voice and gesture annotations on a spatially-aware display. In: (ACM) Symposium on User Interface Software and Technology, Paris, France (2002)
13. Hill, L.C.: Usability of 2D palmtop interaction device in immersive virtual environments. Master's thesis, Iowa State University, Ames, Iowa (2000)
14. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional Computing Series. Addison-Wesley Publishing Company, New York, NY (1995)
15. Object Management Group: The Common Object Request Broker Architecture: Core Specification. 3.0.2 edn. Object Management Group (2002)
16. Hamilton, G., ed.: JavaBeans^{TM}1.01 Specification. Sun Microsystems, Mountain View, CA (1997)
17. Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A., Cruz-Neira, C.: VR Juggler: A virtual platform for virtual reality application development. In: Proceedings of IEEE Virtual Reality, Yokohama, Japan (2001) 89–96

Surface Simplification with Semantic Features Using Texture and Curvature Maps

Soo-Kyun Kim[2], Jung Lee[2], Cheol-Su Lim[1], and Chang-Hun Kim[2,*]

[1] Dept. of Computer Engineering, SeoKyeong University,
{cslim}@skuniv.ac.kr
[2] Dept. of Computer Science and Engineering, Korea University,
{nicesk, airjung, chkim}@korea.ac.kr

Abstract. We propose a polygonal surface simplification algorithm that can preserve semantic features without user control. The semantic features of a model are important for human perception, which are insensitive to small geometric errors. Using an edge detects: Its three kinds of maps are employed to extract these features. First, an image map is generated boundary lines represent changes of chroma in the texture image by using edge detector. Second, the discrete curvatures at 3D vertices are mapped to the curvature map, and their data is also analyzed by an edge detector. Finally, a feature map is generated by combining the image and curvature maps. By finding areas of the 2D map that correspond to areas of the 3D model, semantic features can be preserved after simplification. We demonstrate this experimentally.

1 Introduction

Many simplification algorithms [15, 17] have recently been developed to transform a 3D polygonal model into a simpler version while preserving its original shape and appearance. Garland [7] and Hoppe [9] use an extended error metric that combines geometric and surface attributes such as normals, colors and texture coordinates. This extended method produces a more accurate and high quality mesh than geometric information alone. Hubeli [10] proposed a method to find feature edges in 3D meshes that gives weights to edges based on the difference between two normals and then fits polynomials to the lines of intersection between a parameter plane and a mesh. A drawback of this method is that users must select a threshold value that cannot be found intuitively. Besides, neither of them satisfies the time spent to extract feature points from the 3D mesh. Kim et al.[11] shows that discrete curvature can be a good criterion for a simplification that preserves the shape of an original model.

Although simplification methods produce plausible results in many cases, the semantic or high-level meanings of models are often lost at very low levels of detail. These features are largely a function of human perception. For example,

The corresponding author: Tel. +82-2-3290-3199; Fax. +82-2-953-0771.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1080–1088, 2005.

in the case of a human head, the regions such as nose, eyes, eyebrows, and the lip contact line on a mouth can be crucially affected by small geometric errors.

To address this problem, some previous authors [13, 16] have let the users selectively preserve semantic features during simplification. There are drawbacks to such methods: users must select some threshold value, that cannot be found intuitively and they have to search for the optimal threshold values that correspond to the semantic features of 3D polygonal models, Not only is it a very skillful and tedious take but it is also time consuming task. It threshold values are too high, unnecessary regions such as those that in include noise of high frequency would be included as a feature. If the values are too low, the semantic features are not detected.

We present a new simplification metric that uses both geometrical and iamge maps, obtained by range scanning, to preserve the semantic feature of models without user guidance. We employ three kinds of map to extract these features. First, an image map is generated using image processing techniques. This map is a set of boundary lines which represent the changes of chroma or grayscale. Second, the discrete curvatures of 3D vertices are computed and mapped to a curvature map, and this is analyzed by edge detectors. Finally, the FM is generated by superimposing the image map on the curvature map. So, for instance, while the image map does not help us detect to features on a nose of a head model (because the pixels on a nose are not distinguished from other pixels around it.), the curvature map allows us to find them. In place where the model does not provide texture information, the 2D curvature map is used. By finding how 2D line-type features of the model correspond to location on the 3D mesh, we can assign weights to the error metric during the simplification process. The following is a summary of our contributions:

- Enhancing a semantic feature detection for feature-preserving simplification by superimposing an image map on to a curvature map.
- Freeing uses from the difficult task of specifying semantic features by detecting them automatically.

2 Overview

Our feature-preserving simplification pipeline use begins with a 3D polygonal model that contains texture and geometry information. We can generate a multiresolution representation preserving semantic features without user control. To extract semantic features from a polygonal model, we use image and curvature maps to analyze semantic features such as texture information and change of curvature (Section 2.1 and 2.2). Two maps are generated by edge detection with boundary lines that represent changes of chroma or grayscale. We extract 3D feature areas on the mesh, which correspond to 2D feature lines in the feature map (Section 3), and compose them into line-type features (Section 4). Then we simplify a complicated polygonal model using an extension QSlim [7], to which we have added a weighting term to support feature-preserving simplification (Section 5).

2.1 Image Map

Our polygonal models are acquired from a range scanning system, which provides a texture image as well as geometry data. The texture image enables us to detect contour edges by image processing. Because the image data is represented in a regular fashion (a 2D array of [r,g,b] value), it is easy to handle and used in many algorithms.

From the texture image, we create a image map consisting of an $n \times n$ array of [r,g,b] values. To find the 2D feature lines on an image, we can apply image processing techniques [8] such as Canny edge detection, block binary and skeleton algorithm to this map. We have included the Canny algorithm in our implementation. The Canny edge detection finds more edges than the other methods. After using the block binary technique Initially the edges are found as thick lines by an appropriate segmentation of a grayscale image. We use a skeletonization algorithm to make thin lines.

2.2 Curvature Map

Many techniques [3, 4, 12] for computing differential information from discrete geometry are geared to the estimation of geometric processing such as discrete curvature, normal, dihedral angle and so on. We generate an alternative representation of a polygonal model. It is the form of 2D image, and is easy to handle. We show how to build a curvature map from input geometry data. We can obtain successfully curvature map using edge detector.

Discrete curvature is the approximate measure of curvature on discrete surfaces. Given a mesh, we can estimate the discrete Gaussian curvature K and the absolute discrete mean curvature $|H|$ at a vertex from the lengths of edges, dihedral angles, and areas of adjoining triangles. The sum of the absolute discrete principal curvatures $|\kappa_1|$ and $|\kappa_2|$ can be calculated from the K and $|H|$ curvatures. These curvatures at a vertex describe the shape of surface locally. Although we can easily find feature vertices from these discrete curvatures, it is hard to extract line features. It is possible to construct a piecewise linear geodesic curve that connects feature vertices, and call this a feature line; but we use of a 2D parametric domain to identify 3D features. First, an input model is parameterized on to a 2D rectangular domain We can generate a curvature map through two methods.

1. In the case of a model which has texture coordinates, we parameterize it using these coordinates.
2. In other case, the mesh is parameterized by a mapping function [6, 14] that puts a polygonal surface $\mathbf{S} \subset \mathbf{R}^3$ into one-to-one correspondence with a subset $\Omega \subset \mathbf{R}^2$.

Next, we color this parametric domain as an image of $n \times m$ pixels. The curvature value at a vertex is assigned to the gray level of the pixel to which it corresponds: the minimum and maximum curvatures in the polygonal model transform to 0 and 255 respectively. The color of the other pixels by interpolates

using a linear basis function [5]. Both the Gaussian and mean curvatures may be employed in constructing a curvature map, but Gaussian curvature did not perform well experimentally in searching for features and so we have mainly used the discrete mean curvature to generate curvature map. Finally, we apply image processing operators, such as Canny edge detection, block binary and skeletonization algorithms [8] to look for feature areas in the curvature map.

2.3 Feature Map

This section presents the way to combine iamge and curvature map which has advantage and disadvantage at the same time. An image map does not detect a nose because it can not distinguish of pixel. However, a curvature map is possible to detect the nose that computes the geometry processing. Therefore, we can select the weight function between the image and the curvature map. We add weight in image map to pixels of curvature map. Therefore, the border of FM is clear. The equation of the weight function map is as follows:

$$FM = \alpha \cdot image\ map + (1 - \alpha) \cdot curvature\ map \tag{1}$$

3 Mapping 2D Features to the 3D Model

We have identified sets of edges on the FM which will become semantic feature lines on the model, and now apply an appropriate inverse mapping function and transform these edges back in to 3D from the FM. Let the resolution of FM be $2^n \times 2^n$ and the coordinates of each pixel be $(s_i/2^n, t_j/2^n), i, j = 0, \ldots, 2^n - 1$. A vertex (x_i, y_i, z_i) is parameterized on a pixel (u_i, v_j) of FM.

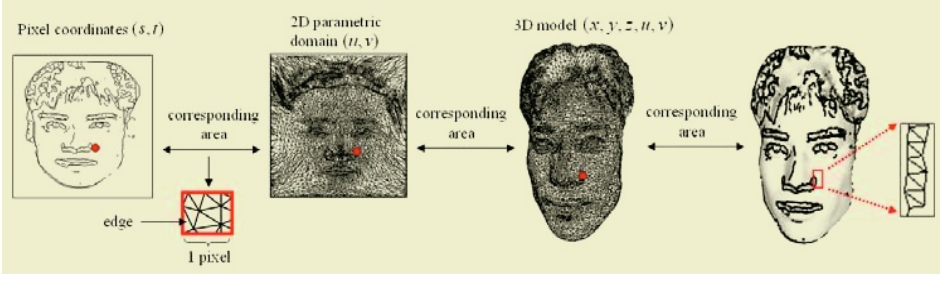


Fig. 1. Extraction of 3D feature areas from the Mr.Kim model: there are several edges on the 3D feature area corresponding to one pixel on the FM

Now, we will find 3D feature area which is the range of $(s_i/2^n, t_j/2^n) \leq (u_i, v_j) \leq (s_{i+1}/2^n, t_{j+1}/2^n)$ by parameterization function. However, a single pixel on the FM usually corresponds to more than one vertex on the 3D mesh (See Fig. 1). (The extent of this problem obviously depends on the resolution of the FM). So the inverse mapping operation actually produces a set of strips on the model, consisting of an over-large set of vertices and their connections. We need to reduce these strips to topologically correct piecewise-linear edges.

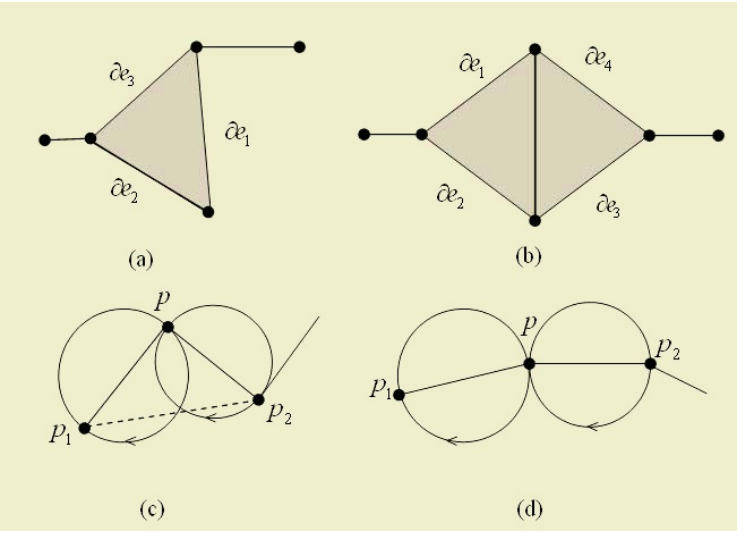


Fig. 2. Two cases of edges to be removed and smoothing a zigzag line strip: (a) case one, (b) case two, (c) intersection between sphere and, (d) not intersection between spheres

4 Thinning and Smoothing the 3D Feature Lines

We will thin out each strip by considering each of the edges that it contains as a candidate for the final feature edge, and eliminating candidates until a topologically valid edge has been created.

For each strip, we now consider all the edges which contribute to a single triangle, and classify this triangle into one of three cases, taking appropriate action in each case:

Case 1: If all the edges of the triangle contribute to that triangle only, then we look at the valency of the triangle's vertices. Any edges which terminate in a vertex of vanlency two are removed from the candidate set. Thus in Fig. 2a, edges ∂e_1 and ∂e_3 will be eliminated. If there are no vertices with a valency as low as 2, then one edge is chosen at random and removed from the set of candidates.

Case 2: If only two edges of the triangle are candidates for elimination, because the third edge contributes to another triangle, then we select one of the two edges which are available, at random, and eliminate it from the set of candidates. Sometimes two triangles of this case meet, as in Fig. 2b. removing one of the candidate edge from one triangle (e.g. ∂e_1 or ∂e_2 from the triangle on the left) will, of course, change the other triangle to Case 1.

Case 3: If there is only one candidate edge in a triangle, it can be straightforwardly eliminated from the candidate set.

Note that, in all cases, at least one edge is eliminated from the set of candidates. This process is repeated until no complete triangles are left in the edge 'strips'. However, there will usually be many hanging edges remaining, and possibly branches consisting of several feature edges linked end to end. These can all be eliminated by simple topological considerations. Now we have feature lines that are topologically correct, but jagged. We use a visibility graph method [2] to smooth them. Again, we consider every segment of the jagged feature edge as a candidate for the smoothed edge, but in this case we will amalgamate edges instead of eliminating them.

Suppose that p is a vertex shared by two candidate segments of the feature edge, labelled $\overline{pp_1}$ and $\overline{pp_2}$. We construct two spheres with $\overline{pp_1}$ and $\overline{pp_2}$ as diameters. If these two spheres intersect, we will remove both $\overline{pp_1}$, $\overline{pp_2}$ and the vertex p, and insert a new edge $\overline{p_1p_2}$, into the feature line. Otherwise, both edges will remain preserved in the set of candidates.

5 Feature-Preserving Simplification

Our method is based on QSlim [7], which uses the iterative edge contraction. To preserve feature edges, we have to change the simplification sequence. If they are assigned heavier simplification cost, they will be reduced later. Controlling the weights of quadrics enables us to put more cost on an edge.

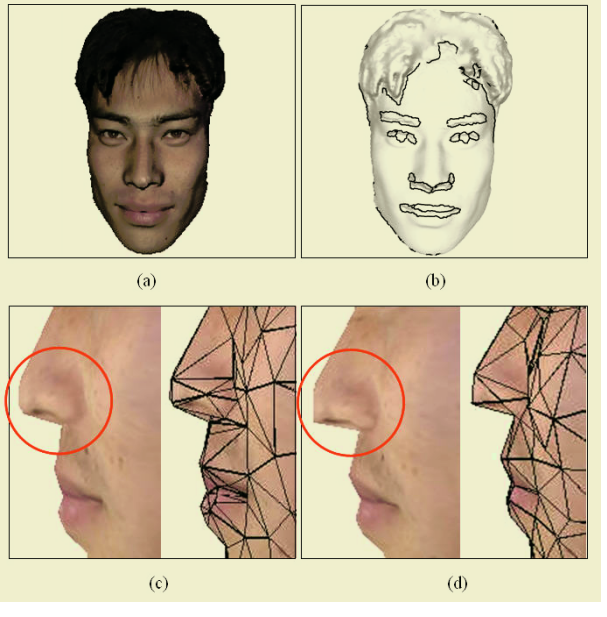


Fig. 3. Feature-preserving simplification of Mr.Kim model. (a) model as input (63,950faces), (b) line-type features using FM (α=0.5), (c) simplified model using our method (850faces), (d) simplified model using an extension QSlim (850faces)

First we choose one or more feature maps among image, curvature, and feature maps, and collect feature edges from it or them. Then the user-given weight δ is set on these feature edges and they give their end-vertices the weight. For each feature vertex, which is connected with one or more feature edges, the sum of the weight is finally assigned.

Therefore we can give more weight to corner vertices, which has three or more feature neighbor-edges. Moreover if a user chooses image map for computing weights, a simplification result mapped by well-preserved texture image will be acquired. A user can select the weight δ and change the computed contraction cost. If a vertex is a feature vertex, then we assign a heavy weight value. A weighted feature vertex v_i and its neighbor vertex v_j are contracted with associated quadrics error, the position of $\overline{\mathbf{v}}$ is selected to minimize $Q(\overline{v})$. For a given feature vertex $\mathbf{v}_i$, the equation 2 shows that the total quadric error metric $a\mathbf{e}$ is expressed as the linear combination of $Q(\mathbf{v}_i)$ and the neighbor vertices $Q(\mathbf{v}_j)$, where δ is the weight coefficient of neighbor vertex $\mathbf{v}_j$. After all weighted quadrics have higher edge contraction costs during mesh simplification. They are delayed order of contraction. Therefore, this equation prevents simplifying feature lines. Moreover, we must consider texture mapping during simplification because most complicated mesh has a texture image.

Our method can acquire good simplification mesh considered a texture image and preserved feature vertices. Before the total error metric $\mathbf{e}$ is calculated,

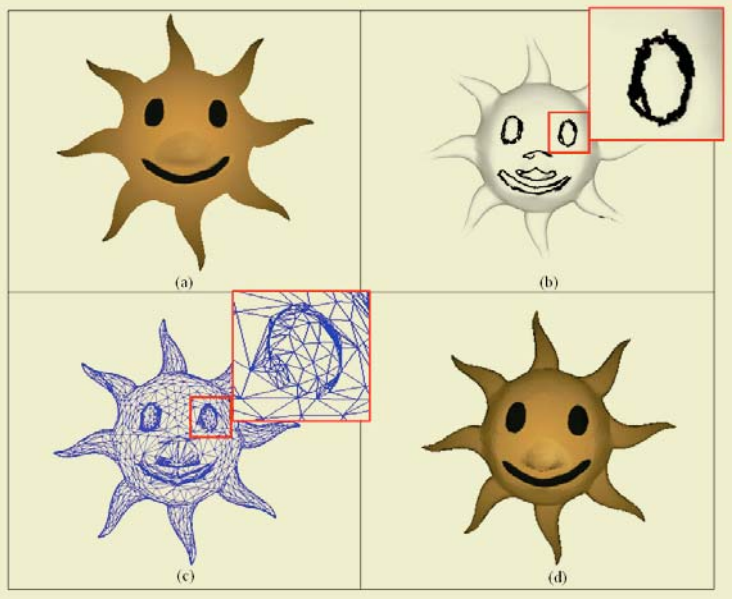


Fig. 4. Feature-preserving simplification of the sun model: (a) original model (65,000 faces), (b) extraction of line-type features, simplified model (3,000faces) (d) texture mapped simplified model (3,000faces)

the weight term of the texture attribute is multiplied by $Q(\mathbf{v}_j)$. The extension quardric error metric for $\mathbf{v} = (p) \in \mathbf{R}^5$ is:

$$\mathbf{Q}(\overline{v} \in \mathbf{R}^5) = \delta_i Q(v_i) + \delta_j Q(v_j) \tag{2}$$

6 Results and Implementation

We have tested our feature extraction approach using a Pentium Processor IV 3GHz PC with 1GB of main memory. Range/RGB complicated data were provided by Solutionix [1]. The resolutions of an image, curvature and feature map was 512×512, and the construction of the maps took a few milliseconds. In order to access the quality of our simplification method, we have created a number of simplified models and we have compared them to simplified models created using extension QSlim [7]. Figure 3 show illustration of Mr.Kim. Figure 3(b) show line-type features of the original model and exhibit distinct perception of semantic feature lines such as eyes, eye brows, nose and mouth in the models. Figure 3 (c) show model using our scheme which does not crumble a nose of a model. But, Figure 3 (d) shows that a nose of a model is crumbled by implementation of extension QSlim. Figure 4 shows feature-preserving simplification of a sun model.

7 Conclusion and Future work

In this paper, we have presented the algorithm for generation of a simplified model to maintain semantic features without user-intervene. We have presented to detect and extract line-type of semantic features using FM and exhibits a simplified model using an extension QSlim [7] that gives more weight to these features. We also showed that simplified models generated by using our scheme are effective by applying feature-preserving. In addition, this scheme can be used in mesh editing, morphing, facial animation, and so on.
Despite of the advantages, we can promote future studies as follows. Line features are detected by improved operators. It is necessary user to interact feature extraction from detected features.

Acknowledgments. This work was supported by University IT Research Center Project and the Basic Research Program of the Korea Science and Engineering Foundation (No. R01-2002-000-00512-0).

References

1. www.solutionix.com
2. de Berg M., van Kreveld M., Overmars M. and Schwarzkopf O.: Computational Geometry Algorithms and Applications VisMath'02, Berlin.

3. Desbrun M., Meyer M., Schroder P. and Barr A.: Discrete Differential Geometry Operators for Triangulated 2-manifolds. Springer, 2nd Edition.
4. Dyn N., Hormann K., Kim S.-J. and Levin D.: Optimizing 3D Triangulations Using Discrete Curvature Analysis. Mathematical Methods for Curves and Surface, pages 135-146, 2001.
5. Dyn N., Levin D. and Rippa S.: Numerical Procedures for Surface Fitting of Scattered Data by Radial Functions. In Proceedings of SIAM, pages 639-659, 1986.
6. Floater M.S.: Parameterization and smooth approximation of surface triangulations. Computer Aided Geometric Design, pages 231-250, 1997.
7. Garland M. and Heckbert P. S.: Simplifying surfaces with color and texture using quadric error metrics. In Proceedings of IEEE Visualization 98, pp.263-269
8. Gonzalez R. and Woods R.: Digital Image Processing, Addison Wesley, 1992, pp. 414 - 428.
9. Hoppe H.: New Quadric Metric for simplifying Meshes with Appearance Attributes. In Proceedings of IEEE Visualization, pp. 59-66, 1999.
10. Hubeli A. and Gross M.: Multiresolution Feature Extraction from Unstructured Meshes. In Proceedings of IEEE Visualization 01, 2001.
11. Kim S.-J., Kim S.-K. and Kim C.-H.: Discrete Differential Error Metric For Surface Simplification. In Proceedings of Pacific Graphics 2002, pp. 276-283, October 2002
12. Kim S.-J., Levin D. and Kim C.-H.: Surface Simplification Using Discrete Curvature Norm. Computers & Graphics, 2002.
13. Kho Y. and Garland M.: User-Guided Simplification. In Proceedings of ACM Symposium on Interactive 3D Graphics. 2003.
14. Levy B.: Constrained Texture Mapping for Polygonal Meshes. In Proceedings of SIGGRAPH, pp. 417-424. 2003.
15. Lindstrom P. and Turk G.: Fast and memory efficient polygonal simplification. In Proceedings of IEEE Visualization'98, pp.279-286, 1998.
16. Pojar E. and Schmalstieg D.: User-controlled creation of multiresolution meshes. In Proceedings of ACM Symposium on Interactive 3D Graphics, pp.127-130, 2003.
17. Rossignac J. and Borrel P.: Multi-resolution 3D approximations for rendering complex scenes. Modeling in Computer Graphics: Methods and Application, pp.455-465, 1993.

Development of a Machining Simulation System Using the Octree Algorithm

Y.H. Kim and S.L. Ko

Department of Mechanical Design and Production Engineering, Konkuk University,
1 Hwayang-dong, Gwangjin-gu, Seoul 143-701, Korea
{therisa, slko }@konkuk.ac.kr

Abstract. The overall goal of this thesis is to develop a new algorithm based on the octree model for geometric and mechanistic milling operation at the same time. To achieve a high level of accuracy, fast computation time and less memory consumption, the advanced octree model is suggested. By adopting the supersampling technique of computer graphics, the accuracy can be significantly improved at approximately equal computation time. The proposed algorithm can verify the NC machining process and estimate the material removal volume at the same time.

1 Introduction

As CNC machining center has been widely used in machining process, it is common thing that using CAD/CAM systems for NC program generation. However, during its generation process, NC program is prone to include errors such as inefficient tool paths and insufficient cutting conditions. The productivity and quality of machined parts highly depend on the NC program. Due to the need for unattended machining and higher productivity, it is becoming more critical to simulate the machining process more precisely to optimize cutting conditions, including the accurate calculation of chip load and the material removal volume. By simulating the machining process prior to the actual cutting operation, errors and inefficiency in the tool path can be corrected at the programming stage [1,2].

However, there are many CAD/CAM systems commercially availabel these days, none of them can produce error-free NC data. Many researches on NC machining simulation and verification had been undertaken, but most of them were limited to relatively simple simulation tasks. Constructive Solid Geometry (CSG) modeling system is the most popular method because it can complete the boolean operation of any 3D part model relatively easy and accurate[3]. The problem with the solid-modeling approach is that it is computationally expensive.

The Z-map model is a special version of cell-decomposition-modeling technique. It is the most widely used model for NC simulation and verification. Most commercial CAM systems are based on the Z-map method. Its advantages include simplicity of implementation, fast computation, approximate calculation of volume removal rates, easy comparison of a Z-map of the work piece and the desired part, and graphical display. Its computation time and memory consumption must be increased drastically, however, to enhance its accuracy [4,5,6].

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1089 – 1098, 2005.

The Voxel and the octree model are 3D decomposition method. Walstra, et al., had developed a simulation system based on the voxel model [7]. The Voxel has a simple data structure, which enables it to generate fast and update a part model. To increase the accuracy of cell decomposition in the model, however, the sizes of the cells have to be reduced. That means, in order to achieve high level of accuracy, it needs memory consumption and much computation time.

The octree model is an efficient representation tool for 3D objects. The major advantages of the octree model are its ability to represent complex objects with simple cubic cells. The octree model's spatial sorting property allows efficient model updating because the size and location of any octant can be retrieved directly from octree model's hierarchical data structure. Due to the nature of the decomposition method, however, the accuracy of the object representation is a major problem. Another limitation is the difficulty in incorporating the octree-based models into general geometric modeling systems such as CSG and B-Rep modeling methods. To address the limitations of the octree model, extended octree modeling methods have been reported in various literatures. Roy and Xu tried to apply the extended octree modeling technique to machining simulation [8]. The suggested algorithm has some defects, though. Since the octree model is a 3D decomposition methods, it is needed to search the cutting area using space partitioning. However, Roy's model search cutting area uses quadtree in an envelope projected on a xy plane. It is thus difficult to say when an octree structure was used for machining simulation. Moreover, the extended octree model has limitations due to its structural defects. As such, it can be applied only to simple part simulation.

In this thesis a new algorithm is proposed, referred to as the advanced octree model, for the prediction of the material removal rate from the NC program. The advanced octree model is based on the octree algorithm. For precise extraction the anti-aliasing theory is used. This theory is generally used with polygon mesh rendering in computer graphics. In this paper, the advanced octree class for the machining was designed and a machining simulator was developed using Visual C++. Comparison with other methods, i.e., Z-map model, the enhanced Z-map model [9], the octree model, the efficiency of suggested algorithm was verified. Considering the properties of decomposition modeling methods, it is expected that the advanced octree method can be applied to the complicated machining process, especially to five-axis machining.

2 The Advanced Octree Model

2.1 Concept of the Octree Model

Space-subdivision techniques are methods that consider the entire object space and in some way label each point in the space according to object occupancy. Space-subdivision techniques are based on a single cubic element, known as a voxel. A voxel is a volumetric element or primitive and is the smallest cube used in the representation. However, this is very costly in terms of memory consumption. To overcome the limitations of the voxel method, many representations replace the underlying regular space subdivision of the pure enumeration with a more efficient and adaptive subdivision. Prime examples of adaptive space subdivision schemes are the oc-

tree representation for solid objects [10,11]. The octree representation uses a recursive subdivision of the space of interest into 8 octants that are arranged into an 8-node tree (hence, the name). Fig. 1 shows the octree data structure. The octree model is a hierarchical data structure that describes how the objects in a scene are distributed throughout the three-dimensional space occupied by the scene. It has a very efficient data structure that stores 3D data in any form.

Properties of octree representations are similar to those of the exhaustive enumeration, with some noticeable exceptions. First, an octree can express all the objects approximately. Up to the limits of the resolution, all octrees unambiguously define a solid. On a fixed resolution, the representation is also unique. Moreover, it is easy to use for computation and application because of its structures.

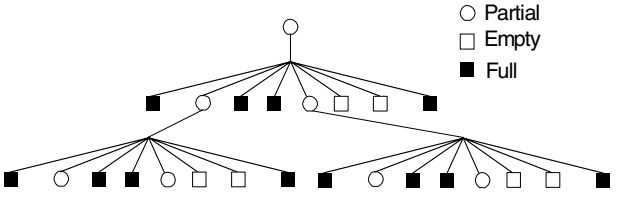


Fig. 1. An octree data structure

2.2 Supersampling Method

The point sampling system detect a polygon primitive at the center of each pixel. This can cause the familiar problem of a staircasing or saw teeth on digitized object boundaries and curves. To obtain properly anti-aliased results, the rendering process must take into account the areas of all the polygons that contribute to the shading of each pixel, rather than just a single sampling point. The supersampling method can be used to address this problem [12]. Samples of more than one point are taken in the region of each pixel and these samples are integrated to obtain the final pixel-shading value. With the multiple steps of intensity in pixel representation, jagged edges are almost diminished.

The octree model is a point sampling method. Therefore, there is same limitations - to get high resolutions, computation time and memory usage increase excessively. To solve this problem, we applied anti-aliasing methods to a simple decomposition model, the Z-map model from previous research [9]. Moreover, these techniques were applied to the octree algorithm.

Many supersampling techniques have been developed to raise the sampling rate and increase the resolution of sampling operations [13]. In this research, regular supersampling method were used, which uses sampling points that are uniformly distributed throughout the sample space.

Although supersampling method can improve the resolution without increasing of memory consumption, it requires heavy computation time in oversampling process. Basically, there is no difference in the computation time between 4x4 points supersampling method and 4x4 pixels point-sampling method. To reduce unwanted overhead in the oversampling process at each pixel, the suggested algorithm detects the boundary grid where the edge of the tool envelope is crossing, as shown in Fig. 2(b).

Supersampling is applied to detected boundary pixels, and conventional point sampling is used inside the tool envelope, as shown in Fig. 2(c). By reducing required oversampling process as Fig. 2, the computation time is 3 ~ 4 times faster than the conventional supersampling method as Fig. 2(a).

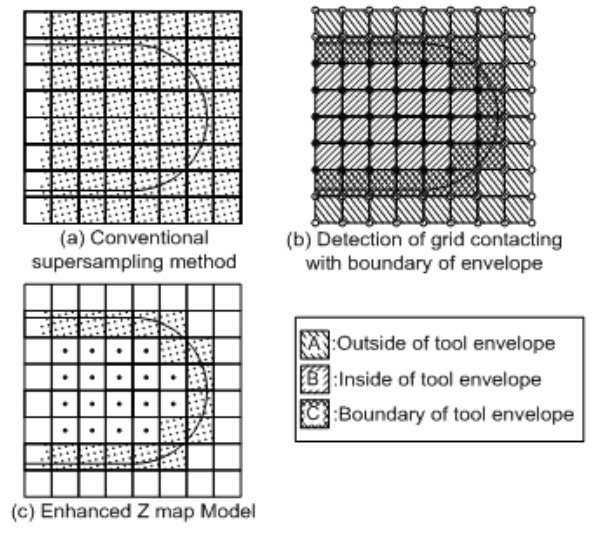


Fig. 2. Adopting supersampling algorithm into the tool envelope boundary

2.3 The Advanced Octree Class

Due to its structural advantages, the octree technique is popular in computer graphics. However, many open classes are insufficient for machining simulation, due to which an octree class for machining simulation is hereby developed.

An advanced octree is a tree-structure that contains data. Every octant has a cube origin, size and node types. The conventional octree node is classified into three types: black, white, intermediate. If the node is completely occupied by the object, it is a full cube (black). Empty cubes (white) are fully outside the object or are entirely free of the object. Partial cubes (intermediate) are partially filled by the object. In addition, the advanced octree model has variables that hold the numbers of the detected sampling points for the application of the supersampling method. This node is classified into gray node. And every octant also contains pointers to 8 children nodes and a pointer to its parent node. To construct an octree, the classification procedure is used in a recursive fashion.

Fig. 3 shows the flowchart that represents the creation of the tree structure in the simulation module. Generating the first root octree, the program starts and reads the NC code. If the tool path does not match the first tool path, it makes a new octree. For

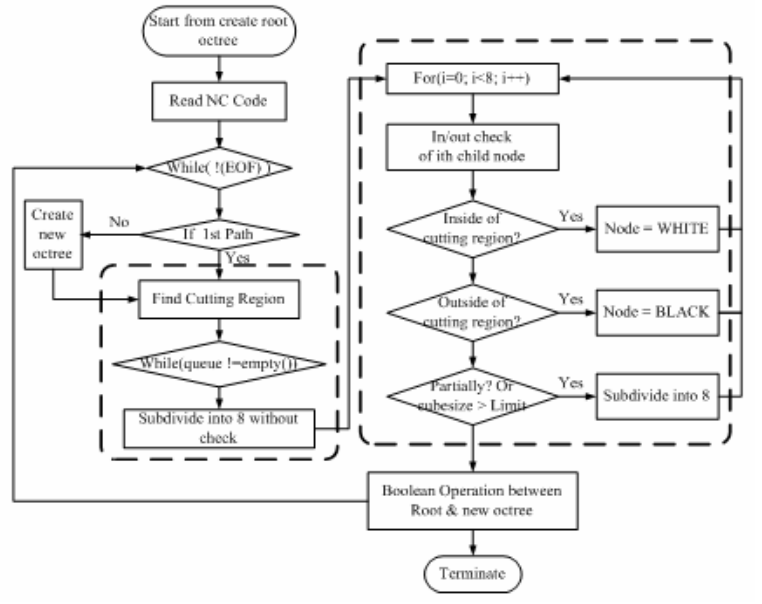


Fig. 3. Flowchart for the machining simulation

the 8 nodes in the subregion falling under each case, the program inspects the interior or exterior of the cutting area on these nodes. When the tool cuts a node completely, the node becomes white, and when the tool does not cut a node, it will be black. When a node is partially cut, if its size is larger than that of the limited cube node (intermediate), the node will be divided into 8 smaller nodes and will undergo the same inspection process.

After every division process is finished, Boolean operations will be performed between two octree models. In case the advanced octree model is applied to the supersampling algorithm, after the division process, Boolean operations will be performed only after supersampling for the boundaries.

3 Application

3.1 Verification of the Octree Algorithm

In order to verify the effectiveness of the advanced octree algorithm in predicting the material removal volume, we performed a machining simulation. The techniques presented in this paper have been implemented on Visual C++. The developmentprogram was used to extract the tool envelope from the NC code, and the recursive subdivision process was accomplished. With this program, we could not only predict the material removal volume but we could also simulate the machining process with a flat-end mill and a ball-end mill.

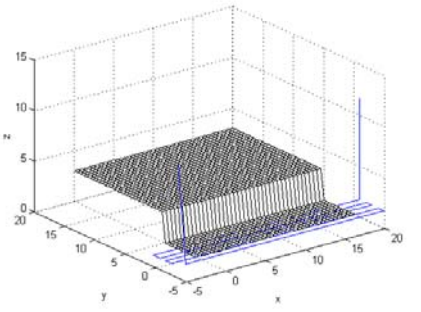


Fig. 4. Tool path for the verification fo the developed model at pick feed is 1mm

To verify the advanced octree algorithm, simulations were accomplished in various conditions–i.e., with various types and sizes of tools and pick feeds. Fig. 4 shows an example of tool path for the verification with the parallel pick feed using Φ6 flat end mill. To inspect its estimate of the material removal volume per unit length (material removal rate) with the cell sizes, four cases were used: 1 mm, 0.5 mm, 0.3 mm and 0.1 mm.The octree model is a 3D decomposition method. Therefore, it has to be compared with a 2D method. As such, simulations were performed with the Z-map model and the enhanced Z-map model, which applied supersampling methods to the Z-map model.

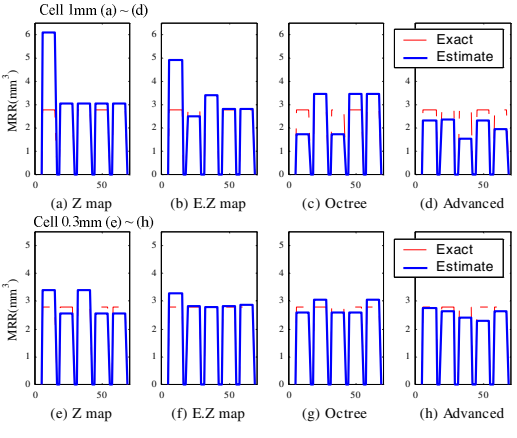


Fig. 5. Simulated material removal volume according to the unit length of a tool movement at a pick feed of 1mm

Fig. 5 shows the simulated material removal rate. The exact solution acquired through the analytical method is plotted as a dotted line in each graph. At the cell size of 1 mm, the Z-map models overestimated, especially in the first path [Fig. 5(a)]. In the first path, the enhanced Z-map model overestimated, but not as significantly as did the Z-map model [Fig. 5(b)]. The octree model [Fig. 5(c)] deviations varied. The

advanced octree model showed especially good prediction results [Fig. 5(d)]. As the cell size decreased, the accuracy of the simulated material removal volume tended to increase in each algorithm [Fig 5(e~h)]. Fig. 6 represents the error deviations from the simulation results. The advanced octree model showed remarkably little error deviation at each cell sizes [Fig. 6(d)]. Fig. 7 represents the relationships between the computation time and the accuracy of the simulated results using the octree and the advanced octree model. The advanced octree model showed good prediction results at the 1 ~ 0.3 mm cell sizes. At the 0.1 mm cell size, however, the computation time was longer than that using the octree model.

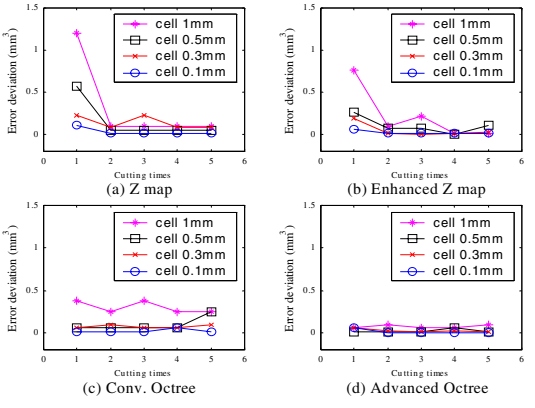


Fig. 6. Error deviations from simulation results with various algorithms

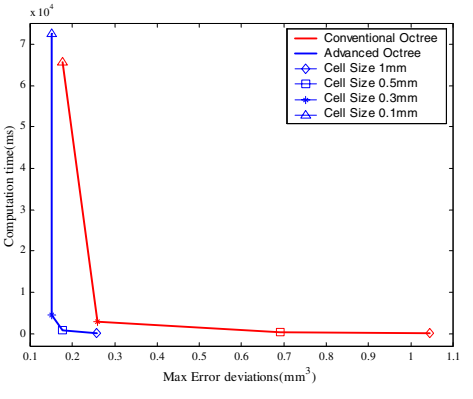


Fig. 7. Relationships between computation time and accuracy of the simulation

3.2 Optimized Cell Size for the Machining Simulation and Its Application Range

To find the optimum cell size, simulations were performed for various cell sizes in radial depths of 0.5 mm, 1 mm and 2 mm. Table 1 shows the simulation results. The

material removal rate predicted by the advanced octree model had an error of about 7.5% when the radial depth and the cell size were the same. On the other hand, the octree model had a 30% error and twice the simulation time of the advanced octree model. To obtain almost the same accuracy in the octree model as that in the advanced octree model, the cell size has to be minimized to three-tenths of the radial depth. In this case, the computation time will be 8.6~92 times longer than in the former condition. Thus, the results can be obtained, including a 7.5% error, when the cell size and the radial size are the same.

To determine the proper application range of the advanced octree method, we considered the relationship of the cell size, computation time and errors. Fig. 8 represents these relations when the radial depth 2 mm. With this results, it is proved that advanced octree model is superior to any other models at cell size 1 ~ 0.3 mm.

Table 2 shows accuracies in each model at the same cell size. The results show that the advanced octree model has an error of about 7.5% when the cell size is 1 mm, which is usually used in commercial CAM system. In this case the Z-map has 32% error. If we decreased the cell size into 0.5 mm, the advanced octree model has 2.4% error. The other hand, the Z-map still has large error percentage-15%.

Table 1. Simulation results

Radial Depth 2mm								
		6mm	4mm	**2mm**	1mm	**0.5mm**	0.3mm	0.1mm
Advanced Octree	Time	15	15	**78**	281	1125	5968	147156
	Error	1.664	0.556	**0.4168**	0.1384	0.104	0.0694	0.0694
	Error (%)	29.9	10.0	**7.5**	2.5	1.9	1.2	1.2
Octree	Time			31	140	**672**	4328	127031
	Error			1.664	0.556	**0.4168**	0.1384	0.104
	Error (%)			29.9	10.0	**7.5**	2.5	1.9
Radial Depth 1mm								
		4mm	2mm	**1mm**	0.5mm	**0.3mm**	0.1mm	
Advanced Octree	Time	32	64	**250**	984	4547	72563	
	Error	0.8328	0.279	**0.209**	0.068	0.0732	0.0404	
	Error (%)	30.0	10.0	**7.5**	2.4	2.6	1.5	
Octree	Time			109	500	**2860**	65563	
	Error			0.8328	0.2792	**0.2089**	0.068	
	Error (%)			30.0	10.0	**7.5**	2.4	
Radial Depth 0.5mm								
		1.5mm	1mm	**0.5mm**	0.3mm	**0.1mm**		
Advanced Octree	Time	93	469	**1890**	8562	157219		
	Error	0.4144	0.1456	**0.1064**	0.0392	0.035		
	Error (%)	29.6	10.4	**7.6**	2.8	2.5		
Octree	Time		204	953	5579	**173875**		
	Error		0.5488	0.4144	0.1456	**0.1064**		
	Error (%)		39.2	29.6	10.4	**7.6**		

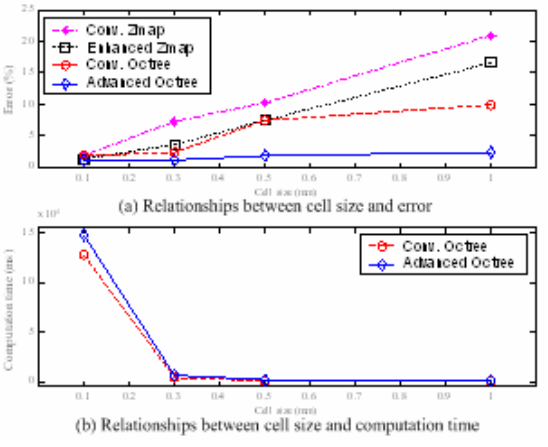


Fig. 8. Relationship between cell size and the prediction efficiency at a pick feed of 2 mm

Table 2. Comparison of the accuracies between four types models

Pick feed 1mm					
		1mm	***0.5mm***	0.3mm	0.1mm
Advanced Octree	Error Dev.(mm^3)	**0.2088**	***0.068***	0.0732	0.0404
	Error(%)	**7.5**	***2.4***	2.4	1.5
Conv. Octree	Error Dev.(mm^3)	0.8328	0.2792	0.2089	0.068
	Error(%)	30	10.0	7.5	2.4
Z map	Error Dev.(mm^3)	0.8867	0.4284	0.386	0.0817
	Error(%)	31.9	15.4	13.9	2.9
Enhanced Z map	Error Dev.(mm^3)	0.6209	0.2848	0.141	0.0543
	Error(%)	22.3	10.2	5.1	2.0

4 Conclusion

The objective of this thesis is to develop a machining simulation system. To achieve a high level of accuracy, fast computation time and less memory consumption, the advanced octree model is suggested. The advanced octree model is based on the octree method.

In the commercial CAM system, the Z-map model, which is a 2D decomposition method, is commonly used due to its advantages. However, this method is difficult to apply in complicated part models. Furthermore, its results are not very accurate. Octree, a 3D decomposition method, can overcome the limitations of the Z-map model. Moreover, the octree algorithm recursively subdivides operational processes only at their boundaries, thus reducing computation time and memory wastage.

To achieve a high level of accuracy using an octree model, the anti-aliasing theory is applied as a form of the supersmpling algorithm in each cubic cell. The supersampling technique in CG applies the entire space of interest. In this model, it applied

boundaries only to the tool envelope. Therefore, this model can achieve a higher level of accuracy with the same memory consumption and in relatively less computation time.

To verify the effectiveness of the developed simulation system, four types of models were compared in various conditions: the Z-map model, the enhanced Z-map model, the octree model, and the advanced octree model. The simulation results show that the advanced octree model is superior in large cell sizes. To apply the optimized cell size while considering the cutting conditions, various simulations were attempted. It was found that with the same cell size as the radial depth, reasonable simulation results could be generated.

Although, the Z-map model has many advantages, it also has several limitations due to the nature of its data structure. Although many researches have already been conducted to overcome these limitations, none of them could solve these problems. It has been verified that the suggested method, i.e., the advanced octree model, can efficiently overcome these problems.

References

1. Tsai, M.D., Takata, M., Inui, M, Kimura, F., Sata, T., 1991, "Operation Planning based on Cutting Process Models", Annals of the CIRP, 40/1:95
2. Schulz, H., Bimschas, K., 1993, "Optimization of Precision Machining by Simulation of the Cutting Process", Annals of the CIRP, 42/1:55
3. Wang, W.P., Wang, K.K., 1986, "Geometric Modeling for Swept Volume of Moving Solids", IEEE CG&A, 6:8-17
4. Anderson, R.O., 1978, "Detecting and Eliminating Collisions in NC Machining", Computer Aided Design, 10/4:231-237
5. Van Hook, T. Aug. 1986. Real-time Shaded NC Milling Display. In Computer Graphics (Proc. SIGGRAPH), 20(4), pp. 255-268.
6. Hsu, P.L. and Yang, W.T. 1993. Real-time 3D Simulation of 3-axis Milling Using Isometric Projection. Vol. 25, No. 4, pp. 215-224.
7. Walstra, W.H.; Bronsvoort, W.F.; and Vergeest, J.S.M. 1994. Interactive Simulation of Robot Milling for Rapid Shape Prototyping. In Computer & Graphics, Vol. 18, No. 6, pp. 861-871.
8. Roy, U., Xu, Y., 1999, "Computation of a Geometric Model of a Machined Part from its NC Machining Programs", Computer Aided Design, 31:401-411
9. Kang, M., Lee, S. L., and Ko, S.L., 2002, "Optimization of Cutting Conditions Using Enhanced Z map Model", Annals of the CIRP 51/1:429-432
10. Jackins, C. L., and Tanimoto, S. L., "Octree and their use in representing three-dimensional objects", CG and Image Processing, 14:249-270
11. Meagher, D., 1982, "Geometric Modeling Using Octree Encoding," CG and Image Processing, 19:129-147
12. Haeberli, P., Akeley, K., 1990, "The Accumulation Buffer: Hardware Support for High-Quality Rendering", Proc. of Siggraph, 24/4:309-317
13. Foley et al, 1996, Computer Graphics: Principles and Practice, Addison-Wesley
14. Choi and Jerard. Sculptured Surface Machining.

A Spherical Point Location Algorithm Based on Barycentric Coordinates

Yong Wu, Yuanjun He, and Haishan Tian

Department of Computer Science, Shanghai Jiao Tong University, China
wuyong916@sjtu.edu.cn

Abstract. An algorithm based on barycentric coordinates is presented to solve the point location problem in spherical triangulation meshes. During the preprocessing stage, a subdivision connectivity mesh is constructed to partition the spherical domain into some subdivision regions. Then we find a representation triangle from the triangle set of original spherical mesh for each subdivision region. During the locating stage, we firstly find the subdivision region containing the query point p and select the corresponding representation triangle as the start one to walk. Then the barycentric coordinates are used to extract local heuristic information about the location of p, so as to find the shortest path from the start triangle to the target one. In comparison with traditional algorithms, our approach has better time-space performance.

1 Introduction

The planar point location problem is one of the basic problems in computational geometry. It is encountered in many fields such as computer graphics, CAD and GIS. Given a planar polygonal mesh M and any query point p, point location means finding the polygon containing p.

In recent years, spherical parametrization based 3D geometry applications have been focused on by many scholars [4], [5], whose basic idea is to transform original irregular meshes into regular ones by spherical remeshing methods [6], [7]. Figure 1 shows the spherical remeshing process. Figure 1-(a) is the spatial Bunny model, which is first parameterized onto the unit sphere (b). Figure 1-(c) shows the regular spherical mesh with subdivision connectivity, which is used to sample (a). Figure 1-(d) is the resulting remesh. During the sampling process, for each vertex p of (c), we find its containing triangle $f(p_0, p_1, p_2)$ in (b) and compute its sampling value according to the counterparts of p_0, p_1 and p_2 in (a). Here how to find the triangle containing p is a spherical point location problem. Many effective methods [1], [2], [3] have been presented to solve the planar point location, but few attentions are paid on the spherical one.

The simplest point location method is the scanning algorithm, which examines triangles of M one by one until the target one containing the query point is found. Suppose M has n_f triangles. Then the expected time complexity of the scanning algorithm is $O(n_f)$.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1099–1108, 2005.

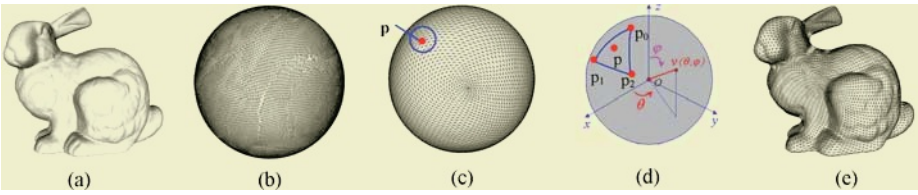


Fig. 1. Spherical remeshing process of Bunny model. (a)Spatial Bunny; (b)Spherical Bunny; (c)Regular spherical mesh; (d)Spherical sampling; (e)Remesh of Bunny

The first efficient algorithm for the point location problem was proposed by Dobkin and Lipton [1]. They divided the plane into some small slabs, which were sorted from left to right and from top to bottom. The time complexity of single point location operation is only $O(\log n_f)$. But $O\left(n_f^2\right)$ space and $O\left(n_f^2 \log n_f\right)$ time have to be spent on preprocessing stage to build the special data structure. So given n_v query points, $O\left(n_f^2\right)$ space and $O\left(n_f^2 \log n_f + n_v \log n_f\right)$ time are required to locate them. Kirkpatrick proposed a triangulation refinement based algorithm to solve the point location optimally [2]. The location time for a given point is also $O(\log n_f)$. But the space and time complexity of preprocessing operation are only $O(n_f)$ and $O(n_f \log n_f)$. So given n_v query points, they can be located in $O(n_f)$ space and $O(n_f \log n_f + n_v \log n_f)$ time. Obviously, this algorithm has better performance than the slab algorithm.

In order to avoid preprocessing and additional storage, Sundareswara and Schrater developed a point location algorithm based on barycentric coordinates [3]. The algorithm can locate a query point in $O\left(n_f^{1/2}\right)$. So given n_v query points, the whole search time is $O\left(n_v \cdot n_f^{1/2}\right)$. The main drawback of this method is that it doesn't discuss how to find a good start triangle to walk.

In this paper we generalize Sundareswara's algorithm from the planar domain to the spherical one. We also present a valid strategy to find a good start triangle to walk.

2 Some Basic Notations

In this paper we concentrate on the set of 2-manifold oriented spherical triangulation meshes. For spherical polygonal meshes, they should be triangulated in advance.

The spherical mesh can be denoted by a pair $C = (P_C, K_C)$, where $P_C = \{p_1, \cdots, p_n\}$ (n represents the number of vertices) is the set of vertex positions defining the shape of mesh in the unit sphere, and K_C is a simplicial complex specifying the connectivity of mesh simplices (the adjacency of vertices, edges and faces). Each element of P_C is represented by a pair of radian values $p_i = (\theta_i, \varphi_i)$. K_C contains three types: vertices $v_i \in V$, edges $e_{i,j} = \{v_i, v_j\} \in E$ and faces $f_{i,j,k} = \{v_i, v_j, v_k\} \in F$. Each element of V has indexes to its neighbor

edges. Each element of E has indexes to its neighbor triangles. And each element of F has indexes to its three vertices. As a result, from any element of K_C, its neighbor can be visited in $O(1)$ time.

3 Outline of Algorithm

Our algorithm contains two stages:

- **Preprocessing stage.** We partition the spherical domain into some subdivision regions and find a representation triangle from the original spherical mesh C for each region.
- **Locating stage.** We locate the query point p using the following strategy:
  - Step 1. Select an appropriate start triangle to walk from the triangle set F of C according to the spherical position of p.
  - Step 2. Compute the barycentric coordinates (α, β, γ) of p with respect to current visited triangle.
  - Step 3. Decide the next triangle to walk according to (α, β, γ).
  - Step 4. Step 2 and Step 3 are executed iteratively until all the barycentric coordinates are nonnegative. Then the last visited triangle is the target one containing p.

4 Preprocessing

During this stage, we first construct a spherical mesh S with subdivision connectivity, which partitions the spherical domain into some subdivision regions. Then we compute the centroid point q of each subdivision region f_S and use the barycentric coordinates approach, which will be described in subsection 5.2, to find the triangle f_C containing q from the triangle set F of the original spherical mesh C. Since f_C lies in the region decided by f_S, we call it the representation triangle of f_S.

4.1 Constructing the Spherical Mesh with Subdivision Connectivity

The aim of constructing spherical subdivision connectivity mesh S is to partition the spherical domain into some subdivision regions. There are many methods

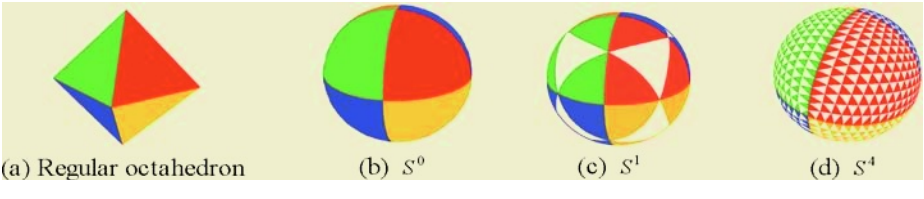


Fig. 2. Spherical subdivision connectivity meshes at several different levels

[8] of constructing S. In order to obtain a spherical mesh with uniform triangle distribution, we start from a regular convex octahedron, whose vertices lie on the unit sphere. After replacing each edge of the octahedron by the corresponding great circle arc, we obtain the spherical base mesh S^0, which is shown in Fig 2-(b). After that, a 1-to-4-subdivision operation is iterated on S^0 to produce a serial of spherical subdivision connectivity meshes $S^0, \cdots, S^i, \cdots, S^l$. The relationship between the subdivision level i and n_f^i, which is the number of triangles in S^i, can be expressed by $n_f^i = 8 \times 4^i$.

4.2 Finding a Representation Triangle for Each Subdivision Triangle

After constructing the spherical subdivision connectivity mesh S^l, we find a representation triangle f_C^l from the triangle set F of original spherical mesh C for each triangle f_{S^l} of S^l. During locating stage, if the query point p lies in f_{S^l}, then its representation triangle f_C^l is selected as the start one to walk.

As $S^0, \cdots, S^i, \cdots, S^l$ is a sequence of subdivision connectivity meshes, we use a hierarchy refinement method to find the representation triangles. Consider the middle mesh S^i. For one triangle f_{S^i} of S^i, we compute its centroid q^i firstly. After selecting the representation triangle f_C^{i-1} of $f_{S^{i-1}}$, which lies in S^{i-1} and is the father of f_{S^i}, as the start one to walk, we use the barycentric coordinates approach of subsection 5.2 to find the representation triangle f_C^i, which lies in C and contains q^i, for f_{S^i}. Above searching strategy of representation triangle is iterated until S^l is reached. Then we can obtain f_C^l corresponding to each f_{S^l}. Since S^0 is the first one of the subdivision mesh sequence, we select a triangle from C at random as the start one to walk while searching f_C^0 for f_{S^0} of S^0.

In above method, we use the representation triangle f_C^{i-1} of $f_{S^{i-1}}$ as the start one to walk while searching f_C^i of f_{S^i}. Since f_C^{i-1} and q^i lie in the same area decided by $f_{S^{i-1}}$, the search time of f_C^i can be reduced greatly.

5 Locating

Given a point p, our algorithm starts off at a selected triangle and walks towards the triangle containing p by a series of decisions based on barycentric coordinates.

5.1 Selecting the Start Triangle to Walk

The first step of our point location algorithm is to select a good start triangle to walk. Its basic idea is to find the triangle $f_{S^l}^p$, which lies in S^l and contains the query point p, and select the representation triangle f_C^p of $f_{S^l}^p$ as the start triangle to walk. As $S^0, \cdots, S^i, \cdots, S^l$ is a sequence of subdivision connectivity meshes, we use another hierarchy refinement method to find $f_{S^l}^p$. First, we compute the barycentric coordinates (α, β, γ) of p with respect to each triangle of S^0 and find the triangle $f_{S^0}^p$ containing p by $\alpha \geq 0, \beta \geq 0, \gamma \geq 0$. Then consider the $1th$ level subdivision mesh S^1. Since p lies in $f_{S^0}^p$, p is sure to be contained by one subdivision triangle of $f_{S^0}^p$. Instead of examining all the four subdivision

triangles shown in Fig 3, we only compute the barycentric coordinates (α, β, γ) of p with respect to the middle triangle $f_4\,(p_1, p_2, p_3)$. If all the coordinates are nonnegative, then f_4 is the triangle containing p in S^1. Otherwise, we find the vertex p_i with negative barycentric coordinate. Then the opposite triangle f_i of p_i is the triangle containing p in S^1. For example, the opposite triangle of p_1, p_2 and p_3 in Fig 3 are f_1, f_2 and f_3, respectively. Above examining operation is iterated on S^i until S^l is reached. After obtaining $f^p_{S^l}$, we select the representation triangle f^p_C of $f^p_{S^l}$ as the start triangle to locating p.

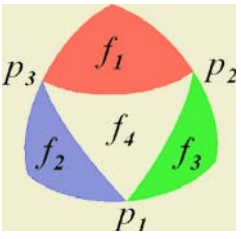


Fig. 3. Four subdivision triangles coming from the same father

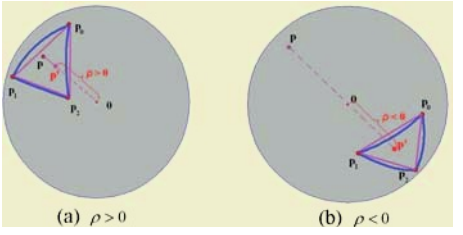


Fig. 4. Two kinds of possible projecting results of p onto f'

5.2 Locating a Query Point by Barycentric Coordinates Approach

After finding the start triangle, our algorithm walks towards the triangle containing p by a series of decisions based on barycentric coordinates. For each visited triangle f during the walking process, we compute the barycentric coordinates (α, β, γ) of p with respect to f. If all the barycentric coordinates are nonnegative, then we can judge that f is the triangle containing p and stop the walking. Otherwise, the algorithm walks to the next triangle and computes the barycentric coordinates again.

While computing the barycentric coordinates (α, β, γ) of p with respect to a triangle $f\,(p_0, p_1, p_2)$ in a plane, we can use

$$\begin{cases} p = \alpha \cdot p_0 + \beta \cdot p_1 + \gamma \cdot p_2 \\ \alpha + \beta + \gamma = 1 \end{cases} \tag{1}$$

But if $f\,(p_0, p_1, p_2)$ lies on a spherical surface, it is difficult to compute the barycentric coordinates by (1). Here we don't directly deal with f, but turn to the corresponding planar triangle $f'\,(p_0, p_1, p_2)$, which is decided by the same three vertices of f. Obviously, there are few chances for p to lie in f'. So we consider the projection p' of p onto f'. Figure 4 shows the projecting approach used in this paper. p' is the intersection point between f' and the line decided by p and the spherical center o.

In Fig 4, suppose the unit vector of $\overrightarrow{op}$ is (l, m, n), the coordinates of o are $(0, 0, 0)$ and the general equation of the plane f' is $Ax + By + Cz + D = 0$. Then the intersection p' can be computed by

$$p' = (\rho l, \rho m, \rho n) \tag{2}$$

where $\rho = -\frac{D}{Al+Bm+Cn}$. After obtaining p', we compute the barycentric coordinates (α, β, γ) of p' with respect to $f'(p_0, p_1, p_2)$.

Now we discuss how to decide the next triangle to walk. According to the value of ρ, three kinds of different situations should be discussed.

First, when $\rho = \infty$, $\overrightarrow{op}$ is parallel with f' and no intersection exists. In this situation, we randomly select an unvisited triangle from the neighbors of f as the next one to walk.

Second, if $\rho > 0$, then the vector $\overrightarrow{op'}$ has the same direction as $\overrightarrow{op}$, which can be seen from Fig 4-(a). In this situation, the projection p' is the intersection of $\overrightarrow{op}$ and f'. We compute the barycentric coordinates (α, β, γ) of p' with respect to f'. According to (α, β, γ), the position relationships between p' and f' can be divided into three kinds of different cases, which are shown in Fig 5. First, if all components of (α, β, γ) are nonnegative, which is shown in Fig 5-(a), then p' is sure to lie in f'. For this case, the corresponding spherical triangle f of f' is the target triangle containing p. Second, if one component of (α, β, γ) such as α is below zero, which is shown in Fig 5-(b), then we can judge that p' lies in the right of the opposite edge $\{p_1, p_2\}$ of p_0. So we select the triangle f_1, which is linked with f by the edge $\{p_1, p_2\}$, as the next one to search. Third, if two components of (α, β, γ) such as α and β are smaller than zero, which is shown in Fig 5-(c), then p' should lie in the intersection region of the two right hemispheres defined by $\{p_1, p_2\}$ and $\{p_2, p_0\}$. In this case, we first find the unvisited 1-ring neighbor triangles set $T^*(p_2)$ of p_2 and sort them by anticlockwise order. Then we start from the element, which is in $T^*(p_2)$ and nearest to $f(p_0, p_1, p_2)$, and select the $\left\lfloor \frac{N^*(p_2)}{2} \right\rfloor$ th element of $T^*(p_2)$ as the next one to search. $N^*(p_2)$ represents the number of elements in $T^*(p_2)$.

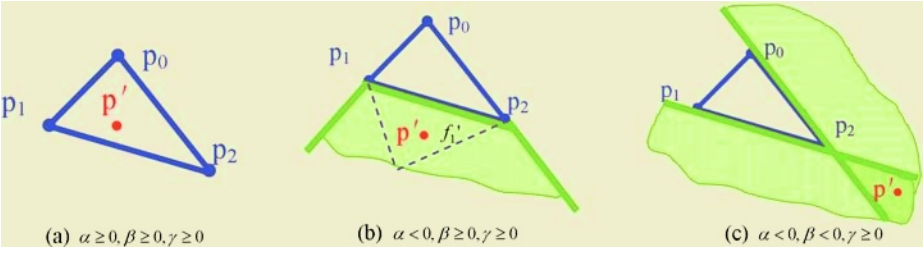


Fig. 5. Three kinds of possible relationships between p' and $f'(p_0, p_1, p_2)$

Third, if $\rho < 0$, then the vector direction of $\overrightarrow{op'}$ is contrary to that of $\overrightarrow{op}$, which can be seen from Fig 4-(b). In this situation, the selection strategy of the next triangle to walk is completely contrary to that of $\rho > 0$. If all components of (α, β, γ) are nonnegative, then p lies in the opposite hemisphere surface of f. We randomly select one unvisited neighbor triangle of f as the next one to walk. If one component of (α, β, γ) such as α is negative, then we suppose $\alpha \geq 0, \beta < 0, \gamma < 0$ and use the same approach as Fig 5-(c) to find the next triangle. If two component of (α, β, γ) such as α and β are negative, then we

suppose $\alpha \geq 0, \beta \geq 0, \gamma < 0$ and use the same approach as Fig 5-(b) to find the next triangle.

6 Algorithm Performance

Before analyzing the algorithm performance, we describe the spherical point location problem again. Given a spherical mesh C with n_f triangles and n_v query points, we are required to find the containing triangle in C for each query point. In order to convenience analysis and comparison of performance, we set $n_v = n_f^{\varepsilon}$, where ε is an exponent to show the relationship between n_v and n_f. Furthermore, since the ratio of triangle number n_f to vertex number n of a 2-manifold oriented triangulation mesh is about 2:1, it is reasonable to replace n by n_f while computing the time-space complexity.

According to our algorithm described in section 4 and 5, the whole running time T is composed of four parts: the constructing time T_b of the spherical subdivision connectivity mesh S^l, the expected search time T_r of representation triangle, the selecting time T_s of start triangle to walk and the expected search time T_h of target triangle containing the given query point. Since the selecting operation of start triangle and the searching operation of target triangle will be executed once for each query point, the whole running time for locating n_v query points can be expressed as $T = T_b + T_r + n_v \cdot T_s + n_v \cdot T_h$. After analyzing the algorithm, we find that $T_b = O\left(4^l\right)$, $T_r = O\left(2^l \cdot T_e\right)$, $T_s = O\left(l\right)$ and $T_h = O\left(\left(\frac{n_f}{4^l}\right)^{\frac{1}{2}}\right)$, where l is the subdivision level of S^l and T_e is the time complexity of locating a point by the barycentric coordinates approach with a random start triangle. Suppose the triangle distribution of original spherical mesh C is uniform, then we have $T_e = O\left((n_f)^{\frac{1}{2}}\right)$. So the expected time complexity of our algorithm can be expressed as $T = O\left(4^l + 2^l \cdot (n_f)^{\frac{1}{2}} + n_v l + n_v \left(\frac{n_f}{4^l}\right)^{\frac{1}{2}}\right)$, which is a function of n_f, n_v and l. As n_f and $n_v = n_f^{\varepsilon}$ are given beforehand, we only consider how to minimize T by selecting the optimal l. The computing result shows that the expected time complexity can be optimized to $\begin{cases} O\left(n_f^{\frac{1}{2}+\frac{1}{2}\varepsilon}\right) & \varepsilon < 1 \\ O\left(n_f^{\varepsilon} \log n_f\right) & \varepsilon \geq 1 \end{cases}$ with $l = \log_2 \left(-\frac{n_v}{2 \ln 2 \cdot n_f^{\frac{1}{2}}} + \frac{1}{2} \left(\left(\frac{n_v}{\ln 2 \cdot n_f^{\frac{1}{2}}} \right)^2 + 4 n_v \right)^{\frac{1}{2}} \right) - 1.5$. The corresponding space complexity is $\begin{cases} O\left(n_f^{\varepsilon}\right) & \varepsilon < 1 \\ O\left(n_f\right) & \varepsilon \geq 1 \end{cases}$.

Table 1 shows the theoretical performance of several different spherical point location algorithms. Although the original algorithms presented by Dobkin [1], Kirkpatrick [2] and Sundareswara [3] are only adapted to planar meshes, they can

be generalized to the spherical domain with unchanged time-space complexity. So they can be listed in Table1 for comparison with our algorithm. While $\varepsilon < 1$, our algorithm has lower time complexity than other four ones. Although $O\left(n_f^{\varepsilon}\right)$ space has to be spent during the preprocessing stage, it is less than $O(n_f)$ of Kirkpatrick's algorithm and $O\left(n_f^2\right)$ of Dobkin's one. While $\varepsilon \geq 1$, the time-space complexity of our algorithm is the same as that of Kirkpatrick's one, which is lower than other three ones.

Table 1. Theoretical performance of several spherical point location algorithms

Algorithm	Preprocessing stage		Search stage	The whole
	Time	Space	Time	time T
Scanning algorithm	——	——	$O\left(n_f^{1+\varepsilon}\right)$	$O\left(n_f^{1+\varepsilon}\right)$
Dobkin's algorithm	$O\left(n_f^2 \log n_f\right)$	$O\left(n_f^2\right)$	$O\left(n_f^{\varepsilon} \log n_f\right)$	$O\left((n_f^2 + n_f^{\varepsilon}) \log n_f\right)$
Kirkpatrick's algorithm	$O(n_f \log n_f)$	$O(n_f)$	$O\left(n_f^{\varepsilon} \log n_f\right)$	$O\left((n_f + n_f^{\varepsilon}) \log n_f\right)$
Sundareswara's algorithm	——	——	$O\left(n_f^{\frac{1}{2}+\varepsilon}\right)$	$O\left(n_f^{\frac{1}{2}+\varepsilon}\right)$
Our algorithm , $\varepsilon < 1$	$O\left(n_f^{\frac{1}{2}+\frac{1}{2}\varepsilon}\right)$	$O\left(n_f^{\varepsilon}\right)$	$O\left(n_f^{\frac{1}{2}+\frac{1}{2}\varepsilon}\right)$	$O\left(n_f^{\frac{1}{2}+\frac{1}{2}\varepsilon}\right)$
Our algorithm , $\varepsilon \geq 1$	$O(n_f)$	$O(n_f)$	$O\left(n_f^{\varepsilon} \log n_f\right)$	$O\left(n_f^{\varepsilon} \log n_f\right)$

7 Experimental Results

We have implemented the spherical point location algorithm based on barycentric coordinates and applied it to several spherical triangulation meshes. All the programs are executed on a 1.4GHz Pentium4 PC with 256MB memory.

For each spherical mesh, we scatter 10,000 query points on its surface by $\left\{(\theta_p, \varphi_p) \,|\theta_p = \frac{2\pi}{100} i, \varphi_p = \frac{\pi}{100} j, i = 1, \cdots, 100, j = 1, \cdots, 100\right\}$. The subdivision level l_0 is computed by $l_0 = \log_2 \left(-\frac{n_v}{2 \ln 2 \cdot n_f^{\frac{1}{2}}} + \frac{1}{2} \left(\left(\frac{n_v}{\ln 2 \cdot n_f^{\frac{1}{2}}} \right)^2 + 4 n_v \right)^{\frac{1}{2}} \right) - 1.5$ and rounded by $l_o = \lfloor l_o \rfloor$.

Figure 6 shows the locating process of our algorithm on the spherical Bunny mesh with $n_f = 69{,}630$. The spherical position of the given query point is $p\left(\theta = \frac{\pi}{4}, \varphi = \frac{\pi}{4}\right)$. During the preprocessing stage, we compute $l_o = 4$ and construct the regular spherical mesh S^4, which is shown in Fig 6-(b). The visited triangles while locating p are illustrated in Fig 6-(d), where f_S^p is the triangle containing p in S^4 and 0 is the selected start triangle containing the centroid q of f_S^p.

Table 2 shows the practical running time of locating $n_v = 10,000$ query points in the spherical meshes by several different algorithms. From Table 2, it can be found that for all spherical meshes, our algorithm has lower running time than

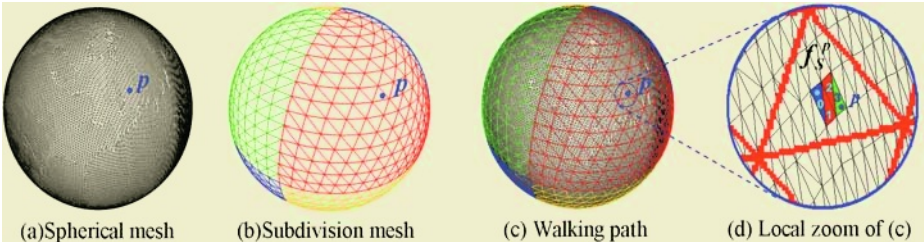


Fig. 6. The spherical point location process of our algorithm

other two ones. With the increment of n_f, the running time will also increase. But n_f is not the only one to decide the running time. Consider Gargoyle and Tyra. Although they have the same n_f, Tyra has higher running time than Gargoyle. This shows that the triangle distribution density of the original spherical mesh C is another factor to influence the running time. In order to reduce this influence, a valid way is to construct a subdivision connectivity mesh S^l with the same triangle distribution as C.

Table 2. The running time of several different spherical point location algorithms. ($n_v = 10,000$, T is the practical running time of locating n_v query points)

Spherical meshes	Pawn	Cow	Bunny	Venus	Gargoyle	Tyra
Number of triangles (n_f)	304	23,216	69,630	100,000	200,000	200,000
Number of query points (n_v)	10,000	10,000	10,000	10,000	10,000	10,000
Exponent $\varepsilon = \log_{n_f}(n_v)$	1.61	0.92	0.83	0.80	0.75	0.75
Optimal subdivision level l_o	2	4	4	4	4	4
Scanning algorithm T(s)	8.81	365.61	1937.52	2618.78	4587.51	6225.21
Sundareswara's algorithm T(s)	1.17	4.80	16.95	17.66	28.61	30.31
Our algorithm T(s)	0.39	0.66	0.92	0.94	1.30	1.48

8 Conclusions and Future Works

In order to solve the spherical point location problem, we present an algorithm based on barycentric coordinates. By constructing the spherical subdivision connectivity mesh S^l, we can select a good start triangle to walk. By using the barycentric coordinates approach, we can find the shortest path from the start triangle to the target triangle containing the given query point. The results of analysis show that while $\varepsilon < 1$ ($n_v < n_f$), our algorithm has lower expected time complexity than traditional ones and lower extra space complexity than those based on preprocessing operation. While $\varepsilon \geq 1$ ($n_v \geq n_f$), our algorithm has the same time-space performance as Kirkpatrick's one, whose time complexity is optimal in theory. The experimental results further show that our algorithm has better running performance than traditional ones.

Although our algorithm has lower expected time complexity than traditional ones, the worst time complexity $O\left(n_f^{1+\varepsilon}\right)$ may appear when the triangle distribution of C is not uniform. So how to construct S^l with the same triangle distribution as C should be studied in future.

Acknowledgments

The 863 program of China (2003AA411310) supports our research. We would like to thank the Stanford Computer Graphics Laboratory for the spatial models. We also appreciate the source code of Qslim offered by Michael Garland.

References

[1] Dobkin. D. and Lipton. R.J. Multidimensional Searching Problems. SIAM J. Comput. 5, 2 (1976) 181-186.
[2] Kirkpatrick. D.G. Optimal Search in Planar Subdivisions. SIAM J. Comput. 12, 1 (1983) 28-35.
[3] Rashmi Sundareswara, Paul R. Schrater: Extensible Point Location Algorithm. GMAG (2003) 84-89
[4] A. Certain, J. Popovi'c, T. DeRose, T. Duchamp, D. Salesin, and W. Stuetzle. Interactive Multiresolution Surface Viewing. In Proceedings of ACM SIGGRAPH (1996), pages 91-98.
[5] U. Labsik, L. Kobbelt, R. Schneider, and H. P. Seidel. Progressive Transmission of Subdivision Surfaces. Computational Geometry, 15:25–39, 2000.
[6] L. Kobbelt, J. Vorsatz, U. Labsik, and H. P. Seidel. A Shrink Wrapping Approach to Remeshing Polygonal Surfaces. In Computer Graphics Forum (EUROGRAPHICS '99 Proceedings), pages 119–130, 1999.
[7] Hormann, U. Labsik, and G. Greiner. Remeshing Triangulated Surfaces with Optimal Parameterizations. Computer-Aided Design, 33(11): 779–788, September 2001
[8] Praun, E., and Hoppe, H. Spherical Parametrization and Remeshing. In Proceedings of ACM SIGGRAPH (2003), pages 340-349.

Realistic Skeleton Driven Skin Deformation

X.S. Yang and Jian J. Zhang*

National Centre for Computer Animation,
Bournemouth University, United Kingdom
{xyang, jzhang}@bournemouth.ac.uk

Abstract. Skeleton driven animation is a popular method for the animation of deformable human and creature characters. The main advantage is its computational performance. However it suffers from a number of problems, such as collapsing elbow and candy wrapper joint. In this paper, we present a new method which is able to solve these defects; reduce the animator's manual work still allowing his/her full control over the process; and realistically simulate the fat bulge effect around a joint.

1 Introduction

Skin deformation is one of the most important issues in realistic character animation and has received a great deal of attention from the animation research community over the last two decades. There are currently two prevalent approaches, one is based on the anatomy of the human/creature characters and the other deforms the character skin directly.

The anatomy based technique tries to mimic the muscle structure of a human or creature. Normally three layers are used, the skeleton, musculature and possibly fat, and the skin layer [1]. This approach usually works by layering individual CG (computer graphics) muscles on the skeleton. These muscles deform (stretch or bulge) following the motion of the skeleton. The final skin takes the overall shape of the muscle and fat layer of the animated character body. Because the skin shape is derived from the underlying structures, this approach affords good graphical realism. The disadvantage, however, is that it is tedious and unintuitive to use. Not many people in the animation industry employ this approach due to these drawbacks.

The second approach, often known as the smooth skinning technique, makes no direct use of the anatomical elements. The skin shape is controlled by the transformations associated with the joints of the skeleton. This technique is simple to use and very fast to compute. Due to these advantages, it is the most popular in animation production and has been incorporated into many animation packages. As it uses a very simple shape blending technique to deal with an inherently very complex matter, it is understandable that this technique is not expected to cover all skin deformation problems. In practice, it suffers from a

* Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1109–1118, 2005.

number of defects, some of which are very difficult to rectify. Typical problems include collapsing elbow, candy-wrapper joint when the arm turns 180 degrees, and intersection between two adjacent bones (links) around a joint. These problems are especially obvious around the shoulder and elbow areas. Manual efforts are often the only option, which can be tedious and undesirable. To understand the cause of these problems, the reader is referred to [2].

1.1 Previous Work

Skin deformation has been an interesting research topic for a long time. Two dominating techniques exist as stated above, one based on anatomy and the other is not.

The anatomy-based technique is able to achieve very good graphical realism on the skin surface. Physical properties [3] can be associated with the individual anatomical elements, such as muscles and other internal elements. When driven by the skeleton, these anatomical elements move and deform accordingly resulting in realistic deformation of the skin surface. There is a large amount of development in this area including Wilhelms and Gelder [4], Scheepers et al. [1], and Aubel and Thalmann [5]. Despite its strength, this approach has not proven popular in animation production, due to the above mentioned disadvantages.

The traditional skinning model, which involves no knowledge of anatomy, goes by various names. Lewis called it Skeleton Subspace Deformation (SSD) [2] and the software package Maya calls it Smooth Skinning. In this paper we will also adopt this name. An introduction of this technique can be found in Lewis et al [2].

Catmull [6] introduced one of the first skeleton-driven techniques to animate an articulated figure. An early 3D skeleton-driven technique that involves deformable skinning was presented by Magnenat-Thalmann et al. [7]. This technique deforms character meshes using the motion of particular joints.

Later improvements have been aiming to both improve realism and reduce tedious manual intervention. One effective idea is to use a large number of examples, which may be obtained either by manual modelling or by range-scanning. Wang and Phillips [8] proposed a least squares multi-weight technique to compute the weights of the elements of the transformation matrices based on examples that are usually hand-modelled. The technique by Allen et al [9] is also based on many examples produced by laser-scanning. The idea is to interpolate the scanned key shapes to derive a fuller set of shapes. Both techniques are able to produce very realistic results, but all rely on the availability of a large number of pre-obtained models.

1.2 Overview

Our aim of this paper is to retain the strength of the popular smooth skinning technique and keep the current production pattern unchanged, in the meanwhile to allow the graphical realism to be significantly improved. Thus the animator can continue with their pattern of practice which they have been used to for a long time. In particular, we make two important contributions:

- Non-linear skin deformation. Current techniques either require heavy manual input or blend shapes in a linear manner. We have found linear shape blending to be a cause for many unpleasant defects. In this paper, we present a non-linear technique to solve this problem.
- Fat bulge effect. When an arm bends, the flesh is pushed out sidewise due to the property of volume preservation. This phenomenon, despite its subtlety, has a profound effect on realism. In this paper, we present an efficient algorithm to model it effectively.

The remainder of the paper is structured as follows. In Section 2 we present our nonlinear smooth skinning method. We will first introduce briefly the traditional technique. Then we discuss how it is improved to produce visually pleasant results. Section 3 discusses the modelling of the subtle fat effect at a bent joint. Section 4 concludes this paper.

2 Nonlinear Smooth Skinning

As will be seen below, the defects of the smooth skinning technique is to some extend caused by the linear interpolation mechanism involved. In this Section, we present a new technique to remedy this problem.

2.1 Smooth Skinning

Current linear blend skinning algorithms work by first placing a skeleton inside a character model, usually in a neutral pose. Each vertex is then assigned a set of influencing joints together with a weight factor corresponding to each influencing joint. Deforming the character into a different pose involves transforming each vertex from the initial pose by all the influencing joints. The transformed positions are finally blended together to give the final position of all the vertices. At a skeletal configuration c, a deformed vertex V_c, can be computed by [10]:

$$V_c = \sum_{i=1}^{n} w_i M_{i,c} M_{i,d}^{-1} V_d \tag{1}$$

where w_i are the weights, V_d is the location of a vertex at its initial pose, $M_{i,c}$ denotes the transformation matrix associated with the ith joint in configuration c and $M_{i,d}^{-1}$ the inverse of the transformation matrix associated with the ith influencing joint.

This skinning algorithm is very fast and widely adopted by animation software packages. However as discussed earlier, it fails for complex deformations and suffers from a number defects such as collapsing elbow and shrinkage around a joint.

2.2 Nonlinear Skin Deformation

All these problems are rooted from how the vertex weight is computed. Most animation software leaves this work to the animator who paints the weight by

hand. Understandably, building a complete sequence of poses entirely by hand is very time-consuming. Attempts were made to relieve the animator from the drudgery, such as [8] and [10] by trying to reverse-engineer the weights from a well defined pose space. However, problems remain if such a pose space, usually constructed from a large number of existing example models, is not available.

In this paper, we present a method for direct assignment of weights to the vertices according to its position around the joint. The method is introduced in three steps as follows.

Weight computation. The weight coefficient modulates the influence of a joint to the point concerned. To develop an effective weight computation model, we first define the properties that an ideal model should satisfy, which are given as follows:

- Smoothness within influenced region. Eq. (1) suggests that smooth skinning transforms a point to a new position defined by blending all relevant transformations together. Linear blending often suffers from lack of smoothness. To ensure satisfying smoothness, a higher degree of continuity is necessary, at least with a tangent continuity. This suggests the necessity of a nonlinear weight computation model.
- Smoothness at the boundary of an influenced region. To avoid sudden changes at the boundaries of an influenced region, weights should change gradually, i.e. avoiding large weight discrepancy at the boundaries.
- (c) Symmetry. To avoid a biased shape, the weights should be distributed approximately symmetrically.

To facilitate discussion, we also need to define some quantities. The closeness of the point concerned, P, to an individual bone (link) j is measured by the *influence angle* α_j (Figure 1a). If $\alpha_j < \alpha$, where α is called the *limit angle*, then point P is influenced only by the transformation associated with joint j, otherwise, P is influenced by the neighbouring joints. The limit angle is specified by the user to indicate how much influence the skin should be affected by the neighbouring joints. Another useful quantity is called the *influence ratio* r, which is defined below:

$$r = \frac{\alpha_{i-1} - \alpha}{\alpha_{i-1} + \alpha_i - 2\alpha} \tag{2}$$

where α_{i-1} or α_i are influence angles (Figure 1a). The default weight value of P influenced by joint i is thus given by:

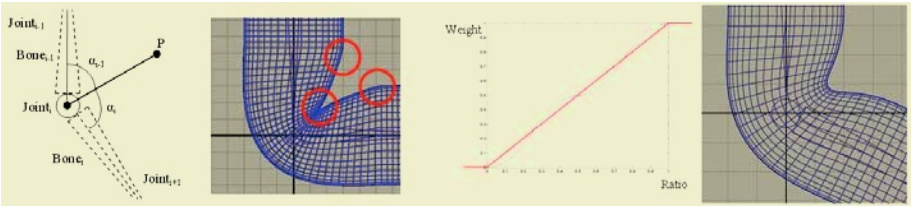
$$W_i = r \tag{3}$$

This however, is not satisfactory. Figure 1b shows the resulting deformation on a Nurbs surface. Discontinuity is evident as marked by the red circles. This is because the linear relationship between the weights and the influence ratio r (the red line shown in Figure 1c). Improvement can be made by defining a higher order function of r, which satisfies the three conditions stipulated above.

higher order function of r, which satisfies the three conditions stipulated above. For this purpose, we propose the following polynomial form:

$$W(r) = \sum_{i=0}^{d} a_i r^i \tag{4}$$

where a_i are the unknowns, which should be determined by satisfying conditions (b) and (c), as (a) will be automatically satisfied by Eq. (4).



a b c d

Fig. 1. Weight computation: (a) Parameters (b)Linear smooth skinning (c)Weight distribution (d) Smooth skinning

It is easy to see that conditions (b) and (c) are equivalent to the following constraints:

$$\begin{aligned} W(0.0) = 0.0 \;\; W(0.5) = 0.5 \;\; W(1.0) = 1.0 \\ W'(0.0) = 0.0 \; W'(0.5) = 1.1 \; W'(1.0) = 0.0 \end{aligned} \tag{5}$$

$W(0.5) = 0.5$ makes the distribution roughly symmetric and $W'(0.5) = 1.1$ is chosen according to our experiment results. Using these 6 constraints (5), one is able to determine the unknowns α_i from (4) by solving a set of linear equations. Thus we obtain the following weight distribution equation:

$$W(r) = -6.4r^5 + 16r^4 - 14.8r^3 + 6.2r^2 \tag{6}$$

Figure 1d shows the improved result using Eq. (6).

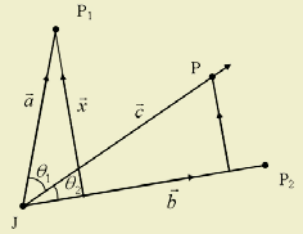


Fig. 2. Interpolation of vertex near joint

Compensation for a collapsing joint. Collapsing joint as a result of big bent angles, is another typical problem. The artefacts occur because the vertices are transformed only in a linear manner, i.e. no consideration is given to its angular contribution.

To rectify this problem, we first compensate the shrinkage by pulling the skin surface away from the joint, and then blend the influences from relevant joints angularly. Assume J is the joint centre (Figure 2), P_1 and P_2 are the transformed positions obtained using the transformation matrices of the two related joints respectively. The new position of the point P is given by:

$$P = \frac{\vec{b}}{\left|\vec{b}\right|}\left|\vec{c}\right|\cos\theta_2 + \frac{\vec{x}}{\left|\vec{x}\right|}\left|\vec{c}\right|\sin\theta_2 \tag{7}$$

where

$$\begin{aligned}
&\vec{a} = \overrightarrow{OP_1} - \overrightarrow{OJ}\ \vec{b} = \overrightarrow{OP_2} - \overrightarrow{OJ}\\
&\theta = \arccos\left(\frac{\vec{a}\bullet\vec{b}}{|\vec{a}||\vec{b}|}\right)\\
&\theta_2 = \theta * (1 - W_i)\ \ \theta_1 = \theta - \theta_2\\
&\vec{x} = \vec{a} - |\vec{a}|\cos(\theta_1 + \theta_2)\frac{\vec{b}}{|\vec{b}|}\\
&l_2 = |(\overrightarrow{OP_2} * W_i + \overrightarrow{OP_1} * (1 - W_i)) - \overrightarrow{OJ}|\\
&W_i' = 1 - 4 * (W_i - 0.5)^2\\
&l_1 = |\vec{b}| * W_i + |\vec{a}| * (1 - W_i)\\
&|\vec{c}| = l_1 * W_i' + l_2 * (1 - W_i')
\end{aligned} \tag{8}$$

where W_i stands for the weight of joint i, O is the origin of the world coordinate system. Figure 3 compares the result for this method. Compensation is effective not only at the inner part but also the outer part of the joint, which is a desirable feature.

Limit angle. The limit angle is another powerful tool, which specifies the influenced scope of a joint. So far this quantity is set as a constant angle. However, by changing this angle in response to the bent angle of a joint, more desirable results can be achieved. In Figure 3, for example, this angle is set to be 30 degrees. This works fine until the bent angle reaches 60 degrees. If the joint bends further, interpenetration starts to show up. From our experience, the joint will deform reasonably if the angle between the two adjacent bones is not smaller than a double of the limit angle. This suggests that for better deformation result,

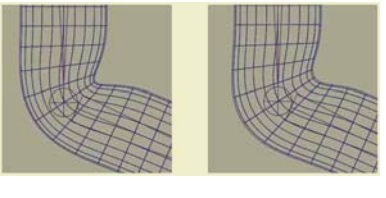


Fig. 3. The compensation result (left: uncompensated, right: compensated)

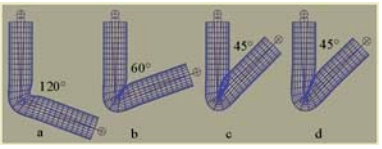


Fig. 4. constant angle for the joint affect scope limit

the limit angle should change dynamically to suit individual cases. In Figure 4 the limit angle is set to be 30 degrees, which is well-behaved for the first two cases where the bent angles are 120 and 60 degrees, respectively. Intersection arises for the third one when the bent angle reduces to 45 degrees. However, if the limit angle is reduced, the problem goes away (Figure 4d).

3 Fat Bulge Effect Around Bent Joint

So far as the visual realism of a human character is concerned, the fat effect comes into play with two contributions. It smooth out the shape obtained from the muscle formation under the skin. The second gives the bulge effect around a joint. When a joint bends, the flesh under the skin is pushed out due to the property of volume preservation.

While the first is relatively easy to do, the second problem proves much harder, as the shape changes dynamically as the joint bends. Current techniques unfortunately struggle to maintain a constant volume level, which result in unrealistic skin shapes. Figure 5 shows our analysis results. The bottom curve records the volume loss from Eq. (3) and the top one from Eq. (7), which compensates the shrinkage around a joint. Although the compensation method Eq. (7) improves the situation, further improvement remains necessary.

Accurate bulge effect can be computed with physical simulation, which nevertheless is by no means an easy task and usually is computationally expensive. Here we present a geometrical method to achieve satisfactory visual look in real-time. As fat is largely incompressible, when a joint bends, flesh between the adjacent bones will be squeezed, producing bulges immediately near the joint and at the sides. The main idea of our method is to formulate a surface function

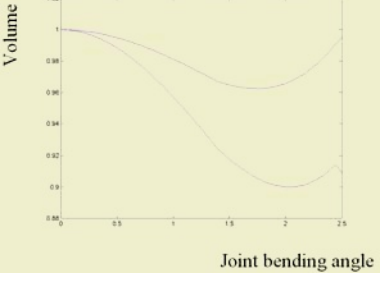


Fig. 5. Volume loss by current methods

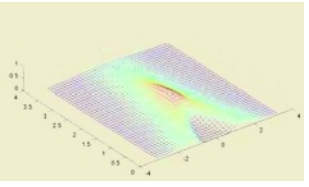


Fig. 6. Fat bulge distribution and ridge curve

able to both represent the bulge distribution and the deformation linked with the joint angle parameter.

Within a cylindrical co-ordinate system, the bulge distribution can be described as a surface, and the ridge forms a curve on this surface as seen in Figure 6.

The ridge curve can be defined as,

$$y = \frac{2\pi - 4\alpha_0}{\pi^2} x^2 + \alpha_0 \tag{9}$$

where α_0 is the limit angle of the joint. The closer to the ridge curve, the bigger bulge occurs and the height roughly takes the normal distribution centred at the ridge curve. This suggests that for each skin vertex $P(x_0, y_0)$, we have the following bulge function,

$$P(R) = \begin{cases} 16A(R-1)^4 - 8A(R-1)^2 + 0.5 & 0.5 < R < 1.5 \\ 0 & R \leq 0.5 \\ 0 & R \geq 1.5 \end{cases} \tag{10}$$

where A is a parameter associated with the volume loss from Fig.5 and

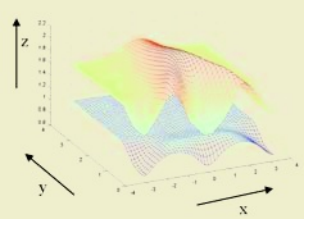


Fig. 7. Fat bulge effect superimposed on the skin shape produced by the compensation method presented above. In this cylindrical co-ordinate system, x stands for the azimuth coordinate, y for the distance from the joint along the limb, z for the radial coordinate. For each vertex on the skin surface, the deformation only affects its z value. The bottom mesh shows the skin deformation around a joint using our compensation method. The top mesh, the final result including fat bulge, adds contribution from the fat layer to each vertex's z coordinate according to the bulge distribution function defined in Eq. (9)

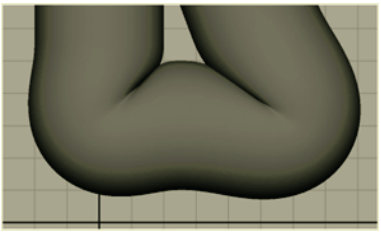


Fig. 8. Fat bulge distribution and ridge curve

$$R = \frac{y_0}{\frac{2\pi - 4\alpha_0}{\pi^2} x_0^2 + \alpha_0} \tag{11}$$

Figure 7 shows the fat bulge distribution surface and Figure 8 shows the deformation of a cylinder-shaped "arm", which was computed by the nonlinear smooth blending and fat bulge compensation methods presented in this paper.

4 Conclusion and Future Work

We have presented a new smooth skinning method for character skin deformation. By setting up a set of criteria that a skinning method should satisfy, we have derived a nonlinear weight distribution function. Unlike other existing weight computation methods, our method does not rely on pre-made examples. Current techniques suffer from a number of defects, such as the clasping elbow problem. In this paper, we have studied the cause of such problems and presented an effective remedy. We also find that dynamically adjusting the limit angle is of a practical value in producing desirable deformation results. Using this technique, the animator is able to produce realistic skin deformation easily and still able to exercise his/her control.

Fat bulge effect adds further realism to a deformable animated character. In this paper we have presented a fast method to model this phenomenon. Although no accurate physics was involved, this method achieves convincing visual results.

We have implemented our method into a plug-in in the Maya Environment alongside the traditional smooth skinning technique. We also present the animator a user-interface to control the important parameters, such as the fat region and bulge dimension. Our method works very well on the character's limbs. On the body, however, because it involves more than two bones, the compensation method is less effective. We would like to improve this matter in the future.

Acknowledgements

This research is funded by the AHRB grant B/RG/AN5263/APN12727. We are grateful for the donation of the Maya licenses from Alias.

References

1. Scheepers, F., R. E. Parent, et al: Anatomy-based modeling of the human musculature. Proceedings of the 24th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co. (1997) 163–172.
2. Lewis, J. P., M. Cordner, et al: Pose space deformation: a unified approach to shape interpolation and skeleton-driven deformation. Proceedings of the 27th annual conference on Computer graphics and interactive techniques, (2000). ACM Press/Addison-Wesley Publishing Co.: 165–172.
3. Chen, D. T. and D. Zeltzer: Pump it up: computer animation of a biomechanically based model of muscle using the finite element method. Proceedings of the 19th annual conference on Computer graphics and interactive techniques, (1992) ACM Press: 89–98.
4. Wilhelms, J. and A. V. Gelder: Anatomically based modeling. Proceedings of the 24th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co. (1997) 173–180.
5. A. Aubel, D. T.: Efficient Muscle Shape Deformation. Deformable Avatars, Kluwer Publ. (2001). 132-142.
6. Catmull, E. E: A system for computer generated movies. Proc. ACM Annual Conf. (1972)
7. Magnenat Thalmann, M., Laperrire, R., And Thalmann, D.: Joint dependent local deformations for hand animation and object grasping. Proceedings of Graphics Interface'88, (1988) 26-33.
8. Wang, X. C. and C. Phillips: Multi-weight enveloping: least-squares approximation techniques for skin animation. Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation, ACM Press (2002) 129–138.
9. Allen, B., B. Curless, et al:Articulated body deformation from range scan data. Proceedings of the 29th annual conference on Computer graphics and interactive techniques, (2002), ACM Press: 612–619.
10. Mohr, A. and M. Gleicher: "Building efficient, accurate character skins from examples." ACM Trans. Graph. **22**(3) (2003) 562–568.

Implementing Immersive Clustering with VR Juggler

Aron Bierbaum[1], Patrick Hartling[1], Pedro Morillo[2], and Carolina Cruz-Neira[1]

[1] Virtual Reality Applications Center, Iowa State University. USA
[2] Departamento de Informática, Universidad de Valencia. Spain
{aronb, patrick, cruz}@iastate.edu
Pedro.Morillo@uv.es

Abstract. Continuous, rapid improvements in commodity hardware have allowed users of immersive visualization to employ high-quality graphics hardware, high-speed processors, and significant amounts of memory for much lower costs than would be possible with high-end, shared memory computers traditionally used for such purposes. Mimicking the features of a single shared memory computer requires that the commodity computers act in concert—namely, as a tightly synchronized cluster. In this paper, we describe the clustering infrastructure of VR Juggler that enables the use of distributed and clustered computers for the display of immersive virtual environments. We discuss each of the potential ways to synchronize a cluster for immersive visualization in use today. Then, we describe the VR Juggler cluster infrastructure in detail, and we show how it allows virtual reality application developers to combine various existing clustering techniques to meet the needs of their specific applications.

1 Introduction

Traditionally, multi-screen immersive visualization systems have relied upon dedicated, high-end, shared memory graphics computers to generate interactive virtual environments. These systems must not be confused with distributed virtual environment (DVE) systems where many users connect from remote computers, typically through the Internet, to share the same 3D virtual world [16]. Such multi-screen immersive systems typically require one or two video outputs for each projection surface, and they often utilize many input devices simultaneously. In recent years, the nearly exclusive use of high-end computers for these purposes has shifted to commodity hardware as it has become a viable alternative [1][8][17]. Continuous, rapid improvements in commodity hardware have allowed users of immersive visualization to employ high-quality graphics hardware, high-speed processors, and significant amounts of memory for much lower costs than would be possible with high-end, shared memory computers. However, to drive a multi-screen immersive visualization system, we need multiple commodity systems working as a single unit, thereby mimicking the behavior of a single, shared memory computer. This can be accomplished through the use of a tightly synchronized cluster.

Clustering techniques have been utilized to parallelize complex computations for many years in high-performance computing (HPC) [4][17]. While HPC clusters offer an alternative to expensive supercomputers to drive multi-screen visualization systems, the existing parallelization techniques used for HPC cannot be applied directly to graphics

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1119–1128, 2005.

clusters. Graphics programmers have to be aware of the following issues when writing virtual reality (VR) software:

- **High-performance network:** Interactive graphics require extremely low latency communication networks in order to maintain real-time frame rates.
- **Swap buffer synchronization:** In order to prevent tearing while combining the rendered images, we must carefully synchronize the swapping of the front and back frame buffers.
- **Consistent random number generation:** Applications that use random numbers must get identical results on each node.
- **Frame delta:** Many applications use the elapsed time since the last frame in various calculations, so time information must be the same on all nodes for each frame.
- **Start barrier:** Certain clustering techniques require that each node starts the first frame of the application at the same time.
- **Multiple input devices:** Most VR input devices are connected to a computer via a serial port, but most commodity computers have only one or two serial ports. This can be addressed by allowing multiple computers to be used as input device hosts and providing device location transparency at the application level.

It is important to note that the above constraints are all solved in the hardware of shared memory computers. For a graphics cluster, they must be solved at the software level.

In this paper, we describe the clustering infrastructure of VR Juggler [3], which enables the use of distributed and clustered computers for the display of immersive virtual environments. The main goals of the VR Juggler cluster infrastructure are to allow the cluster software to adapt to the particular hardware configuration of the virtual reality system; to provide application portability and scalability from high-end systems to commodity clusters by hiding the clustering from developers; and to allow users to customize the clustering methods being used to best meet their specific needs.

2 Background

In recent years, there have been several efforts to create software libraries that generate immersive environments by utilizing clusters of commodity computers. Different techniques are used to address the issues listed in the previous section, and each of these solutions attacks the task of data synchronization at one of four locations: input data, remote shared memory, scene graph change lists, or graphics primitives. In this section, we will briefly discuss each of these potential synchronization points and the software libraries that utilize them. The following subsections are organized according to the pipeline utilized by VR software in which input from physical devices is processed to effect changes in the graphics being rendered.

In order for a user to become fully immersed in a virtual environment, they must interact with it using one or more physical input devices such as a position tracker or glove. Since the objective is to provide the user with a sense of immersion these devices obtain all the input needed to determine the changes in application state. Clustering solutions utilizing *input data* sharing start a distinct complete copy of the application on each node in the cluster. All input data is then synchronized across the cluster at the

beginning of each frame loop. Thus, application state remains consistent as long as it depends solely on input events. This approach does not require any changes to the application relative to a shared memory architecture (since the application still has access to the same input data and rendering targets), random number generation, consistent frame deltas, and a start barrier are not addressed. Examples of multi-screen immersive visualization systems based on input data sharing are Net Juggler [1] and Syzygy [15]. While Net Juggler uses message passing via the Message Passing Interface (MPI) [7], frameworks such as VR Juggler and Syzygy use the TCP/IP suite directly.

The remote shared memory approach offers another way to ensure that each node has an identical snapshot of that state for rendering each frame. Implementations of remote shared memory often require that the application programmer take special steps to use it. Special storage areas must be created, and in some cases, access to the shared memory must be controlled so that there may be multiple readers but only one writer. Different designs put more or less of the burden on the application programmer for understanding and managing these details. Implementations such as DIVERSE [9] are based on a shared memory architecture where a general inter-process communication programming tool guarantees identical copies among cluster nodes.

Since most graphic applications are based on the manipulation of scene graphs [18], if one node keeps track of all changes made to the scene each frame, that node can send the changes to each of the other nodes to be reapplied to the local memory copy of the scene graph. Therefore, each node always has the information it needs to render an accurate version of the scene. This approach, called *scene graph change lists*, takes advantage of the fact that visual consistency and coherency is the critical aspect of all graphics clusters. Both OpenSG [14] and Syzygy implement this clustering method.

At the lowest level, all immersive applications generate a stream of graphics commands that are delivered to the graphics hardware for rendering. This is accomplished by making calls to a low level graphics application programming interface (API) such as OpenGL. Software libraries such as Chromium [8] and DGL [10] are designed to intercept the *graphics primitives* for the rendering of each frame and distribute them over a network in order to divide the rendering task among multiple nodes. This approach tends to require more bandwidth than any of the previously mentioned methods [17].

3 Conceptual Model and Architecture

The clustering techniques presented above have their own unique benefits and drawbacks. We argue that there is no single technique that best meets the needs of all possible VR applications. In other words, any given application may require the use of one or more of the above approaches to sharing data across a graphics cluster. Hence, we have developed an architecture that allows the application developer to choose the method (or methods) that meets their specific needs.

Based on these ideals and on our past experiences, we present the current design of the VR Juggler clustering infrastructure, a modular and extensible architecture that allows the use of any of the methodologies for clustering described above. The VR Juggler clustering infrastructure provides code that performs input data sharing and remote shared memory. The sharing of graphics primitives or scene graph change lists can be

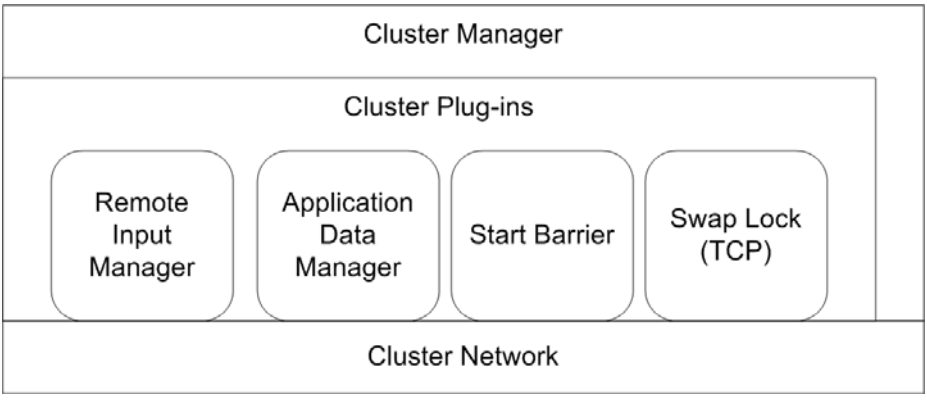


Fig. 1. Architecture of the VR Juggler clustering infrastructure

accomplished by combining VR Juggler with tools such as Chromium or OpenSG. The current work represents a major step forward with respect to the original clustering support for VR Juggler, written in 2002 [11]. The design of the VR Juggler cluster infrastructure contributes several key features not found in other clustering architectures: run-time reconfiguration, a component-based system, and a layered architecture.

3.1 Run-Time Reconfiguration

VR Juggler has an advanced configuration infrastructure that allows configuration changes to be made to running applications [2][3]. Configuration information arrives in the form of *config elements* which are the fundamental unit of configuration in VR Juggler. New config elements may arrive at any time during the lifetime of an application, thus allowing run-time reconfiguration of the software. The clustering infrastructure takes advantage of this feature by allowing nodes, displays, and input devices to be added, removed, or reconfigured as needed at run time.

3.2 Extensible Component-Based Design

The VR Juggler clustering infrastructure provides a framework for dynamically loading cluster components as plug-ins, formally called Cluster Plug-ins. We have followed the traditional component-based approach [19] for developing this architecture. This means that the interface of the component is physically separated from the implementation. Each plug-in is a standalone component loaded at run time based on the user-specified cluster configuration. The plug-ins extend the clustering infrastructure with specialized clustering functionality. Users can choose any of the plug-ins needed for their applications and their specific cluster configuration.

3.3 Layered Architecture

As Fig. 1 shows, the VR Juggler clustering infrastructure is comprised of a set of components that are arranged into layers. Each of these layers builds on the functionality of the layers below it. At the lowest level, the Cluster Network provides a messaging interface for communicating with the entire cluster. The Cluster Plug-ins are built on top

of the Cluster Network and provide the application developer with a set of components to construct the best solution for their applications needs. The top layer is the Cluster Manager which acts as a façade to the clustering infrastructure, which is accessed by higher level code.

Cluster Manager Layer. The Cluster Manager is responsible for handling the cluster configuration and synchronizing the calls to each plug-in. Because the Cluster Manager receives the configuration information about what Cluster Plug-ins to load, it is responsible for performing the steps of loading the plug-ins dynamically and managing their execution. The loading of plug-ins follows the traditional approach used by component-based software. Shared object code is loaded into memory, and an entry point function is called to create an instance of the plug-in.

Before a plug-in can be used, however, it may need to be configured. For example, a plug-in that implements a master/slave network protocol would need to be configured to identify which node in the cluster is the master. The Cluster Manager performs the task of registering plug-ins so that they may receive run-time configuration information.

The design of VR Juggler is one of a micro-kernel architecture where the main thread of control for all applications is the "kernel loop" [3], which corresponds to the concept of a "frame loop" in the terms of real-time rendering software. The kernel contains well-defined points where input data is current and graphics contexts are active. The Cluster Manager can in turn invoke the methods of the plug-ins at well-defined times during the kernel loop. Any given plug-in could need to synchronize its data at a variety of points during the execution of the kernel loop. In order to accommodate all possible needs, each plug-in has a well-defined interface and a contract that specifies the invocation timing.

Cluster Network Layer. The Cluster Network layer maintains an abstract representation of the system of interconnected nodes that comprise the cluster. This abstraction provides the VR Juggler clustering infrastructure with a messaging interface for communicating with the entire cluster. Internally, it maintains a list of the nodes in the cluster along with the current network connections used to communicate with them.

Communication between the cluster nodes is done using a custom protocol. The protocol is based on the well-known, standard TCP/IP suite to avoid requiring any special hardware. Furthermore, the protocol is designed to be efficient because it uses domain-specific knowledge about how VR Juggler applications are written and how they execute. Presently, the Cluster Network uses only the Transmission Control Protocol (TCP) for communicating between nodes. The design of the Cluster Network Layer allows for future implementations using other networking protocols such as the User Datagram Protocol (UDP).

Cluster Plug-Ins. The Cluster Plug-ins are the point of extension for the VR Juggler clustering infrastructure. New plug-ins can be added to address cluster-specific application issues not handled by the standard set of Cluster Plug-ins. VR Juggler includes several commonly used plug-ins that serve as the base for clustering in VR Juggler. These are the *Remote Input Manager Plug-in* (RIM Plug-in), the *Application Data Manager* (ADM), the *Start Barrier Plug-in*, and the *Swap Lock Plug-in*s.

The RIM Plug-in was created to share device input across a cluster. It allows input devices to be connected to one or more nodes in the cluster. The RIM Plug-in ensures that all nodes in the cluster have a consistent snapshot of all input data, regardless of the location of the physical hardware. The RIM Plug-in helps us achieve our goal of designing a distributed shared memory system for VR application input/output (I/O) data. As a result of distributed I/O, application development and execution can move transparently between shared memory VR systems and cluster VR systems. Ideally, the same VR applications will run on both types of systems without any changes to the application. VR Juggler already has an input manager that handles local input data on a single computer. The RIM Plug-in extends the existing input manager to allow input from multiple computers.

The ADM was created to facilitate the sharing of application-specific data structures across the cluster. For example, random numbers can be generated on a single node and shared each frame. This capability extends the fundamental input data sharing and demonstrates that the VR Juggler clustering infrastructure design allows multiple clustering techniques to be utilized. Sharing of application-specific data structures works by providing the application developers with a base class that they extend with their own type. Application developers must give each of their application-specific types an identifier so that the ADM can handle instances of their types correctly at run time. For this, we use 128-bit identifiers called Globally Unique Identifiers (GUIDs) [20].

When the graphics cards in the cluster nodes cannot be synchronized at the hardware level, the Swap Lock Plug-in addresses the need for inter-node card synchronization at a software level. In abstract terms, the plug-in creates a software barrier by sending signals between the cluster nodes. The plug-in uses a master/slave paradigm where each slave sends a signal to the master and waits to swap its frame buffers. The master responds with a "go" message immediately before it invokes the frame buffer swap operation. Communication latency is a major factor when the master sends the signals to the slaves because the swap should begin simultaneously on all nodes. Implementations of this concept have used different communication channels for sending the signals with each channel having a corresponding plug-in implementation. Depending on how the nodes are connected, there are three versions of the Swap Lock Plug-in available: TCP swap lock, parallel port swap lock, and hybrid TCP/serial port swap lock.

When using the input data sharing clustering technique, we must ensure that all the application instances begin their frame loop at the same time. The Start Barrier Plug-in makes this possible through a master/slave paradigm similar to the Swap Lock Plug-ins. When each slave is ready to begin its frame loop, it sends a message to the master and waits for a response. When the master has received all the messages from the slaves, it sends the responses to the slaves. At that point, all nodes may begin their frame loop, thereby guaranteeing that all nodes begin their execution on the same frame.

4 Performance Evaluation

In this section, we present the performance evaluation of the VR Juggler clustering infrastructure. Instead of analyzing the efficiency of our clustering platform by using simple and well-known 3D models [17], we have performed our experiments on real VR

applications. We selected the following four representative VR applications that cover the spectrum of graphics-intensive and computationally intensive workloads: “Cubes,” “Agua,” “Hindu,” and “MPApp.”

Cubes is an extremely simple application where 1000 cubes are drawn floating in the space. The application generates low levels of rendering and computational workload. The next application, Agua, takes advantage of special hardware techniques, such as vertex shading [6], in order to recreate real-time travel around a complete deep-sea reef. While Agua is not computationally intensive, its extensive use of the graphics card capabilities makes it highly graphics intensive. The third application, Hindu, is a virtual walk through which allows users to explore the Radharaman Temple (Vrindavan, India). This application uses a set of animated virtual characters in order to perform a traditional religious ceremony inside the temple. Hindu is graphically intensive due to the large amounts of polygons and textures (for both temple and characters), and it is computationally intensive due to the movement and shadow generation for all characters during each frame. Finally, on the opposite extreme of the application spectrum, MPApp performs a real-time simulation of a square piece of cloth represented as a simple mesh surface. The mesh is generated by a complex, computationally intensive 3D polynomial equation, but it requires little graphical workload. Figure 2 depicts different snapshots of the four chosen VR Juggler applications using a single-node configuration.

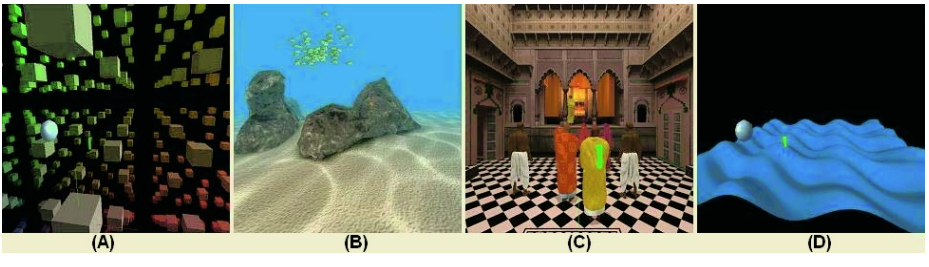


Fig. 2. Snapshots obtained from VR Juggler for: (a) Cubes (b) Agua (c) Hindu and d) MPApp

Our test-environment is composed of eight desktop computers, each running RedHat 8.0 with a NVIDIA GeForce3 Ti200 (128 MB of RAM) graphics card, a 2 GHz Intel Pentium4 Processor, 1 GB of RAM, and 512 KB of cache memory. The machines are connected to a Cisco Catalyst 3750 Gigabit Ethernet switch.

Since we have considered the system throughput as the number of frames per second (FPS) performed by the graphics cluster, we have measured this parameter by varying both the number of nodes and the network bandwidth in the cluster. The results of these variations are shown in Fig. 3a and Fig. 3b, respectively. The Y-axis of both figures show FPS values for the simulations performed with each system configuration. Each point in both plots represents the average FPS obtained after 25 executions of the same application benchmark. The standard deviation for any of the points shown in the plots was not higher than 4 FPS in any case.

Figure 3a shows the FPS reached by VR Juggler depending on the number of nodes in the cluster. This figure shows the number of nodes ranging from one to eight on

the X-axis. Figure 3a shows that the FPS is almost linearly reduced as more nodes are added to the cluster. Moreover, the linear factor of the throughput reduction decreases with the workload generated by the application in a single-node configuration. In this sense, applications such as Hindu or MPApp only have an average reduction of nine and seven FPS when they are ported from a single- to an eight-node configuration. This behavior is related to the linear network overhead that is incurred as new nodes are added to the cluster system, which is caused by the master/slave configuration of the Swap Lock Plug-in.

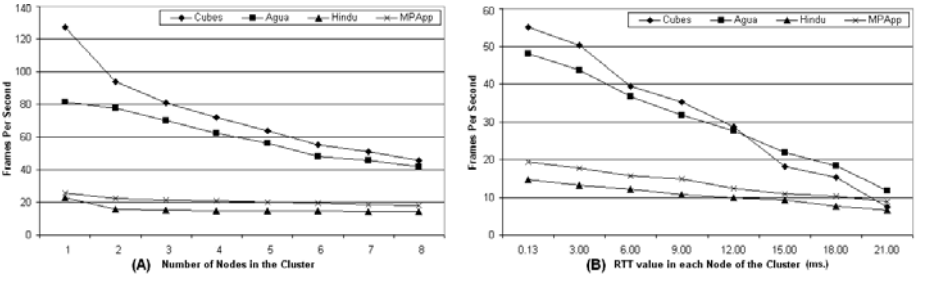


Fig. 3. Values of system throughput (FPS) for different (A) number of nodes and (B) network bandwidth

Analysis of VR clusters often lacks an investigation into the relationship between network bandwidth and system throughput, even though this has been shown to be an important topic [13]. In order to study this relationship, we have taken advantage of a software tool called *Netem* [12] (Network Emulator). It is a general-purpose tool for emulating bandwidth-limited links in real time. By operating at the IP level, Netem can emulate the critical end-to-end performance characteristics imposed by various wide area network (WAN) situations or by various underlying subnetwork technologies (e.g., Ethernet, Fast Ethernet, cable modems). Netem allows each node of the cluster to ensure a non-uniform round trip time (RTT) value for the transmitted TCP packets according to the specifications of the subnetwork technology. Since the correlation between RTT delay and the type of the physical network connection has been widely described in the literature of cluster computing [4][5], Netem becomes an excellent tool for emulating performance dynamics in our test environment.

Figure 3b shows the performance results when different RTT values are considered in a six-node configuration. The first value on the X-axis (0.13 ms) corresponds to the case in which no delay was added to all packets going out of the local Ethernet. This case shows the effective (and minimum) RTT value obtained in this configuration composed of six nodes and based on a Gigabit Ethernet backbone. In order to decrease the network throughput of the system, the above values correspond to the situations where Netem is used. While the main goal of our study was to determine the performance of VR Juggler clustering in LAN configurations, a variety of delays were considered. This figure shows how the FPS decreases linearly as the communication link delay increases for all the considered benchmark applications. Unlike the above case, the FPS value

tends to converge towards a similar threshold level when high latencies are emulated for all the benchmark applications. In this situation, VR Juggler spends most of the rendering period waiting for the synchronization from the Swap Lock and the Start Barrier plug-ins. Figure 3b shows that VR Juggler has only a small impact on performance levels, in terms of FPS, for immersive applications launched on commodity LAN clusters. Since the RTT values in these systems are not higher than a couple of milliseconds [4], these results show that VR Juggler can be considered as an efficient tool to utilize multi-screen immersive visualization systems on a cluster of computers.

5 Conclusion and Future Work

In this paper, we have described the architecture and the performance evaluation of the VR Juggler clustering infrastructure. The open architecture of the clustering infrastructure has been specifically designed to allow VR application developers to combine various existing clustering techniques, both at the hardware level and at the software level, to meet their own specific needs. The VR Juggler clustering infrastructure allows VR application developers to combine various existing clustering techniques, both at the hardware level and at the software level, to meet their own specific needs. The design of the VR Juggler clustering infrastructure allows users to configure the software to meet the needs of their specific hardware, and its plug-in framework allows programmers to extend the clustering infrastructure with new features. In our case, the VR Juggler clustering support has allowed us to migrate existing applications (designed initially for high-end shared memory computers) to the newer cluster-based configurations while keeping high levels of frame-rate and without changes required to the application code.

As future work to be done, we plan to add the ability to use additional network protocols such as IP multicast. Since the current version of Cluster Network layer is limited to using point-to-point TCP connections, our intention is to provide further network efficiency beyond those domain-specific optimizations presented in the current implementation. The addition of a plug-in for monitoring the performance of the cluster at run time and for validating correct synchronization is also planned.

References

1. Jérémie Allard, Valérie Gouranton, Emmanuel Melin, and Bruno Raffin. Net juggler and softgenlock: Running vr juggler with active stereo and multiple displays on a commodity component cluster. In *IEEE Virtual Reality 2002 Proceedings*, Orlando, Florida, United States, March 2002.
2. Allen Bierbaum and Carolina Cruz-Neira. Run-time reconfiguration of vr juggler. In *Proceedings of IPT 2000*, Ames, Iowa, United States, June 2000.
3. Allen Bierbaum, Christopher Just, Patrick Hartling, Kevin Meinert, Albert Baker, and Carolina Cruz-Neira. VR Juggler: A virtual platform for virtual reality application development. In *Proceedings of IEEE Virtual Reality*, pages 89–96, Yokohama, Japan, March 2001.
4. Richard Charles Booth. A system area network characterization in a commercial cluster. Master's thesis, University of Minnesota, 1998.
5. Mark Carson and Darrin Santay. NIST Net: A linux-based network emulation tool. *ACM SIGCOMM Computer Communication Review*, 33(3):111–126, 2003.

6. Wolfgang Engel. *Programming Vertex and Pixel Shaders*. Programming Series. Charles River Media, September 2004.
7. Message Passing Interface Forum. MPI: A message-passing interface standard. *International Journal of Supercomputer Applications and High Performance Computing*, 8(3/4):165–414, Fall/Winter 1994.
8. Greg Humphreys, Mike Houston, Ren Ng, Randall Frank, Sean Ahern, Peter D. Kirchner, and James T. Klosowski. Chromium: A stream processing framework for interactive graphics on clusters. In *ACM SIGGRAPH 2002 Sketches and Applications*, San Antonio, Texas, United States, July 2002. ACM Press.
9. John Kelso, Lance Arsenault, Steven Satterfield, and Ronald Kriz. Diverse: A framework for building extensible and reconfigurable device independent virtual environments. In *IEEE Virtual Reality 2002 Proceedings*, pages 183–190, Orlando, Florida, United States, March 2002.
10. Kai Li, Han Chen, Yuqun Chen, Douglas W. Clark, Perry Cook, Stefanos Damianakis, Georg Essl, Adam Finkelstein, Thomas Funkhouser, Allison Klein, Zhiyan Liu, Emil Praun, Rudrajit Samanta, Ben Shedd, Jaswinder Pal Singh, George Tzanetakis, and Jiannan Zheng. Early experiences and challenges in building and using a scalable display wall system. *IEEE Computer Graphics and Applications*, 20(4):671–680, 2000.
11. Eric Charles Olson. ClusterJuggler – PC cluster virtual reality. Master's thesis, Iowa State University, Ames, Iowa, 2002.
12. Netem: Network Emulator Home Page. http://developer.osdl.org/shemminger/netem/.
13. Aske Plaat, Henri E. Bal, and Rutger H. F. Hofman. Sensitivity of parallel applications to large differences in bandwidth and latency in two-layer interconnects. In *Proceedings 5th IEEE HPCA '99*, pages 244–253, Orlando, Florida, United States, January 1999.
14. Marcus Roth, Gerrit Voß, and Dirk Reiners. Multi-threading and clustering for scene graph systems. *Computers & Graphics*, 28(1):63–66, February 2004.
15. Benjamin Schaeffer and Camille Goudeseune. Syzygy: Native pc cluster vr. In *IEEE Virtual Reality 2003 Proceedings*, pages 15–22, Los Angeles, California, United States, March 2003.
16. Sandeep Singhal and Michael Zyda. *Networked Virtual Environments: Design and Implementation*. Addison-Wesley, 1999.
17. Oliver G. Staadt, Justin Walker, Christof Nuber, and Bernd Hamann. A survey and performance analysis of software platforms for interactive cluster-based multi-screen rendering. In *Proceedings of the Workshop on Virtual Environments 2003*, pages 261–270, Zürich, Switzerland, 2003. ACM Press.
18. Paul Strauss and Rikk Carey. An object-oriented 3d graphics toolkit. In *Proceedings of 19th ACM SIGGRAPH*, pages 341–349. ACM Press, August 1992.
19. Clemens Szyperski, Dominik Gruntz, and Stephan Murer. *Component Software: Beyond Object Oriented Programming*. Component Software Series. Addison-Wesley Publishing Company, New York, NY, second edition, 2002.
20. The Open Group. *DCE 1.1: Remote Procedure Call*. The Open Group, August 1997.

Adaptive Space Carving with Texture Mapping

Yoo-Kil Yang[1], Jung Lee[2], Soo-Kyun Kim[2], and Chang-Hun Kim[2,⋆]

[1] Dept. of Visual Information Processing, Korea University,
ykyang@software.or.kr
[2] Dept. of Computer Science and Engineering, Korea University,
{airjung, nicesk, chkim}@korea.ac.kr

Abstract. Space carving reconstructs a 3D object from multiple images, but existing algorithms rely on a regular grid which makes poor use of memory. By using the image information, adaptive space carving uses a recursively generated structure which reduces memory requirements and thus allows a finer grid. After reconstruction, models are triangulated to facilitate texture mapping. Experimental results show the enhanced appearance of models reconstructed in this way.

1 Introduction

Several methods are available to create realistic models of 3D object, including range scanning, 3D modeling tools, and image-based modeling. Range scanning requires expensive capture devices, and 3D modeling tools need a great deal of effort and skill. Image-based modeling approach produces realistically colored models, by mapping textures from the source images onto the recovered shape although the geometry of the model is not exactly same as that of the original object. This method is relatively effective in terms of time and cost. There are various approaches to image-based modeling, as follows.

Seitz and Kutulakos [5] used a plenoptic function to modify virtual 3D objects and propagate these modifications to all the source images. Its basic concept is derived from the approach of Adelson and Bergen [1]. In this paper, the concept of 'ordinal visibility' is introduced to improve and simplify the visibility computation.

Matusik and et al. [3] used epipolar geometry to reconstruct the visual hull of the target object. This is subsequently rendered using silhouette cones from the corresponding source images. The results look photorealistic, but this technique is only suitable for rendering because no geometry is reconstructed.

Seitz and et al. [4, 6] proposed a space carving algorithm that recovers 3D volumetric models. This work is free from vision constraints such as occlusion, and recovers the maximal photo-consistent shape that accounts for the source images. However two important problems exist in aspect of rendering quality and memory management. These problems stem from the voxel quantization

⋆ The corresponding author: Tel. +82-2-3290-3199; Fax. +82-2-953-0771.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1129–1138, 2005.

that occurs when a voxel is larger than a pixel. Due to the color interpolation process that has to be performed, the resulting pictures are so blurred appearance and details are lost as shown in Fig. 1. This should be obvious from the figure. To reduce the loss of quality, smaller voxels can be used. However as the voxel size is decreased, the number of voxels in the initial static volume increases exponentially. This problem cannot be ignored when using a system with limited memory capacity.

Fig. 1. Details lost by the space carving technique [6]. Left: source image, right: rendered result from the same viewpoint

To solve these problems, various modified carving methods have been proposed.

Vedula and et al. [7] added motion parameters to the carving criteria. Using a hexel structure, they recover motion flow as well as geometry. However, the voxel quantization problem is not addressed.

Broadhurst and Cipolla [2] suggest that the voxel quantization problem is mainly a result of using a simple voxel centroid sampling method which cannot represent the color distribution on the voxel's projected area exactly. They improved the rendered result by using a statistical sampling method (F-ratio). But this approach only reduces the voxel quantization problem a little.

We think that an effective solution of the quantization problem is to make the voxels smaller which also changing the rendering method so that it does not need to use the color interpolation. By carving the target object adaptively, we can eliminate a large portion of empty space early in the process. The memory saved in this way can be used in generating smaller voxels.

Additionally, we apply a texture mapping that converts reconstructed voxels into polygonal cubes. This technique can largely avoid quantization effects. Our approach offers two main contributions.

• By applying a space carving method adaptively, a large amount of memory is saved and smaller voxels can be used.
• By using texture mapping, color interpolation process is avoided during the rendering of the reconstructed shape.

The remainder of this paper is organized as follows. Section 2 introduces some assumptions. Section 3 describes 'photogrammetric occupancy and emptiness'. The detailed reconstruction process is explained in Section 4. Experimental results and comparisons are shown in Section 5 and we conclude this paper in Section 6.

2 Assumption

In this section, we briefly introduce the two major assumptions that underlie our method. In most respects, they are similar to those of previous space carving approaches.

2.1 Source Images

Like most image-based approaches, we obtain camera-calibrated images which are background-segmented before starting the reconstruction. In this paper, we use synthetic images whose background color is white because we only want to evaluate the efficiency and accuracy of our space carving method. Silhouette information is used to distinguish the foreground from the background (See Fig. 2).

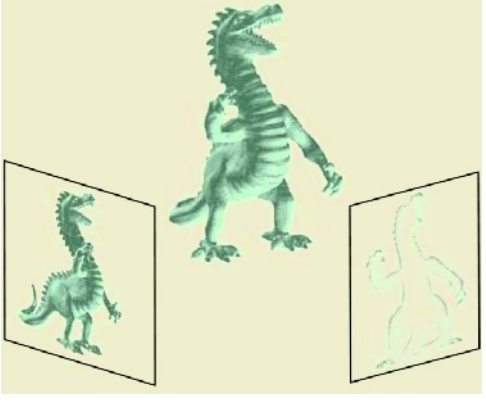


Fig. 2. Source images acquisition. Left: the source image itself. right: silhouette information

2.2 Cubic Shape of Voxels

We are carving volumetric space to recover a 3D shape. Using cubic voxels makes the carving and triangulation processes much more straightforward. After the carving process, the voxels are triangulated into cubes.

3 Adaptive Space Carving

3.1 Photogrammetric Occupancy and Emptiness

To determine whether each voxel is inside or outside the target object, we introduce the concept of photogrammetric occupancy and emptiness, based on projection of voxels.

Theorem: If a voxel is completely projected onto the background area, it resides outside the target object (emptiness condition).

The theorem above can be also restated as its converse: if a voxel is projected onto an area containing at least one foreground pixel, it may be inside or on the target object (occupancy condition).

As can be seen in Fig. 3, only voxels C and D intersect the surface of the target object. Camera C1 determines voxel A is empty because it projects entirely onto background pixels in the image plane of C1. Voxel B is partially projected onto the foreground pixels, so it is evaluated to be on the surface of the object. If we add a second camera C2, it becomes apparent that voxel B is empty. Using the approach of Broadhurst and Cipolla [2], the front face of each voxel is projected to calculate its influence on each image. But, we also project the side faces, and determine the occupancy and emptiness of each primitive, not just its influence. If we projected only the front face, then voxel C would be seen as empty by camera C2 because its front face is entirely projected on the background pixels of C2.

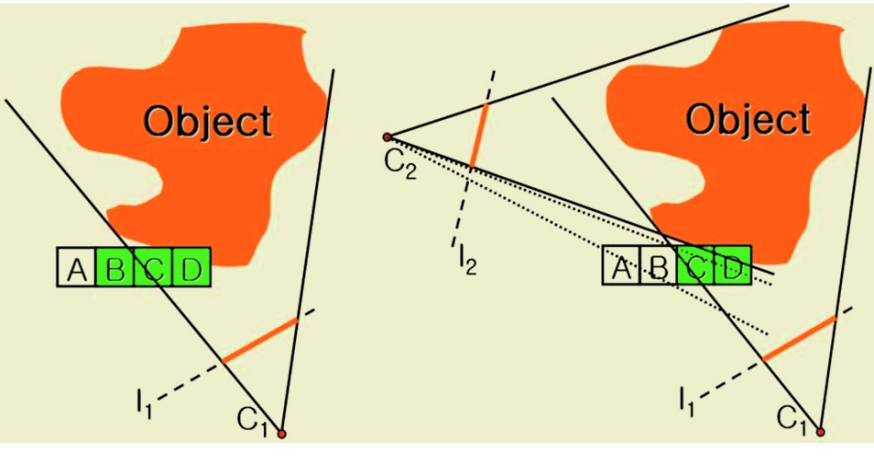


Fig. 3. Photogrammetric occupancy and emptiness. The dotted lines show the boundaries of the background area. The solid lines show the boundaries of the foreground area. Left: carving from camera C1, right: carving from both cameras C1 and C2

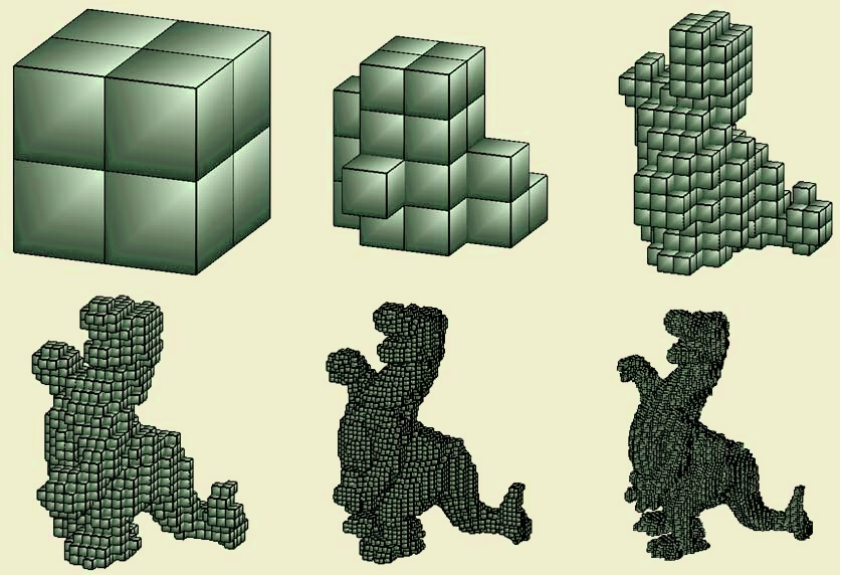


Fig. 4. Adaptive carving of the Dragon model. top left: level 1, top middle: level 2, top right: level 4, bottom left: level 5, bottom middle: level 6, bottom right: level 7

3.2 Adaptive Space Carving

Previous approaches to reconstructing a 3D shape from an initial static volume use a uniform grid of voxels. But the number of voxels that compose the final reconstructed shape is small in comparison with the initial volume. By detecting large empty spaces at an early stage, a lot of memory can be saved to achieve a denser voxel distribution. The octree structure is effective for dividing space adaptively. The original octree algorithm adaptively divides the space based on the complete geometry information of the object. In our approach, the geometry of the object is the goal to be reconstructed and we have no information about that, except for its images. In this section, we explain how we divide and construct the shape adaptively by using given source images.

Our carving process is divided into two steps. First, we must determine whether each voxel is to be carved or not, by placing one large voxel that completely covers the target object (see Fig. 4). Then we determine whether each voxel is inside or on the surface of the target object based on the photogrammetric occupancy and emptiness test. If a voxel is found to be outside the target object, it is eliminated from the space. Each remaining voxel is subdivided into eight smaller voxels, and their occupancy and emptiness are investigated. This process proceeds recursively until the allocated memory has been used. The detailed algorithm of this process is described as follows.

Fig. 4 shows the adaptive carving of a Dragon model at successive levels. The shape becomes closer to the target model as adaptive construction proceeds.

4 Polygonal Conversion of Voxels for Texture-Mapping

4.1 Polygonal Conversion of Voxels

The remaining voxels are converted to triangulated cubes before texture mapping. The triangulation process is simple. The six rectangular faces of each voxel are triangulated into twelve triangles.

4.2 Visibility Calculation and Texture Mapping

To eliminate redundant interior triangles and apply texture mapping, we calculate the visibility of all the triangles that come from the converted cubes, using a depth-marking method similar to the plane-sweep visibility technique of Seitz and Kutulakos. But we do not need to consider its ordinal visibility condition, because the created structure does not change after the carving process has been performed. The visibility calculation uses a depth buffer. If a triangle is visible from a certain camera, its depth value will be same as that of the corresponding pixel position in the depth buffer. Using this property, we can robustly and simply obtain visibility information for all the triangles.

After triangulation, we map the source images onto the triangulated model. Before assigning the texture coordinates to each triangle, we must determine the texture which will represent the most accurate color. The quality of the resulting image is heavily reliant on the selection of an appropriate texture. Like most image-based approaches, we select the image that is closest to the front

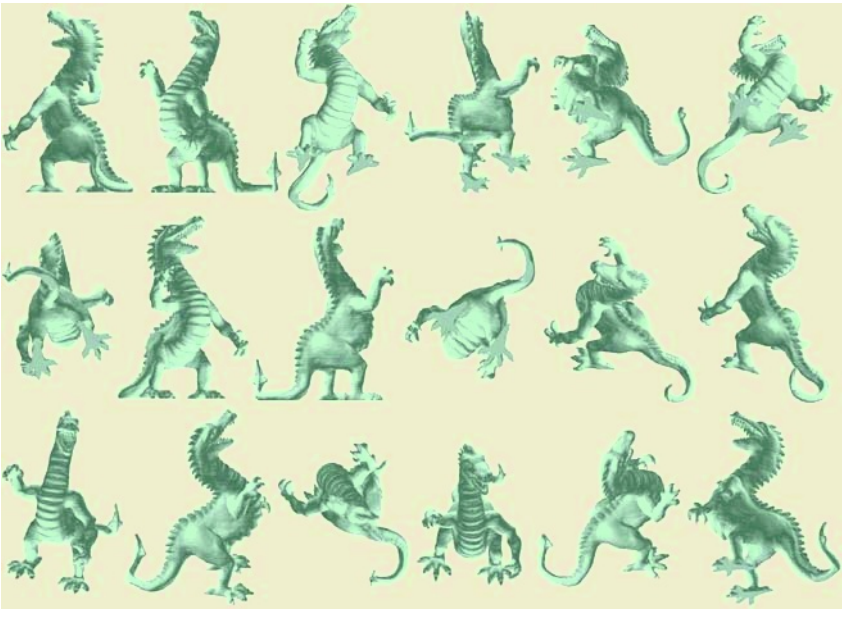


Fig. 5. 18 source images of the Dragon model

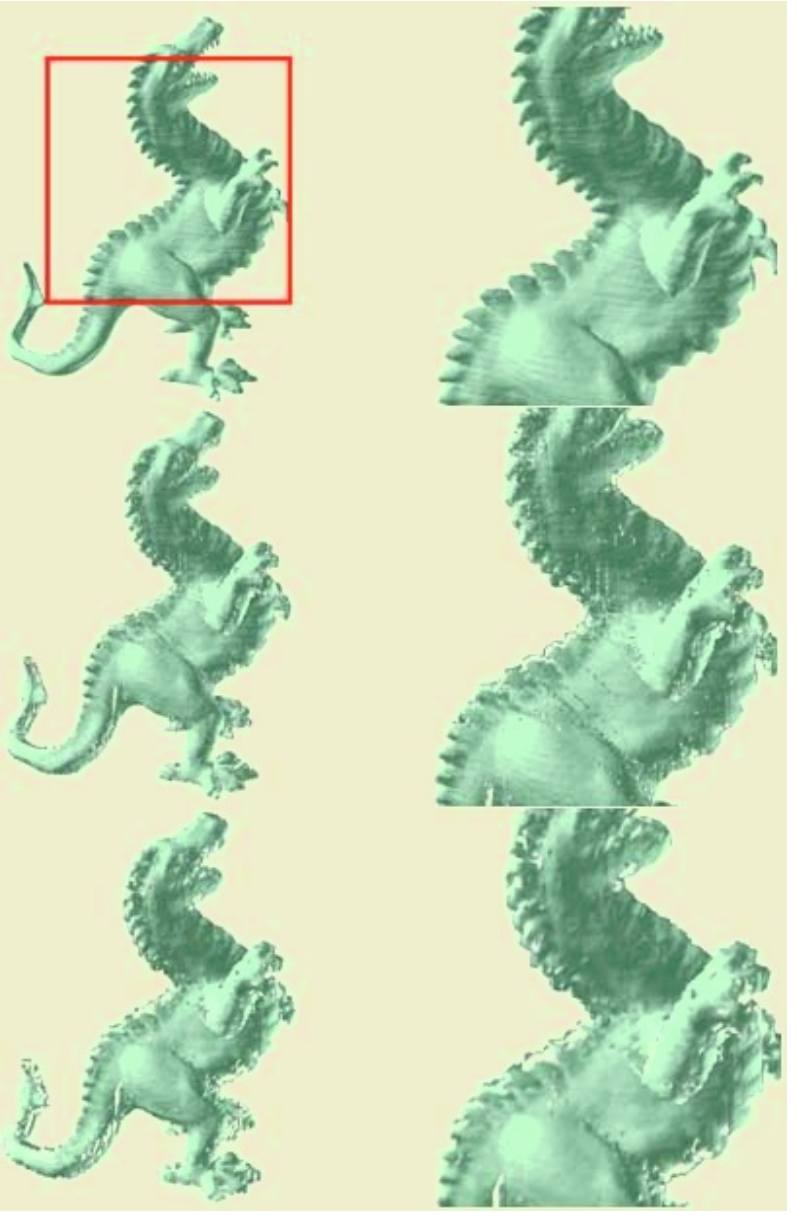


Fig. 6. Reconstruction of the Dragon model using our method and the previous voxel-carving method. Top row: source image, middle row: our reconstruction, bottom row: reconstruction by previous method [6]

view of each triangle. The texture coordinates are calculated by projecting the vertices of each triangle onto the selected texture.

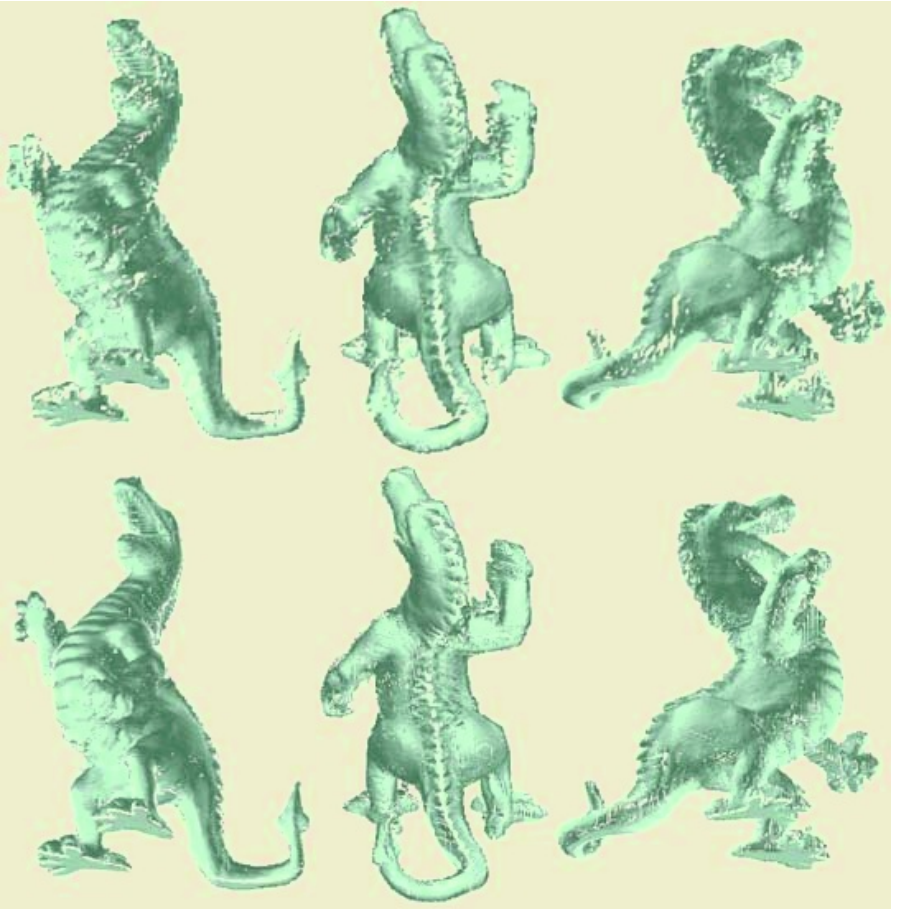


Fig. 7. Reconstruction of the Dragon model rendered from arbitrary viewpoints. Top row: results by previous method [6], bottom row: our results

5 Results and Discussion

Experimental results and comparisons with previous voxel-carving approaches are presented in this section. All the experiments used 18 source images, and were performed on a PC with a Pentium IV 2.8 GHz CPU, 256 MB memory, and an ATI Radeon 9800 3D-accelerated graphic card.

Fig. 5 shows the 18 source images that we used to reconstruct the result of Figs. 6 and 7; the camera parameters were already known. Fig. 6 compares the results obtained using our method and previous voxel-carving approach [6] at the same viewpoint of the source image. The white noises in the rendered image appear due to the exterior part of the surface-intersecting voxels.

Table 1 shows their efficiency and accuracy by pixel-by-pixel comparison between the source images and the reconstruction results rendered at the view-

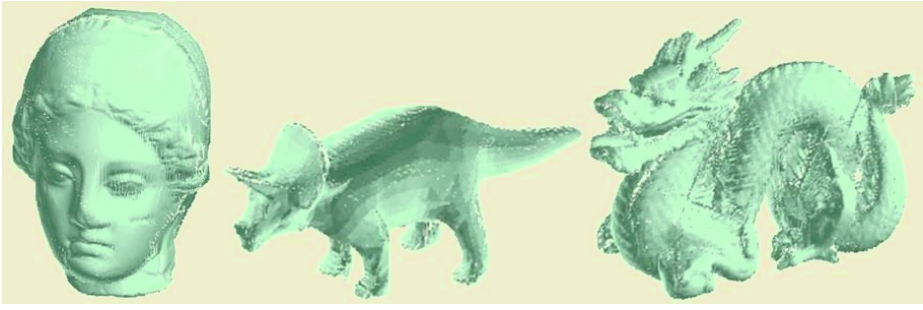


Fig. 8. Various models reconstructed by our method. Left: 'Igea' model, middle: 'triceratops' model, right: 'Chinese dragon' model

Table 1. Comparison of the reconstructed Dragon model between our method and the previous method [6]

	Memory (MBs)	Voxel Dimension	Reconstruction Time (sec)	Average RGB Error
Previous Method	127	128x128x128	97	(35, 27, 27)
Our Method	106	256x256x256	44	(25, 20, 20)

points of the corresponding images. Although the previous method requires the larger amount memory than ours, its result produces severe aliasing and blurred regions due to its bigger voxel. To carve the shape out of the initial volume with the same voxel size of ours, it costs 806 MBs memory.

Fig. 7 compares the results at arbitrary viewpoints that are not included in the source images. As seen in this figure, our reconstruction shows more photorealistic appearance than the previous one. Additionally, our result shows the plausible appearance even if it is magnified.

Fig. 8 shows the results of various models rendered at arbitrary viewpoints. Even with the complex shape such as a Chinese dragon model, our method reconstructs the detailed description of its shape and appearance.

6 Conclusion

We have proposed an effective space carving algorithm that considerably reduces the artifacts of previous voxel-carving approaches. Adaptive carving saves a large amount of memory by eliminating a large portion of empty space at an early stage. The saved memory can be used for a denser voxel distribution. Our final reconstructed shape corresponds to the source images well and shows more photorealistic appearance than the previous voxel-carving approaches.

In a future work, we plan to study techniques for reducing the aliasing artifacts in rendered images, and to explore various applications of our approach.

Acknowledgments. This work was supported by University IT Research Center Project and the Basic Research Program of the Korea Science and Engineering Foundation (No. R01-2002-000-00512-0).

References

1. Adelson, E. H., Bergen J. R. The plenoptic function and the elements of early vision. Computational Models of Visual Processing (1991), pp. 3-20.
2. Broadhurst A., Cipolla R. A statistical consistency check for the space carving algorithm. In Proc. 11th British Machine Vision Conference (2000), pp. 282-291.
3. Matusik W., Buehler C., Raskar R., Gortler S. J., McMillan L. Image-based visual hulls. In Proc. SIGGRAPH '00 (2000), pp. 369-374.
4. Seitz S. M., Dyer C. R. Photorealistic scene reconstruction by voxel coloring. International Journal of Computer Vision (1999), pp. 151-173.
5. Seitz S. M., Kutulakos K. N. Plenoptic image editing. In Proc. 6th International Conference of Computer Vision '98 (1998), pp. 17-24.
6. Seitz S. M., Kutulakos K. N. A theory of shape by space carving. International Journal of Computer Vision (2000), pp. 199-218.
7. Vedula S., Baker S., Seitz S. M., Kanade T. Shape and motion carving in 6D. In Proc. IEEE International Conference on Computer Vision and Pattern Recognition (2000), pp. 592-598.

User-Guided 3D Su-Muk Painting

Jung Lee, Joon-Yong Ji, Soo-Kyun Kim, and Chang-Hun Kim*

Department of Computer Science and Engineering, Korea University
{airjung, guybrushji, nicesk, chkim}@korea.ac.kr

Abstract. We present a technique for rendering animated 3D models in a Su-Muk painting style with user's guide. First, a user can sketch directly over 3D models by varying ink and water values. And we simulate the behavior of ink-water directly on 3D model for consistent shade information over 3D models. After simulation, ink and water are spread over a 3D surface appropriately. Second, to achieve a real hand-crafted look of Su-Muk painting, ink-water behavior is simulated again on a 2D screen by using overall ink-water shade information from previous output of 3D ink-water simulation. We demonstrate some images for Su-Muk painting with our user-guided drawing system.

1 Introduction

Su-Muk painting is Oriental black-ink painting, and a traditional art form in China, Japan, and Korea. It is a simple and elegant art that produces black images on white paper. Until now many researchers have investigated on various traditional art form, e.g., pen and ink, watercolor, charcoal, etc. in addition to Su-Muk painting.

The early research on Su-Muk rendering focused on the simulation of brush work. Strassmann swept a one-dimensional texture to achieve shading[1]. Pham modeled a brush stroke based on a variable offset approximation of uniform cubic B-splines[2]. Lee modeled a brush as a collection of rods with simulated elasticity[3]. And until now, many researchers have been researching on brush work[8, 7, 9] Though elaborate and realistic Su-Muk images can be generated by these 2D interactive drawing systems, they can not support 3D animation with temporal coherence.

Su-Muk rendering is now being extended to 3D animation. Chan utilized existing rendering packages to animate Chinese painting by means of procedural shaders[4]. Kang and Kim used sphere mapping and 2D image processing techniques within the graphics processing unit to generate 3D Su-Muk images in real-time[5].

Zhang et al. simulated the behavior of ink-water using a cellular automaton in 2D screen[6]. This system is simple to implement, but also produces realistic results. And they apply this simulation model to 3D tree animation in Su-Muk

* The corresponding author: Tel. +82-2-3290-3199; Fax. +82-2-953-0771.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1139–1147, 2005.

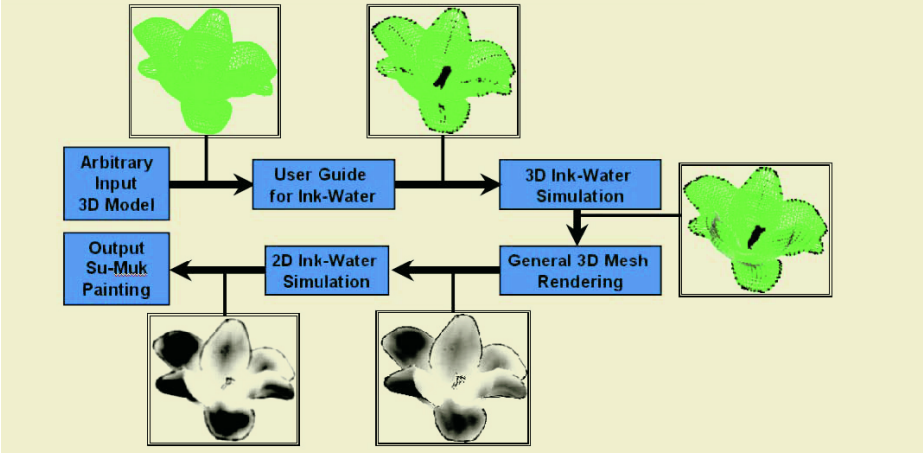


Fig. 1. Overall process of proposed Su-Muk system

style. But this method is only focused on tree models which are easy to shade, not all general and various models. But unlike other painting style, e.g., cartoon, oil painting, there are no standard or simple technique to draw various models. So it is hard to render general and various models in Su-Muk style.

In this paper, we present a method that a user can generate 3D animation of various 3D models in elaborate Su-Muk painting style very easily. First, we simulate behavior of ink and water directly on 3D mesh from the values which are given from a user. After simulation, ink and water are spread over a 3D surface appropriately. Therefore a user can imitate brush stroke intuitively by just determining appropriate ink and water values and areas on 3D vertices. And also ink and water information about how to draw object are already stored in 3D mesh, we can keep consistent stroke style information all over different views. Second, we simulate again behavior of ink and water on 2D screen from the values which are given from above 3D simulation by using Zhang's method[6]. This is due to the fact that ouputs of 3D model from 3D ink-water behavior model are solid-looking. Because ouputs of 3D model are rendered by using vertices color in general 3D rendering pipeline. Therefore through this step, we can generate important features of Su-Muk style such as ink diffusion and paper effect pretty realistic. Figure 1 shows process flow of our simulation model.

2 Ink-Water Behavior Model on 3D

We assume that input to our method is 3D triangle mesh which has connectivity information. Ink-water flows along the edges of the 3D polygonal model and remains at its vertices. When simulation on 3D mesh is complete, each cell contains a certain amount of ink and water. We set a color value for each vertex that corresponds to the number of the ink particles in that cell. More particles reduce the brightness, within a range suitable for the rendering software. Also,

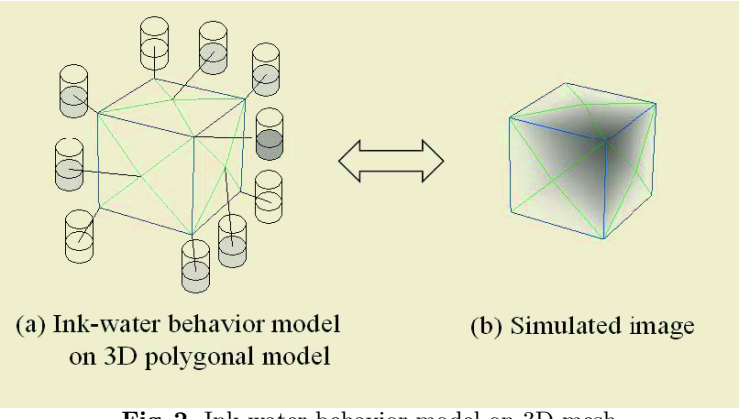


Fig. 2. Ink-water behavior model on 3D mesh

we set an alpha value for each vertex appropriately for alpha blending: The higher the brightness, the greater the transparency. Using these values, we render the 3D polygonal model using software which supports color interpolation, hidden surface removal, and alpha blending. In our paper, we use this scheme for maintaining consistent stroke(shade) information over 3D surfaces and allowing a user draw stroke very easily on 3D mesh. Figure 2 shows our 3D simulation model and its rendering method.

We will now show how to extend the 2D cellular automaton model proposed by Zhang [6] to simulate the behavior of ink-water directly on a 3D polygonal model.

2.1 Definitions of Ink-Water Behavior Model

A cellular automaton consists of an infinite, regular grid of cells, each in one of a finite number of states. Time is also discretized, and the state of a cell at time t is a function of the state of a finite number of cells, called the neighborhood of that cell, at time $t-1$. Every cell has the same rule for updating, based on the values in its neighborhood. Each time the rules are applied to the whole grid a new generation of the automaton is produced.

A 3D polygonal model consists of a list of vertices $V := \{v_o | 1 \leq o \leq n_v\}$, where n_v is the total number of vertices, and a list of edges $E := \{f_o | 1 \leq o \leq n_e\}$, where n_e is the total number of edges. The set of vertices that share a edge with v_o will be denoted V^o. If the vertex v_k belongs to V^o, we say that v_k is in the neighborhood of v_o. We can then construct a cellular automaton $C := \{c_o | 1 \leq o \leq n_v\}$, where c_o is a cell of the cellular automaton and corresponds to the vertex v_o. The set of cells c_k which correspond to the vertices v_k which belong to V^o will be denoted C^o. If the cell c_k belongs to C^o, we say that cell c_k is in the neighborhood of c_o. The number of elements of the set C^o will be denoted by $|C^o|$.

There are two state variables in each cell c_o. One is the number of 'water particles', denoted by W_o, and the other is the number of 'ink particles' dissolved in the water, denoted by I_o. To simulate the paper texture, each cell is considered as a 'water tower'. The capacity of the tower is T_o, and its height of bottom is

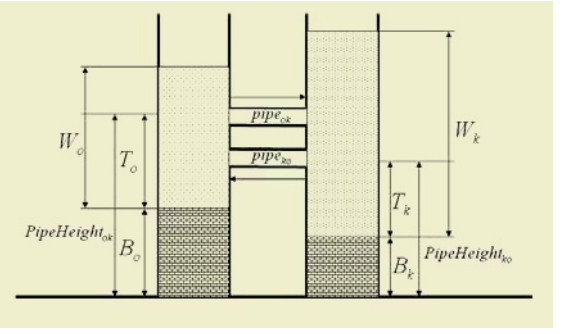


Fig. 3. Cellular automaton structure

B_o. Together, these determine the maximum quantity and the potential energy of the ink and water stored in that cell.

2.2 Modeling the Transfer and Diffusion of Ink-Water

Let $S_o(t)$ be a state variable of cell c_o at time t. We define it to be:

$$S_o(t) := (W_o(t), I_o(t))$$

The state is updated by the function f:

$$S_o(t) := f(S_o(t-1), S_k(t-1))$$

S_k are the state variables of the cells c_k which belong to C^o.

Transfer of Water Particles. Through $pipe_{ko}$ water and ink can flow from cell c_k to cell c_o; through $pipe_{ok}$ vice versa. Each pipe has a 'height' denoted by $PipeHeight_{ko}$ and $PipeHeight_{ok}$, and defined (see Figure 3) as:

$$PipeHeight_{ko} = \max(B_o, B_k + T_k)$$
$$PipeHeight_{ok} = \max(B_k, B_o + T_o)$$

We denote the amount of water transferable from cell c_k to cell c_o by ΔW_{ko}, and ΔW_{ok} vice versa. ΔW_{ko} is determined by the difference in water between cell c_o and its neighborhood cell c_k, as shown in Figure 3. ΔW_{ok} is calculated in the same way as we define ΔW_{ko}. We define ΔW_{ko} and ΔW_{ok} as:

$$\Delta W_{ko} = \max(0, \alpha \cdot 1/|C_o| \cdot \min((B_k + W_k) - (B_o + W_o), (B_k + W_k) - PipeHeight_{ko})$$
$$\Delta W_{ok} = \max(0, \alpha \cdot 1/|C_o| \cdot \min((B_o + W_o) - (B_k + W_k), (B_o + W_o) - PipeHeight_{ok})$$

α : the transfer coefficient for water particles.

The number of water particles in cell c_o at time t:

$$W_o(t) = max(0, W_o(t-1) + \sum_{k|c_k \in C^o} (\Delta W_{ko}(t-1) - \Delta W_{ok}(t-1)))$$

Transfer of Ink Particles Accompanying Water Particles. The number of ink particles which are transferred with the water depends on the quantity of water and the concentration of the ink. ΔI_{ko} is the number of ink particles in ΔW_{ko}, and ΔI_{ok} is simultaneously defined. The number of ink particles in cell c_o at time t is obtained in a similar way to the number of water particles:

$$\begin{aligned}\Delta I_{ko}(t-1) &= \Delta W_{ko}(t-1) \cdot (I_k(t-1)/W_k(t-1)) \\ \Delta I_{ok}(t-1) &= \Delta W_{ok}(t-1) \cdot (I_o(t-1)/W_o(t-1)) \\ I_o(t) &= I_o(t-1) + \sum_{k|c_k \in C^o} (\Delta I_{ko}(t-1) - \Delta I_{ok}(t-1))\end{aligned}$$

Transfer of Ink Particles to Balance the Concentration. Ink diffusion force results from the difference in density between the ink at cell c_o and that at its neighborhood cell c_k. We define ΔI_{dko} as the number of ink particles transferred from cell c_k to cell c_o due to this ink diffusion force.

$$\begin{aligned}\Delta I_{dko} &= \beta \cdot (I_k - W_k \cdot (I_o + I_k)/(W_o + W_k)) \\ I_o(t) &= I_o(t-1) + \sum_{k|c_k \in C^o} \Delta I_{dko}(t-1)\end{aligned}$$

β : the diffusion coefficient for ink

Evaporation of Water. We can simulate evaporation by subtracting some water particles from the water state variable.

$$W_o(t) = W_o(t-1) - \Delta W$$

where ΔW is the amount of water lost by evaporation

3 Ink-Water Behavior Model on 2D

The 2D simulation model is same as the 3D simulation model, except from each cell now corresponds to each pixel on 2D screen. So same rule can be derived as the case of 3D simulation model. After 2D simulation is over, we set pixel colors proportional to ink values which are remained in pixel. In our paper, we use this scheme for making images like hand-crafted look on 2D paper. Figure 4 show Zhang's 2D simulation model and its rendering method.

In addition, we use same Zhang's method to generate paper effect. We generate 2D line segments, with random position and direction, which they consider as fibers, and raise the water storage capacity of cells which lie under these line segments. This is achieved by increasing the capacity of the cell. So we can imitate paper texture by varying the water storage capacity of each cell. Figure 5 shows final results of 2D simulation system from 3D simulation system.

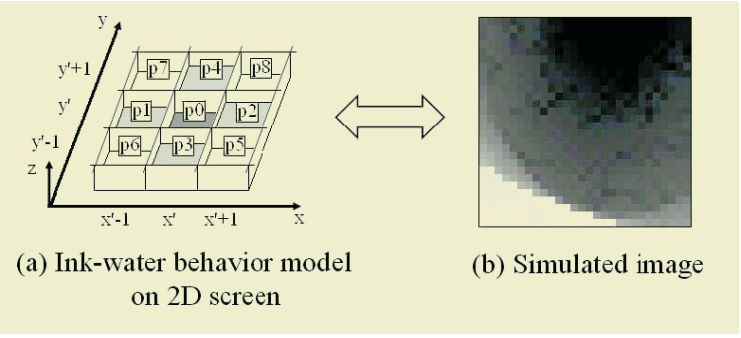


Fig. 4. Ink-water behavior model on 2D screen

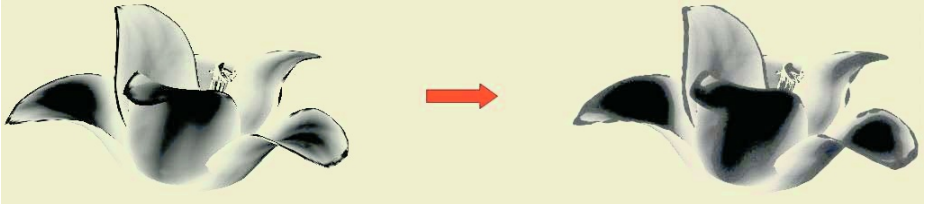


Fig. 5. Left: Result of 3D Simulation Model and Input of 2D Simulaiton Model, Right: Result of 2D Simulation Model

4 User-Guided Simulation

We now show how our 3D ink-water behavior model can be used as the basis of an user-guide drawing system. First we prepare input parameters interface, those are parameters of brush, paper, and simulation. The parameters that control the ink and paper properties, and the simulation itself, can be set by the user in the dialog box shown in Figure 6.

We can simulate various drawing situations by changing these control parameters. As the values of 'Fiber Ratio in Paper' and 'Fiber Length' increase, the characteristic texture of the paper is expressed more distinctly. We can use the brush stroke and simulation parameters to simulate various drawing styles. For example we draw outline boundaries and creases using low values of 'Water Particles' and 'Brush Width', and fill the interior of these outlines with washes created using larger values of 'Water Particles' and 'Brush Width'. At the same time, by changing parameters such as 'Water Transfer Coefficient', blurred brush strokes can be simulated.

Next a simple interactive device, such as a mouse, is used to draw directly on to the 3D polygonal model. The path of a brush stroke is composed of vertices of the 3D polygonal model. The user can toggle to a drawing mode in which edges, vertices and picked vertices are shown. Using existing rendering software which supports vertex selection algorithm, a vertex can be selected as a point on the

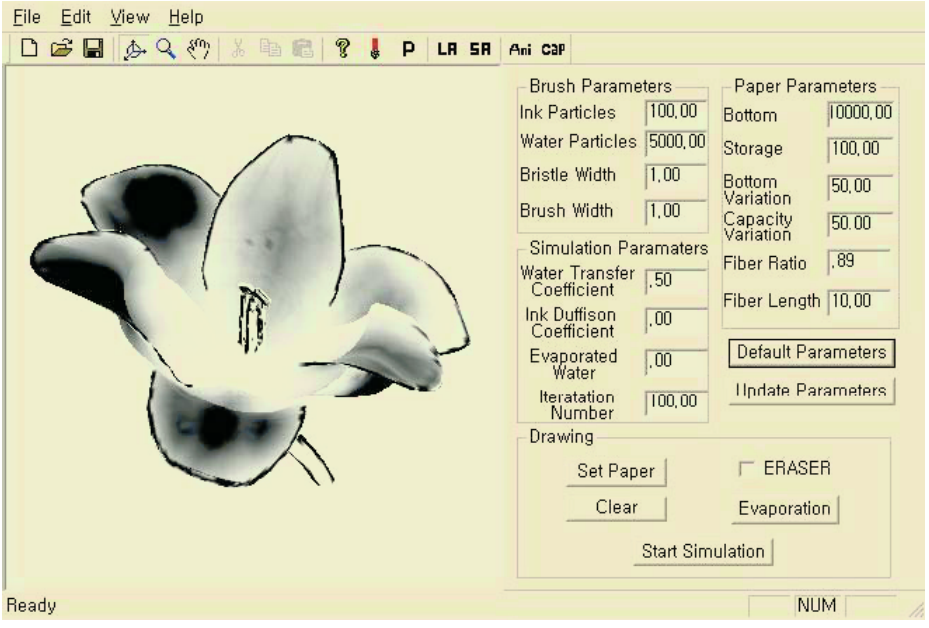


Fig. 6. Control parameters interface

path. As the value of the 'Brush Width' parameter increases, the selection region becomes wider. When a vertex is selected, the current values of 'Ink Particles' and 'Water Particles' are stored at that vertex. After all the brush strokes have been drawn on the 3D polygonal model, we start the simulation of ink and water behavior, and the water and ink diffuse over the vertices in a manner that reflects the parameters that have been inputed.

Fig. 7. Stages in drawing using our system

Figure 7 shows the stages in making a drawing using our system. The left most image shows the vertices of the 3D polygonal model in drawing mode. Brush strokes are shown in drawing mode as a bright vertices in the second image. The third image shows how the ink and water are attached to selected vertices before simulation. After simulation the ink and water have propagated across the 3D model in the right most image.

5 Results

Figure 8, 9 show various models rendered from different views. In figure 9, the user draws strokes along center lines of each petal with much water and boundaries with little water. Table 1 reports the complexity of the various models, the corresponding simulation times. In this table, We can find that 3D simulation time is proportional to the number of faces, but 2D simulation time is constant due to same 2D screen resolution. Our system consists of Pentium IV 1.7GHz, 640M RAM, nVIDIA GeForce4 Ti 4400 video card.

Fig. 8. Two views of a orchid model rendered with our system

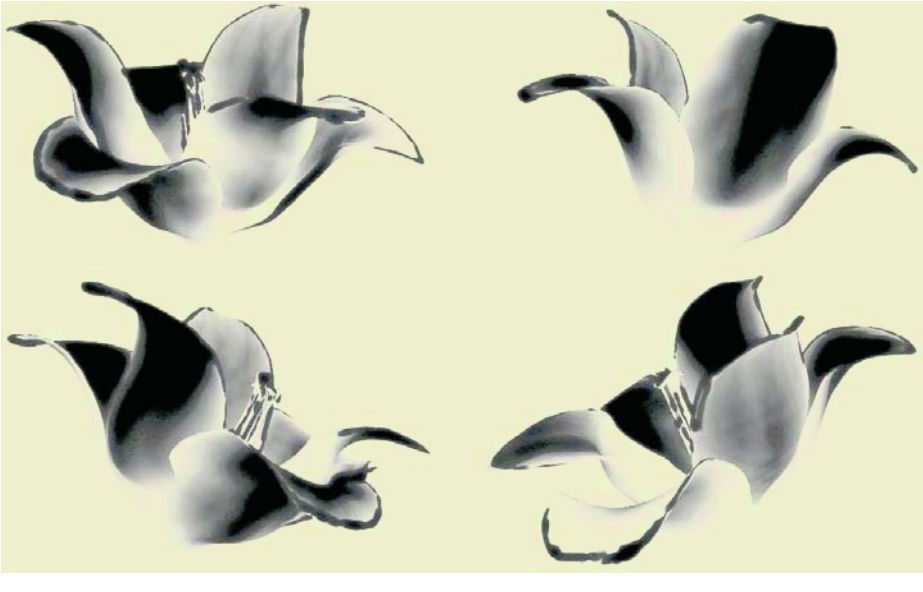


Fig. 9. Different views of a lily model rendered with our system

Table 1. Model complexity and simulation times

Model	# of faces	3D Simulation(s)	2D Simulation(s) (800x600)	Total Simulation(s)
Lily	89744	4.23	5.21	9.44
Ginseng	55568	2.10	5.21	7.31
Korean melons	152768	8.23	5.21	13.44
Bune mask	59279	3.65	5.21	8.86

6 Conclusion

We present a technique for rendering animated 3D models in a Su-Muk painting style with user's guide. A user can draw Su-Muk painting for animation, very easily and intuitively. And output results are as real as elaborate hand-crafted look of Su-Muk paintings.

As future work, we would also like to design a automatic 3D sumuk painting system which varies the brush parameter values automatically, according to the shapes and features of the 3D model.

Acknowledgments. This work was supported by University IT Research Center Project and the Basic Research Program of the Korea Science and Engineering Foundation (No. R01-2002-000-00512-0).

References

1. Strassmann, Steve.: Hairy Brush. In Proceeding of SIGGRAPH, pp. 225-232, 1986.
2. Pham, Blin.: Expressive Brush Strokes. Graphical Model and Image Processing, 1991.
3. Lee, Jintae.: Diffusion rendering of black ink painting using new paper and ink models. Computers & Graphics, pp. 295-308, 2001.
4. Chan, Ching and Akleman, Ergun and Chen, Jianer.: Two Methods for Creating Chinese Painting. Pacific Graphics, 2002.
5. Kang, Shin-Jin and Kim, Sun-Jeong and Kim, Chang-Hun.: Hardware-Accelerated Real-Time Rendering for 3D Sumi-e Painting. LNCS 2669, pp. 599-608, 2003.
6. Zhang, Qing and Sato, Youetsu and Takahashi, Junya and Muraoka, Kazunobu and Chiba, Norishige.: imple Cellular Automaton-based Simulation of Ink Behaviour and Its Application to Suibokuga-like 3D Rendering of Trees The Journal of Visualization and Computer Animation, pp. 27-37, 1999.
7. Baxter, Bill and Scheib, Vincent and Lin, Ming C. and Manocha, Dinesh.: DAB: Interactive Haptic Painting with 3D Virtual Brushes. In Proceeding of SIGGRAPH, 2001.
8. Wong, Helena T. F. and Ip, Horace H. S.: Virtual brush: a model-based synthesis of Chinese calligraphy Computers & Graphics, pp. 99-113, 2000.
9. Xu, Songhua and Lau, Francis C. M. and Tang, Feng and Pan, Yunhe.: Advanced Design for a Realistic Virtual Brush. In Proceeding of EUROGRAPHICS, 2003.

Sports Equipment Based Motion Deformation

Jong-In Choi[2], Chang-Hun Kim[2,⋆], and Cheol-Su Lim[1]

[1] Dept. of Computer Engineering, SeoKyeong University
cslim@skuniv.ac.kr
[2] Dept. of Computer Science and Engineering, Korea University
{gmcg2001, chkim}@korea.ac.kr

Abstract. We propose a novel scheme for generating the locomotion of sports equipment along with motion deformation. To accomplish this, we search the event frames in order to find the points where the character is in contact with the sports equipment (contact points) or not (split points). To find these points, it is necessary to locate the corresponding event points in the signal using Discrete Fourier Transformation. This technique generates a left and right hand-cross analysis and sports equipment locomotion adapted to the character motion. Also, the locomotion of the sports equipment is changed by adjusting its properties, and this may require the posture of the character to be modified. In conclusion, we are able to deform the motion of both the sports equipment and the character simply by adjusting the properties of the sports equipment. Our scheme can be used for the motion deformation of any character with an associated sports equipment.

1 Introduction

It is difficult to deform the motion of physical sports equipments, because the natural harmony between the character and the sports equipment is lost if we capture the motion of the character and sports equipment separately. To solve this problem, we introduce an automatic algorithm, which can deform the motion automatically by adjusting the properties of the sports equipment. The motion of the sports equipment is generally made by hand or captured using a special technique, and this is expensive and time consuming. Since our scheme generates the locomotion of sports equipments automatically, the task becomes easier. Also, we can automatically generate natural deformed motion by adjusting the properties of the sports equipment. Because the physical properties of the sports equipment are retained, we can refer to the proposed method as a safe deformation method. The properties of the sports equipment are its unique physical characteristics, sports equipment such as its elasticity, weight, etc. If we make its elasticity higher, the sports equipment bounces higher or vice versa.

Motion editing is a technique which allows the motion in a moving image to be modified. Bruderlin and Williams [2] showed that techniques from the signal

⋆ The corresponding author: Tel. +82-2-3290-3574; Fax. +82-2-953-0771.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1148–1156, 2005.

processing domain can be applied to the manipulation of animated motions. They introduced the idea of displacement mapping to alter a motion clip. Witkin and Popovic [14] presented a motion warping technique for the same purpose. Bruderlin and Williams also presented a multi-target interpolation technique with dynamic time warping to blend two motions. Unuma et al. [13] used Fourier analysis techniques to interpolate and extrapolate motion data in the frequency domain.

Traditionally, inverse kinematics solvers can be divided into two categories: analytic and numerical solvers. Most industrial manipulators are designed to have analytic solutions for efficient and robust control. Kahan [7] and Paden [12] independently discussed methods of solving an inverse kinematics problem by reducing it into a series of simpler sub-problems whose closed-form solutions are known. A numerical method relies on an iterative process to obtain a solution. Girard and Maciejewski [3] addressed the locomotion of a legged figure using a Jacobian matrix and its pseudo inverse. Koga et al. [9] made use of the results obtained from neurophysiology to achieve an experimentally good initial guess and then employed a numerical procedure for the purpose of fine tuning the solution.

In motion analysis, we can use the zero-crossings of the second derivative of the motion data to detect significant changes in the motion. The zero-crossing point in a trajectory implies changes in motion such as starting from rest, coming to a stop, or changing the direction. These events were noted to have a descriptive significance in [1]. When the zero-crossing point also coincides with an end-effector which is proximal to another type of sports equipment, it implies the contact of the end-effector with the sports equipment. The zero-crossing point enable us the proximities only at possibly relevant frames.

Previous studies were focused on adapting the motion of a person to his or her path of movement or environment. In this study, however, we are interested in the relationship between a person or character and sports equipment. We generate the locomotion of the sports equipment according to the motion capture data. To accomplish this, we search the event frames in order determine if the character is in contact with the sports equipment or not, and classify the frame accordingly, while also registering other attributes, such as whether the left or right hand is in contact with the sports equipment and number of the frame. Then, we deform the character motion by adjusting the properties of the sports equipment. This change in the properties of the sports equipment is designed to make certain physical properties of the sports equipment higher or lower. The trajectory of the sports equipment can also be changed, in which case the motion of the character deforms to fit the modified trajectory of the sports equipment. To accomplish this, we use displacement mapping and inverse kinematics with constraints.

2 Overview

Our scheme is divided into three stages. Figure 1 shows an overview of our scheme. Firstly, we classify the events in terms of the relationship between the

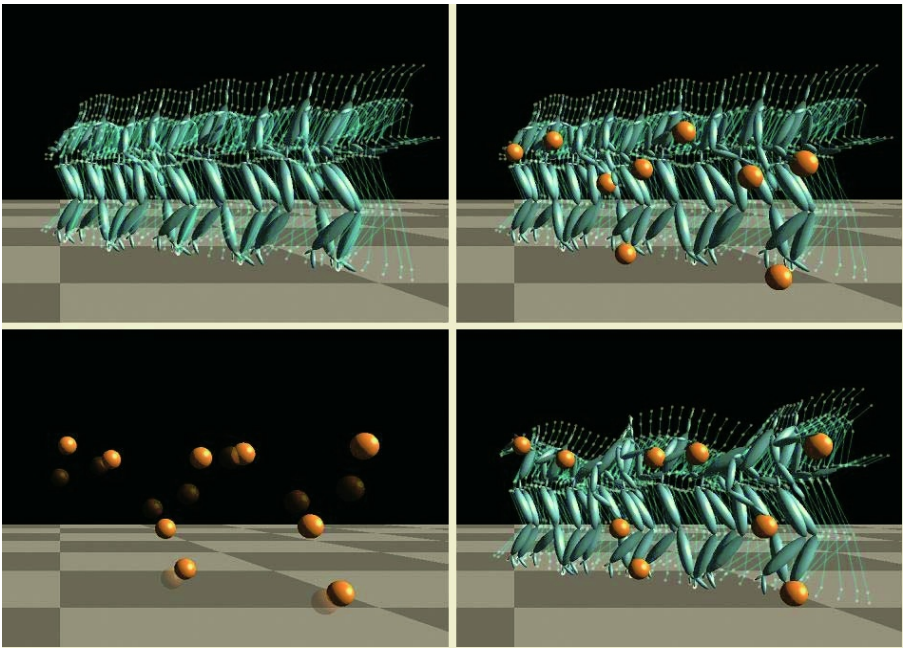


Fig. 1. Motion deformation using the trajectory of the sports equipment

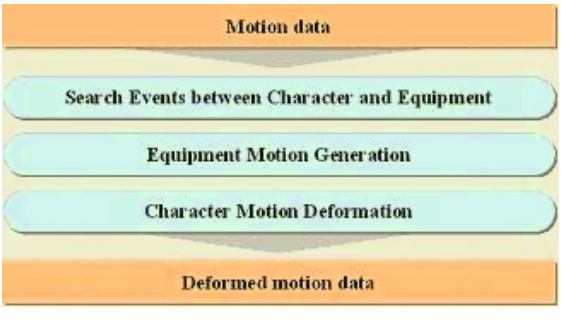


Fig. 2. Flowchart of our scheme

character and the sports equipment. To accomplish this, we search the event frames in order determine if the character is in contact with the sports equipment (contact frames) or not (split frames), and classify the frame accordingly, while also registering other attributes, such as whether the left or right hand is in contact with the sports equipment and number of the frame. Secondly, we generate the locomotion of the sports equipment by applying the appropriate physical formulae. In this way, we compute the positions of the sports equipment ein the split frames and the palm position in the contact frames. Lastly, we deform the character motion according to the trajectory of the sports equipment, by means of displacement mapping and inverse kinematics. The final result is a satisfac-

tory trajectory for the sports equipment and correctly deformed motion data. Figure 2 shows a flowchart for our scheme. In the next section, we explain our algorithm in more detail.

3 Search for Events Between the Character and the Sports Equipment

First of all, we must find those events in which there is separation or contact between the character and the sports equipment at a particular frame, which we refer to as event frames. To accomplish this, we employ signal processing. Firstly, we decide the event ranges and search for event frames, and then we classify the event frames. This process is fully automatic. We input the motion data and the computer performs all of the processes automatically, thereby providing us with all of the event frames in the motion data.

3.1 Determination of Event Ranges

An event range includes only one event frame. It starts at a zero point and ends at a zero point. Namely, both ends of the range are the zero-crossing of the union velocity graph. The union velocity signal is made by adding two signals, viz. the left and right velocity signals. The upper part of figure 3 shows the result of this

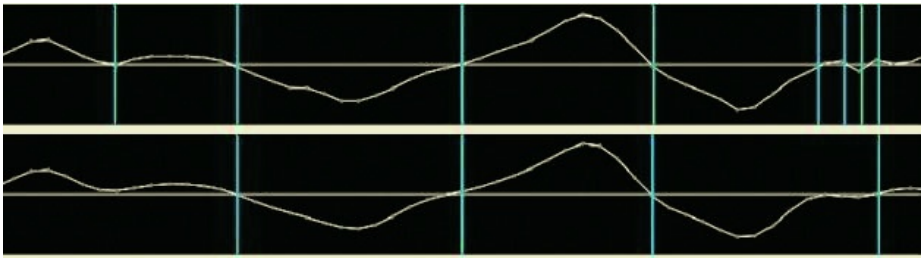


Fig. 3. Noise reduction. Top: Original signal. Bottom: Filtered signal

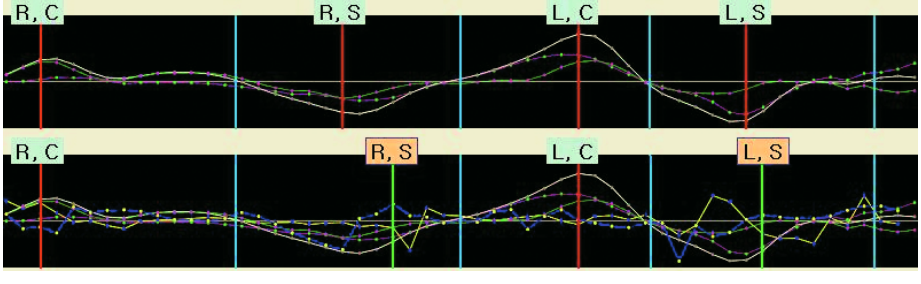


Fig. 4. Zero-crossing of events. L and R refer to the left and right hands, respectively. C and S refer to contact and split, respectively. Top: Original zero-crossing points. Bottom: Fixed split points

operation. It has some short ranges, which are due to noise. Next, we reduce this noise using low pass filtering, which is accomplished by Fourier transformation. After the low pass filtering, the noise disappears, as shown in the lower part of figure 3. The blue lines are the boundaries of the ranges, which are the zero-crossing positions of the signal. The DFT (Discrete Fourier Transformation) and inverse DFT operations are described by equation 1. Figure 3 shows the effect of low pass filtering.

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n] exp\left(\frac{-j2\pi kn}{N}\right) \\ x[n] &= \frac{1}{N}\sum_{k=0}^{N-1} X[k] exp\left(\frac{j2\pi kn}{N}\right) \end{aligned} \tag{1}$$

3.2 Detection of Zero-Crossing Points

We detect the event which has the maximum speed and classify the events frames. The event frames are classified according to three parameters, viz. their number, state and hand, where the number is the frame count, the state is either split or contact and the hand is either left or right. The hand is decided by the difference between the maximum and minimum velocity in the range. We select the bigger of the two differences. Also, the maximum velocity corresponds to contact, while the minimum velocity corresponds to split. For more natural motion, we must fix split frame. The fixed split frame is the first frame which has maximum acceleration after the original split frame.

4 Sports Equipment Locomotion Generation

We construct the trajectory of the sports equipment according to the motion data. The state of the sports equipment is divided into two stages, viz. the split and contact states. In the case of the split state, we can obtain the position of the sports equipment simply by computing thee hand position. The hand position can be computed by determining the position of the center of the palm, where the center of the palm is considered to be located midway between the fingers and the wrist joint. In the case of contact, the position of the sports equipment is computed by means of the appropriate physical formulae. Because plausible equipment locomotion is sufficient for our purposes, we only use the

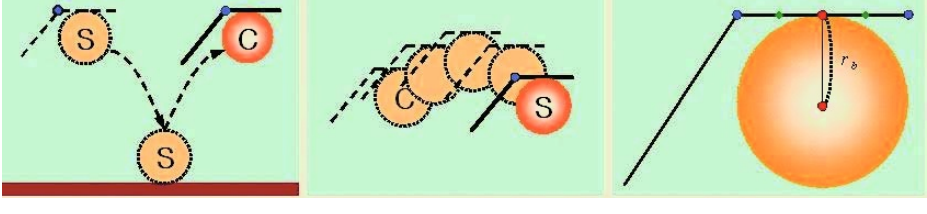


Fig. 5. Basketball locomotion. Left: Split to contact. Middle: Contact to split. Right: Ball position in hand (r_b is radius of ball)

acceleration interpolation method. In order to obtain more realistic movement, another method needs to be used, such as space-time constraints. Figure 5 shows the computed locomotion of a ball. The hand is presented by black stick figures.

5 Character Motion Deformation

To obtain deformed character motion, we adjust the properties of the sports equipment. Since, in our example, we used a ball for the implementation of the proposed method, we adjust the elasticity of the ball, by either decreasing or increasing it by one third, thereby causing the ball to bounce one third lower or higher, respectively. As a result, the event position is also reduced or increased by one third. We cause the motion to fit the changed trajectory of the sports equipment by means of inverse kinematics [15]. To accomplish this, constraints are added to both feet of the character, and we adjust his or her hands in all event frames. To do this, the character pose is changed adaptively. Figure 6 shows the result. Displacement mapping is used to generate a smooth motion

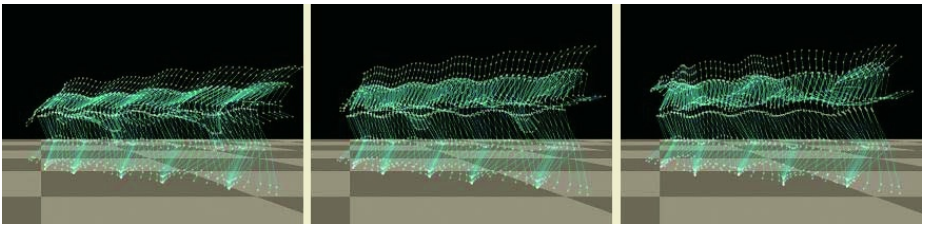


Fig. 6. Dribble motion deformation. Left: Lower posture. Middle: Original posture. Right: Higher posture

stream. This warps the range of the motion between the event frames. It is a very powerful and simple method of motion deformation. The following equation calculates the number of degrees added to the orientation of each joint.

$$\theta_{add} = w(t)\theta_1(t) + (1 - w(t))\,\theta_2(t) \tag{2}$$

where $\theta_1(t)$ is the angle at the beginning of the range and $\theta_2(t)$ is the angle at the end of the range. We add θ_{add} to all joint angles. w(t) is the effective weight of the two degrees, $\theta_1(t)$ and $\theta_2(t)$, which are computed by inverse kinematics.

6 Results

We implemented our scheme in Pentium IV 1.7 GHz, 512 RAM. It was performed in real-time. We made elasticity highly or lowly and mixed two trajectories, which mean split position is lower and contact position is higher. Following figures show our results. The 'k' means elasticity. We can see characters which bows waist

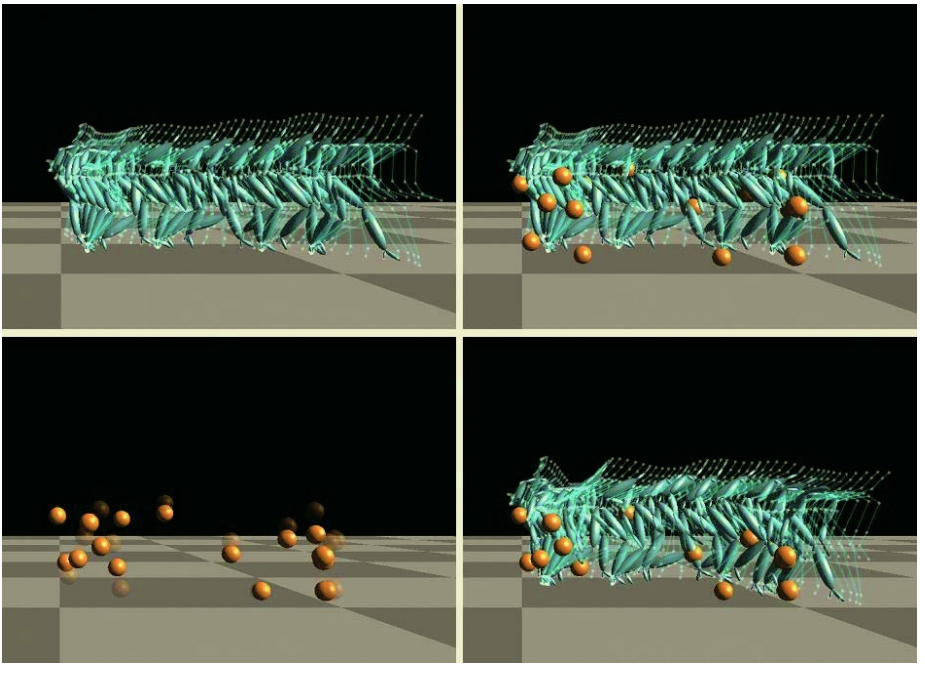


Fig. 7. Arbitrary dribble motion

down. Ball trajectory is lower. Transparent balls indicate original trajectory and clear balls changed trajectory. Figure 7 shows another dribble motion which changes hand. Figure 8 shows our results at k = 0.66, 1.33 and mixed. Lower elasticity causes low ball trajectory. In contrast high elasticity causes high ball trajectory. And split is adapted to lower position and contact is higher position in bottom of figure 8.

We implemented our scheme using a PC with a Pentium IV 1.7 GHz and 512 MB of RAM. The experiment was performed in real-time. We made the elasticity of the ball higher or lower and used two trajectories, in which the split position is lower and the contact position is higher. Figures 7 and 8 show the results, where 'k' refers to the elasticity. We can see the characters who is bent over at the waist. The ball trajectory is lower. The transparent balls indicate the original trajectory and the opaque balls indicate the changed trajectory. Figure 7 shows another example, involving dribble motion in which the ball passes from one hand to the other. Figure 8 shows the results for k = 0.66, 1.33 and mixed. Lower elasticity causes the ball trajectory to be lower. In contrast, higher elasticity causes the ball trajectory to be higher. The split is adapted to lower elasticity and contact is higher elasticity in bottom of figure 8.

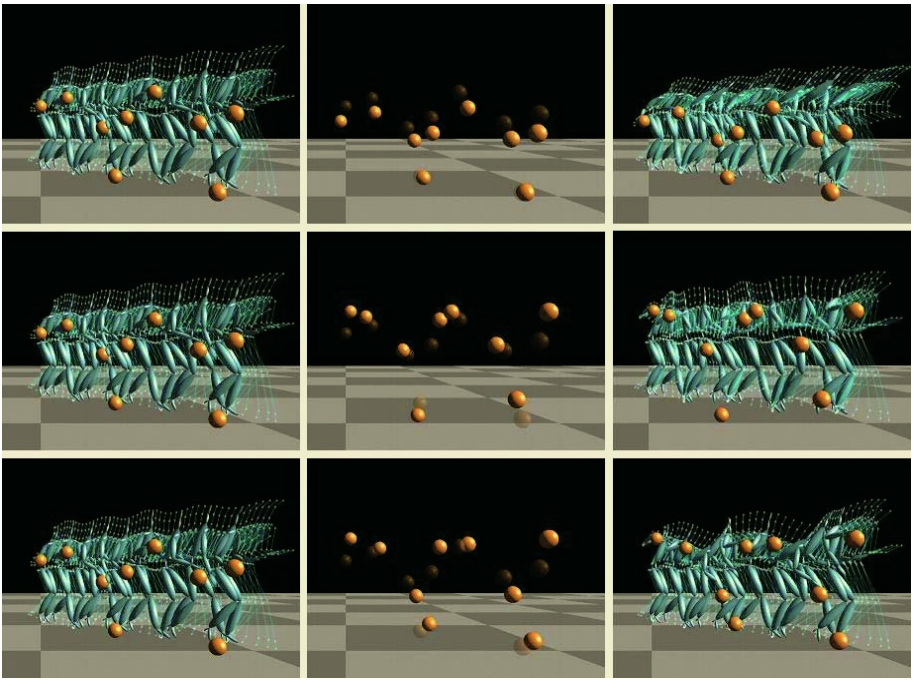


Fig. 8. Top: Lower ball elasticity (k = 0.66). Middle: Higher ball elasticity (k = 1.33). Bottom: Mixing two levels of elasticity

7 Conclusion and Future Work

In this study, we explained the generation of locomotion for sports equipment and motion deformation by adjusting the properties of they sports equipment. We generated the locomotion of sports equipment automatically using motion capture data, using signal processing. We were able to find the zero-crossing points of the signal by removing the short ranges. After generating the locomotion of the sports equipment, we changed its properties. Changing the properties of the sports equipment caused its trajectory to be changed. We used this technique to bring about character motion deformation. Our scheme is very simple but effective. In the future, we will adapt our scheme for a variety of different motions, and. generalize it so that it can be used with other kinds of sports equipments.

Acknowledgements

This work was supported by University IT Research Center Project and the Basic Research Program of the Korea Science and Engineering Foundation (No. R01-2002-000-00512-0).

References

1. R. Bindinganavale and N.I. Badler. Motion Abstraction and Mapping with Spatial Constraints. *International Workshop*, CAPTECH, 251–276, 1998
2. A. Bruderlin and L. Williams. Motion signal processing. *SIGGRAPH*, 97–104, 1995
3. M. Girard and A. A. Maciejewski. Computational modeling for the computer animation of legged figures. *SIGGRAPH*, 263-270, 1985
4. M. Gleicher. Retargetting Motion to New Characters. *SIGGRAPH*, 33–42, 1998
5. M. Gleicher. Motion Path Editing. *ACM Symposium on Interactive 3D Graphics*, 195–202, 2001
6. M. Gleicher, H.J. Shin, L. Kovar and A. Jepsen. Snap-Together Motion Assembling Run-Time Animations. *ACM Symposium on Interactive 3D Graphics*, 181–188, 2003
7. GW. Kahan. Lectures on computational aspects of geometry. *Unpublished manuscripts*, 1983
8. T.H Kim, S.I. Park and S.Y. Shin. *Rhythmic-Motion Synthesis Based on Motion-Beat Analysis. SIGGRAPH*, 392–401, 2003
9. Y. Koga, K. Kondo, J. Kuffer, and J. Latombe. Planning motions with intentions. *SIGGRAPH*, 395–408, 1994
10. L. Kovar, M. Gleicher and F. Pighin. Motion Graphs. *SIGGRAPH*, 473–482, 2002
11. J.H. Lee and S.Y. Shin. A Hierarchical Approach to Interactive Motion Deformation for Human like Figures. *SIGGRAPH*, 39–48, 1999
12. B. Paden. Kinematics and Control Robot Manipulators. In *PhD thesis*, 1986
13. M. Unuma, K. Anjyo, and R. Takeuchi. Fourier principles for emotion-based human figure animation. *SIGGRAPH*, 91–96, 1995
14. A. Witkin and Z. Popovic. Motion warping. *SIGGRAPH*, 105–108, 1995
15. J. Zhao and N. I. Badler. Inverse Kinematics Positioning Using Nonlinear Programming for Highly Articulated Figures. *ACM Transactions on Graphics*, 313–336, 1994

Designing an Action Selection Engine for Behavioral Animation of Intelligent Virtual Agents

F. Luengo[1,2] and A. Iglesias[2,3,⋆]

[1] Department of Computer Science, University of Zulia,
Post Office Box #527, Maracaibo, Venezuela
fluengo@cantv.net
[2] Department of Applied Mathematics and Computational Sciences,
University of Cantabria, Avda. de los Castros, s/n, E-39005, Santander, Spain
[3] Laboratory of Advanced Research, Building B, Room #1025,
University of Tsukuba, Kaede dori, Third Cluster of Colleges, Tsukuba, Japan
iglesias@unican.es
http://personales.unican.es/iglesias

Abstract. This paper presents a new action selection scheme for behavioral animation in computer graphics. This scheme provides a powerful mechanism for the determination of the sequence of actions to be performed by the virtual agents emulating real world's life. In particular, the present contribution focuses on the description of the system architecture and some implementation issues. Then, the performance of our approach is analyzed by means of a simple yet illustrative example. Finally, some advantages of our scheme and comparison wih previous approaches are also briefly discussed.

1 Introduction

The issue of *action selection* has been largely analyzed in the framework of ethology and cognitive sciences [2, 22, 23], psychology [14] and robotics [1, 15]. More recently, it has also become an interesting challenge for behavioral simulation in computer graphics [4, 5, 6, 20, 21]. Roughly speaking, it can be established as follows: *at each moment of time, given a set of feasible goals to be performed, we want to choose the most appropriate one based on the agent's internal and external conditions*. In other words, the central problem to deal with is the determination of the sequence of actions to be performed by the virtual agents as a function of internal and/or external factors. Of course, this determination is expected to be *realistic*, since we are going to use virtual agents to simulate human beings with a certain level of realism.

From the previous definition, it becomes clear that the construction of appropriate schemes for action selection is a key component in behavioral animation

⋆ Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1157–1166, 2005.

of virtual characters. Because of that, a number of different proposals have been described in the literature (see, for instance, [1, 2, 5, 6, 7, 12, 13, 16, 17, 18, 20, 21] and references therein).

In this paper, a new framework for action selection is presented. We point out here that this action selection system is actually a module of a whole behavioral animation system already described in previous references [8, 11]. The reader is also referred to [9] and [10] for more details about such a behavioral system.

The structure of this paper is as follows: Section 2 describes the architecture of this new approach and its simulation flow as well as some implementation issues. The performance of this new scheme is analyzed in Section 3 by means of a simple yet illustrative example. Finally, some advantages of our scheme and comparison wih previous approaches are briefly discussed in Section 4. Conclusions and further remarks close the paper.

2 The Action Selection System

This section describes the action selection system introduced in this paper. Firstly, we focus on the description of the system architecture. Then, some implementation details are also given. Finally, the simulation flow is briefly analyzed.

2.1 System Architecture

The architecture of the action selection system described in this paper is displayed in Figure 1. It consists of a *goal database* and three different modules (the emotional analyzer, the intention planning and the action planning) intended to perform specific tasks as described below.

The first component of our system is a database that stores a list of arrays (associated with each of the available goals at each time) having the structure:

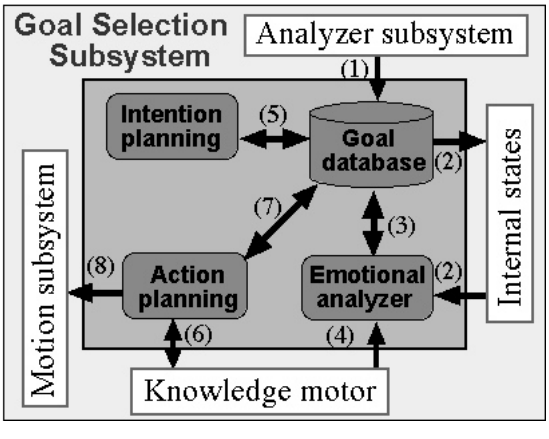


Fig. 1. Architecture of the goal selection system

Table 1. Variables associated with each goal stored into the database: variable names (left) and their meaning (right)

Variable		Meaning
`goalid`	-	goal identification code (see Section 3 for a list of feasible goals for the example described in this paper)
`feasr`	-	goal's feasibility rate
`priority`	-	goal's priority (determined by the intention planning)
`wishr`	-	wish rate (determined by the emotional analyzer)
`t`	-	time at which the goal is selected
`successr`	-	goal's success rate

```
[goalid, feasr,priority, wishr, t, successr]
```

where the meaning of each variable is indicated in Table 1. On the other hand, an additional array is stored only for the current goal in progress: `[tsel,tatt]`, the components being the time at which the goal has been selected and attained, respectively.

The *emotional analyzer* is the module responsible to update the wish rate of a goal (regardless its feasibility). Such a rate takes values on the interval $[0, 100]$ according to the mathematical functions describing the agent's internal states[4]. Those functions involve the internal state variables described in Section 3 as well as two parameters:

(1) the *dynamic rate*, D, which expresses the agent's predilection for dynamical activities (such as walk or run) over the intellectual ones and
(2) a temporal parameter Ω_k, defined as:

$$\Omega_k = \alpha_k \, \delta_k \, \frac{t - t_m}{t - t_k} \tag{1}$$

where Θ is the set of all possible goals, $t_m = \min_{j \in \Theta, j \neq k} t_j$, where t_j is the simulation step at which the j-th goal was selected for the last time and k the current goal, t is the current time, $\delta_k \in \{1, 1.2\}$ is a parameter that accounts for the goal's success (successful goals exhibit higher wish rate than those unsuccessful), and $\alpha_k \in [0, 2]$ is a parameter used to promote some particular goals with respect to others, depending on the agent's personality. Note that the role of Ω_k is to increase the wish rate of the oldest goals in the priority list (i.e. the older a goal, the higher its wish). This simple procedure assures that, for a sufficiently long span, all possible goals will be finally selected. Note also that this condition can be easily skipped by simply omitting this factor in the equations of the internal states.

[4] The mathematical description of the internal state functions corresponding to the example described in this paper is not included here because of limitations of space. For the definition of those functions for a different example, the interested reader is referred to [10].

The *intention planning* module determines the priority of each goal. To this aim, it uses information such as the factibility and wish rate. From this point of view, it is rather similar to the "intention generator" of [20] except by the fact that decision for that system is exclusively based on rules. Our intention planning module also comprises a buffer to store temporarily those goals interrupted for a while, so that the agent exhibits a certain "persistence of goals". This feature is specially valuable to prevent agents from the oscillatory behavior appearing when the current goal changes continuously.

The last module is the *action planning*, a based-on-rules expert system that gets information from the environment (via the knowledge motor described in [8]), determines the sequence of actions to be carried out in order to achieve a particular goal and updates the goal's status accordingly. These actions are transferred to the motion subsystem to be converted into graphical instructions subsequently sent to the graphics pipeline.

2.2 Implementation Issues

Concerning the implementation details, the action selection module presented here has been developed in Visual C++ v6.0 on a PC platform with Pentium IV processor and 256 MB. of RAM. The graphical output has been implemented on Open GL with GLUT and subsequently compiled in Visual C++.

It is interesting to remark that our decomposition of the goal selection module into four subsystems as described in Section 2.1 is very useful from the programmers' viewpoint: on one hand, maintenance, debugging and updating of the system components are much easier and simpler. On the other hand, any function can be modified by simply rewritting some code lines of the particular subsystem at which this function is implemented.

2.3 Simulation Flow

Figure 1 depictes the simulation flow of the goal selection system described above. Firstly, the analyzer subsystem updates the factibility, which is stored into the goal database (step (1) of that figure). Then, the emotional analyzer gets information about:

- the internal states from the internal states subsystem (2),
- the time at which each goal is selected/attained from the goal database (3) and
- relevant parameters from the knowledge motor (4).

This information is used by the emotional analyzer to update the goals' wish rate at the goal database. The factibility and wish rates are sent to the intention planning module (5) to determine the priority of each goal, which is subsequently updated at the goal database. Then, the current goal is sent to the action planning module. It takes additional information on the environment from the knowledge motor (6) in order to run the set of actions associated with such a goal. This will modify the agent's status within the virtual 3D world (and, hence, the knowledge motor as well). Information about the actions is sent to

the goal database (7) to update the goal's status (failed, candidate, in progress). Finally, those actions are sent to the motion subsystem (8) to be converted into graphical instructions.

3 An Illustrative Example

In this section, the performance of the goal selection scheme is analyzed by means of a very simple yet illustrative example. We remark that this example is considered here for illustrative purposes only. In fact, more complex scenarios can be easily generated from our system.

The scene consists of a shopping center at which the virtual agents can perform a number of different actions, such as eat, drink, play videogames, sit down to rest and, of course, do shopping. The environment also comprises different static (such as trees, tables, shops) and smart objects (such as benches, videogame machines, drink machines). Therefore, it is a convenient place for a wide range of potential agent-object and agent-agent interactions. To this aim, we consider four virtual agents, three kids and a woman.

Figure 2 shows the temporal evolution of the internal states (top) and the goals' wishes (bottom) for one of the kids. Similar graphics can be obtained for the other agents (they are not included here because of limitations of space). The picture on the top displays the temporal evolution of the five internal state functions (valued onto the interval $[0, 100]$) considered in this paper, namely, `energy`, `shyness`, `anxiety`, `hunger` and `thirsty`. On the bottom, the wish rate (also valued onto the interval $[0, 100]$) of the feasible goals (`have a rest`, `eat something`, `drink water`, `take a walk` and `play videogame`) is depicted.

In the example described in this paper, the following initial values for the agent's internal states and parameters have been chosen: `energy=100`, `shyness=0`, `anxiety=0`, `hunger=0` and `thirsty=0`. Therefore, the kid is very sociable and dynamic and likes activity very much, while being neither hunger, nor anxious nor thirsty at all. Both pictures in Figure 2 are labelled with eight numbers indicating the different simulation's milestones (the associated animation screenshots for those time units are displayed in Figure 3):

(1) At the initial step, the three kids go to play with the videogame machines, while the woman moves towards the eating area (indicate by the tables in the scene). Note that the internal state with the highest value for the agent analyzed in this work is the energy, so the agent is going to perform some kind of dynamic activity, such as to play.
(2) The kid keeps playing (and their energy level going down) until his/her satisfaction reaches the maximum value. At that time, the anxiety increases, and the agent'wish turns into performing a different activity. However, the goal `play videogame` is still that with the highest wish rate, so this goal will be in progress for a while.
(3) At this simulation step, the anxiety reaches a local maximum again, meaning that the kid is getting bored about playing videogames. Simultaneously, the

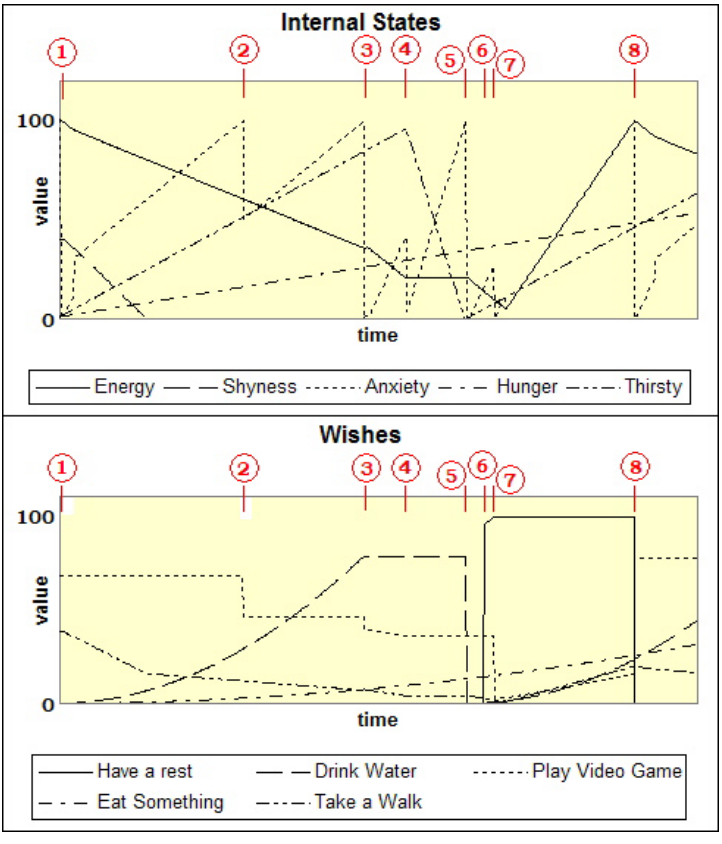


Fig. 2. Temporal evolution of the internal states (top) and available goals' wishes (bottom) for the example in this paper

goal with the highest value is `drink water`, so the agent stops playing and starts to look for a drink machine.

(4) At this time, the kid gets the drink machine and starts to drink. Consequently, the internal state function `thirsty` decreases as the agent drinks until the status of this goal becomes *goal attained*.

(5) Once this goal is satisfied, the goal `play videogames` is the new current goal. So, the kid comes back towards the videogame machines.

(6) However, the energy level is very low, so the goal `play videogames` is interrupted, and the kid looks for a bench to sit down and have a rest.

(7) Once seated, the energy level turns up and the goal `have a rest` does not apply anymore.

(8) Since the previous goal `play videogames` is still in progress, the agent comes back to it, and plays again.

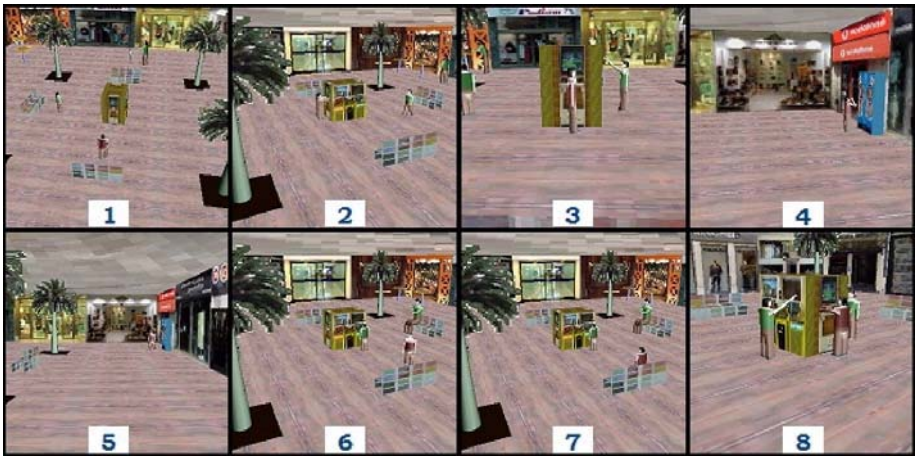


Fig. 3. Screenshots of the shopping center example

The results show the excellent performance of the present scheme. In particular, the six criteria used in [18] to validate a motivational model for action selection are also fulfilled here:

- the motivations are taken into account (criterion 1) via the goal's wish function. Additionally, "the persistence of motivations" is also included into the system. Note, for instance, that at step (5) of our example the goal `play videogames` exhibits the highest value (and hence, it is the current goal). However, it is interrupted by the goal `have a rest` as the agent is too much tired to keep playing at that time. After a while, the agent's energy level is high enough to return playing, and the former goal is subsequently recovered at step (8).
- In addition, the environment information (criterion 2) is provided by the knowledge motor and used to determine the goal's feasibility and perform the actions in sequence for the current goal accordingly (criterion 4).
- On the other hand, the opportunistic behavior can interrupt the current goal (criterion 5). The compromise behavior (to choose the action which satisfies the greatest number of motivations) is also considered here via the intention planning subsystem (criterion 6).
- Finally, criterion 3 (to prefer motivated actions over locomotion actions) is also considered here. In fact, the most important goal for the kids in our example is to `play videogames`, as it is the most dynamic activity available in the shopping center environment. However, we do not expect this criterion to be true everytime-everyone. Back to front, our system allows a richer variety of behaviors ranging from dynamic to softer goals. Such a choice is carried out by using the dynamic parameter of Eq. (1), as described above.

4 Comparison with Previous Approaches

Some interesting features of the present action selection system exhibit certain similarities with others from previous schemes. In particular, the architecture of our goal selection system represents a substantial improvement of that in [6], in which the characters' behavior mechanism is based on compact table-based descriptions and flexible scripts. However, the proposal in [6] is much simpler since it is restricted to a particular case and it is environment and input-device dependent, while ours is extremely flexible: agents can adapt to any environment without modifying the undelying structure. In fact, the process only requires the simple addition and/or modification of the internal states and parameters.

On the other hand, the short-term memory and computer redundancy avoidance via cascading and reusing of [3] are actually applied in our approach, although not exactly in the same way (see [10] for more details). Another advantages are the inclusion of personality (described in terms of different parameter values and functions) and uncertainty (performed through some probability terms, so that different agent parameters lead to a drastically different reactions). Another interesting feature of our system is the use of different Artificial Intelligence techniques for autonomous reasoning. They will be reported in detail in a future publication somewehere else.

5 Conclusions and Future Work

In this paper, a new action selection scheme for behavioral animation of virtual agents is introduced. The paper describes its design and implementation issues as well as its simulation flow. The performance of this approach has been shown by means of an illustrative example. Finally, comparison with previous (similar) approaches is briefly discussed.

Despite of the encouranging results, this is just one step to reproduce realistically the huge range of complex human behaviors and there is a long way ahead. In particular, further research is still needed in order to describe many human behaviors in mathematical terms: some functions are to be improved, others have to be defined yet. On the other hand, we are interested to describe human emotions and how they do influence the decision process [19]. Subsequent versions of this model will include many additional modifications and improvements. However, we do not expect to modify the current design and implementation of our action selection system significantly.

Acknowledgements

This paper has been written while the second author was at the Laboratory of Advanced Research of the University of Tsukuba (Japan) for a sabbatical year stay. He would like to thank the laboratory staff, and especially Prof. Tetsuo Ida, for their wonderful hospitality and great collaboration. *Domo arigato gozaimazu!*

References

1. T. Anderson and M. Donath, Animal behavior as a paradigm for developing robot autonomy. Robotics and Autonomous Systems, **6** (1990) 145-168
2. B. Blumberg, Action-Selection in Hamsterdam: Lessons from Ethology, Proceedings of the 3rd International Conference on the Simulation of Adaptive Behavior, MIT Press, Cambridge, MA (1994)
3. C. Bordeux, R. Boulic and D. Thalmann, An Efficient and Flexible Perception Pipeline for Autonomous Agents, Computer Graphics Forum, **18**(3) (1999) 23-30
4. L. Chen, K. Bechkoum and G. Clapworthy, A Logical Approach to High-Level Agent Control, Proceedings of the Fifth International Conference on Autonomous Agents, ACM Press, NY (2001) 1-8
5. C. Geiger and M. Latzel, Prototyping of Complex Plan Based Behavior for 3D Actors, Proceedings of the Fourth International Conference on Autonomous Agents, ACM Press, NY (2000) 451-458
6. S.H. Guan, S.Y. Cho, Y.T. Shen, R.H. Liang, B.Y. Chen and M. Ouhyoung, Conceptual Farm, Proc. of IEEE Multimedia and Expo, ICME'2004, (2004) TP9-2 (CD-ROM)
7. Iglesias A., Luengo, F.: Behavioral Animation of Virtual Agents (invited paper). Proc. of the Fourth International Conference on Computer Graphics and Artificial Intelligence - 3IA, Limoges, France (2003) 99-114
8. Iglesias A., Luengo, F.: A New Based-on-Artificial-Intelligence Framework for Behavioral Animation of Virtual Actors. Proceedings of Computer Graphics, Imaging and Visualization - CGIV'2004 IEEE Computer Society Press, Los Alamitos, CA (2004) 245-250
9. Iglesias A., Luengo, F.: Intelligent Agents in Virtual Worlds. Proceedings of Cyberworlds - CW'2004, IEEE Computer Society Press, Los Alamitos, CA (2004) 62-69
10. Iglesias A., Luengo, F.: New Goal Selection Scheme for Behavioral Animation of Intelligent Virtual Agents. IEICE Transactions on Information and Systems (2005) (*in press*)
11. Luengo, F., Iglesias A.: Framework for Simulating the Human Behavior for Intelligent Virtual Agents. Lectures Notes in Computer Science, **3039** (2004) Part I: Framework Architecture. 229-236; Part II: Behavioral System 237-244
12. J. Liu and H. Qin, Behavioral Self-Organization in Lifelike Agents, Proceedings of the Second International Conference on Autonomous Agents, ACM Press, NY (1998) 254-260 (See also: J. Liu and H. Qin, Behavioral Self-Organization in Lifelike Synthetic Agents, Autonomous Agents and Multi-Agent Systems, **5**(4) (2002) 397-428).
13. M.L. Maher and N. Gu, Situated Design of Virtual Worlds Using Rational Agents, Proceedings of the Second International Conference on Entertainment Computing, ACM Press, NY (2003) 1-9
14. D. McFarland, Animal Behaviour: Psychobiology, Ethology, and Evolution (2nd edition), Longman Scientific and Technical, Harlow, England (1993)
15. D. McFarland, Intelligent Behavior in Animals and Robots. MIT Press, Cambridge, MA (1993)
16. J.S. Monzani, A. Caicedo and D. Thalmann, Integrating behavioral animation techniques. In Proceedings of EUROGRAPHICS'2001, Computer Graphics Forum, **20**(3) (2001) 309-318

17. S. Sanchez, O. Balet, H. Luga and Y. Dutheu, Autonomous Virtual Actors, Proceedings of the Second International Conference on Technologies for Interactive Digital Storytelling and Entertainment - TIDSE'2004, Springer-Verlag, Berlin Heidelberg, Lectures Notes in Computer Science, Vol. 3015 (2004) 68-78
18. Sevin, E., Thalmann, D.: The Complexity of Testing a Motivational Model of Action Selection for Virtual Humans, Proceedings of Computer Graphics International, IEEE Computer Society Press, Los Alamitos, CA (2004) 540-543
19. D.R. Traum, S. Marsella and J. Gratch, Emotion and Dialogue in the MRE Virtual Humans, Tutorial and Research Workshop, Proceedings of Affective Dialogue Systems - ADS'2004, Kloster Irsee, Germany, 2004.
20. X. Tu and D. Terzopoulos, Artificial fishes: Physics, Locomotion, Perception, Behavior. Proceedings of ACM SIGGRAPH'94 (1994) 43-50
21. X. Tu, Artificial Animals for Computer Animation: Biomechanics, Locomotion, Perception, and Behavior, Ph.D. thesis, Dept. of Computer Science, University of Toronto (1996)
22. T. Tyrrell, Defining the Action Selection Problem, Fourteenth Anual Conference of the Cognitive Society (1992)
23. T. Tyrrell, Computational Mechanisms for Action Selection, Center for Cognitive Science, University of Edimburg (1993)

Interactive Transmission of Highly Detailed Surfaces*

Junfeng Ji[1], Sheng Li[1,3], Enhua Wu[1,2], and Xuehui Liu[1]

[1] Lab of Computer Science, Institute of Software,
Chinese Academy of Sciences, Beijing, China
{jjf, ls, weh, lxh}@ios.ac.cn
[2] Department of Computer and Information Science,
University of Macau, Macao, China
[3] School of Electronics Engineering and Computer Science,
Peking University, Beijing, China

Abstract. It is a challenging work for transmitting the highly detailed surfaces interactively to meet real time requirement of the large-scale model visualization from clients. In this paper, we propose a novel approach to interactive transmission of highly detailed surfaces according to the viewpoint. We firstly map the 3D surfaces onto the parameter space through surface parametrization, and the *geometry images* (GIM) and normal map atlas are obtained by regular re-sampling. Then the quadtree-based hierarchical representation can be constructed based on GIM. Since the hierarchical structure is regular, an efficient compression scheme can be applied to encode the structure and vertices of its nodes. The encoded nodes can be transmitted in arbitrary order, so the extreme flexibility in transmission could be achieved. By taking advantage of normal texture atlas, the rendering result of the partial transmitted model is improved greatly, and then only the geometry on the silhouette to the current viewpoint need be refined and transmitted, and so the amount of data needed to transfer each frame is greatly reduced.

1 Introduction

Due to the availability of the 3D scanner for obtaining the highly detailed object, the size of model currently obtained is increasing rapidly. It poses a challenge to representing, transmitting and rendering such large-scale data sets.

Steady growth of e-commerce and entertainment over Internet has brought increasing interest in transmission of 3D graphics models. In the Internet environment, it is impractical in an interactive speed to download a large full model with complex details because of the limitation of the network bandwidth. Currently, the dominant strategy to deal with the problem is to simplify the model

* The work in this paper is supported by National Key Fundamental Research and Development Project (973) with Grant No. 2002CB312102, 2004CB719403, NSFC (60033010) and Research Grant of University of Macau.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1167–1176, 2005.

and transmit it progressively. The progressive approach allows a very coarse approximation of a model to be sent at first, followed by subsequent transmission of incremental geometry to add more and more details. This process allows the user to obtain an early visual perception of the geometry, and incrementally with detail up to user's satisfaction.

Streaming a triangle mesh $(K,\ V)$ must process two kinds of information: the topological connectivity K (adjacency graph of triangles) and the geometry V (positions of the vertices). To the triangle meshes with arbitrary topology, progressive meshes [4][5] are served as the framework of streaming [11][12][7]. Similar in spirit to progressive meshes, some commercial systems have been developed, such as Meta MTS products [1]. However, the complex topology is not amiable to the interactive transmission. On the contrary, the models with regular topological connectivity [2][3][9] lend themselves in an advantageous position over the arbitrary topology models in the expression and transmission [8].

The transmission of the geometry V is a heavy work, because each vertex requires a vector represented by 3 float-point data. The number of vertices can be reduced through mesh simplification, and the removed details on the surfaces can be visualized by using texture mapping, which can be accelerated by even the low-end graphics hardware. Normal mapping can add 3D shading detail to the surface, and can be accelerated by modern GPU. The details, such as color and normal vector on the surface can be stored compactly (only 1 to 3 bytes for each texel required), and can be converted into textures using surface parametrization [16][14].

Normal mapping only modifies the surface's normal, and so the visual effect at the silhouette of the texture-mapped object is very coarse due to loss of the geometry information. The silhouette can be enhanced using the view-dependent *level of detail* (LOD) techniques.

In this paper, we propose an extremely flexible scheme to transmit the details and geometry progressively and interactively. We employ the surface parametrization to convert the arbitrary surfaces into geometry images (GIM)[2], and the details on the surface can be mapped onto the texture atlas. Since GIM is the completely regular representation of 3D surface in the parameter space, we construct a quadtree structure based on GIM, dubbed P-Quadtree. Since the details texture atlas is utilized to enhance realism of the partial transmitted models, the key to our approach is to refine and stream the silhouette geometry view-dependently. To transmit the P-Quadtree efficiently and flexibly, we encode and encapsulate it into a localized communication packet. The topological connectivity of the node is encoded as an integer number by its localization code in pointless quadtree structure [13], and the coordinates of each vertex can be encoded as a scalar float number by using the normal meshes [3]. Since all hierarchical and topological structure of the node is packed into a communication packet, the nodes can be transmitted in arbitrary order.

The contribution of this paper comes from two aspects: Firstly, a very efficient encoding scheme is applied to pack the structure and the vertices of the nodes efficiently. Secondly, a transmitting scheme with high flexibility is designed so

that the nodes can be transmitted in arbitrary order. Since the realism of the rendered image would be improved greatly by using normal texture atlas, only the silhouette need to be refined view-dependently. This guarantees the minimal amount of data to be transmitted.

2 Overview

Our approach contains two stages: data preprocessing stage and data transmission stage. The transmission scheme of our approach includes two parts: data preparing on server and data receiving and 3D surface construction on the clients. The overall process is shown in Fig. 1.

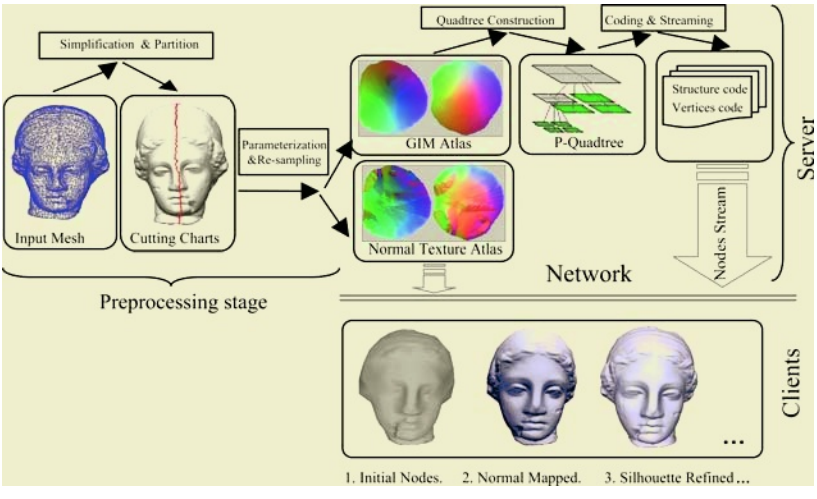


Fig. 1. The framework of the interactive transmission

On the server side, the meshes with arbitrary topological connections must be processed, and the geometry and details on the surface should be converted into the regular representation in the parameter space: GIM atlas and normal texture atlas. The P-Quadtree structure is constructed based on the GIM atlas.

When the transmission process begins, P-Quadtree can be streamed according to the requests with different viewpoints from the clients. The data of detailed surfaces are compressed and sent to the clients through the network.

Once the data are received on the client, the nodes' structure and vertices will be decoded from the data packages, and the structure of P-Quadtree and the surface could be reconstructed. This process is progressive and view-dependent. The normal mapping can be accelerated by modern graphics hardware, and thus the client re-quires only the refinement of the silhouette to reduce the data to be transmitted.

3 P-Quadtree and the Node Encoding

P-Quadtree is a multi-resolution representation of geometry image, whose nodes can be encoded efficiently by using the regular structure.

3.1 Geometry Images

Geometry image [2] is an "image-like" representation of surfaces. It is constructed by converting the surfaces into a topological disk using a network of cuts and parameterizing the resulting disk onto a square domain using surface parameterizations, which is a mapping U: $(x, y, z) \rightarrow (u, v)$ from 3D space to 2D space. It samples the surface geometry as a $n \times n$ array of (x, y, z) values, where n is the size of the image. A pixel in the geometry image maps the 3D coordinates of a vertex on the surface in object space, and four neighbored pixels correspond to the vertices of a quadrilateral in the object space.

Mapping the whole surface of the model onto a square domain will introduce high distortion, especially for the shapes with high genus or long extremities. To mitigate this, the surface can be cut into several pieces, i.e. charts, and many charts can be packed into multi-chart geometry images [15] to represent the complex model. Each chart has a natural boundary, and is parameterized onto irregular polygon. Such a parametrization atlas reduces the distortion, and distributes samples more uniformly over the surfaces and therefore better captures the surface signals. The boundary must be zippered carefully [15].

3.2 The Structure Encoding of P-Quadtree

P-Quadtree construction. P-Quadtree is constructed based on GIM. It requires the size of square to be $(2^n+1) \times (2^n+1)$ to construct a quadtree, where n is the number of levels. To meet this requirement, the packed atlas is sampled by (2^n+1) pixels on the longer side, and then we enlarge its shorter side to be (2^n+1). The new pixels in the expanded area are assigned with null values. We construct P-Quadtree top-down. The process includes construction of quads hierarchy and computation of the quad attributes, such as node's errors, radius of bounding sphere (Fig. 2(b)) and normal cone (Fig. 2(c)). These attributes are used in view-dependent data streaming.

Node Structure. A node in our algorithm is a 33 pixels block in image space, whose structure is shown in Fig. 2(a), and represents a quadrilateral patch in object space (as shown in Fig. 2(b)). A node has a center vertex and four vertices at the corners, called *basic vertices* (the gray dots in Fig. 2(a, b)), which are indispensable to construct the surface. When we render the surfaces using P-Quadtree nodes, crack may appear on the edge whose adjacent patches have the different level. In order to prevent the crack, new vertices need to be added on the middle of these edges. We call these vertices as *optional vertices* (the white vertices in Fig. 2(a, b)). Only the basic vertices need to be transmitted. A quadrilateral patch of the node can be subdivided recursively. From the view of topology, we could use a 2-tuple (K, V) to represent it, where K is a simplicial complex, expressing the connectivity of vertices, edges and surfaces; $V = \{V_i$

$\in R^3 \mid i \geq 5\}$, which gives the vertices of node or the quadrilateral patch. The attributes of a vertex may include the geometry coordinates (x, y, z), the normal vector $\mathbf{n}$, and the radius of the bounding sphere r.

Since the shape of the chart is irregular polygon, it may contain null pixels. If all of its pixels are defined pixels, the node is called a *valid node*. On the contrary, if the valid pixels in a node could not map at least one valid triangle in object space, it is called *invalid node*. The others nodes are the *gray nodes*. The higher level gray nodes may be subdivided into valid nodes and invalid nodes.

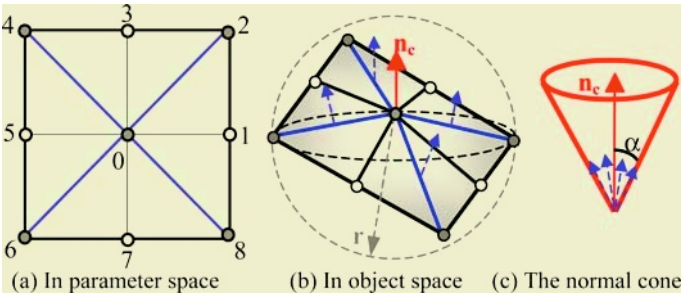


Fig. 2. The structure of a P-Quadtree node

The position of a pixel in the image space of GIM is called the *image position* (u, v) *of the pixel*. The image position of the pixel at left-bottom corner of a node is called *image position* (u_0, v_0) *of the node*, and its size by pixels is called *image size*. We could reconstruct the topological connectivity of the patch of the node by its image position and the image size.

Node Encoding. In order to transmit the data of patches efficiently, we must encode the data of the node for compression. A pointless quadtree [13] is a linear array of nodes in quadtree, and the topological and hierarchical structure of a node could be determined by its localization in the array, or localization code. Therefore we could encode the structure of the node into an integer number by using the localization code of the nodes in the pointless P-Quadtree. This code is the *structure code* of the node, which is noted as P. The encoding process is similar to converting a quinary code into a decimal number.

Besides the topological structure, we need to transmit the geometry of the vertices on the quadrilateral patches (the nodes), which corresponds to the value of pixels in GIM. We note the pixels array as VA. Normal meshes could represent the regular mesh using one float-point data per vertex. The valid nodes on the most top of P-Quadtree would be served as the base meshes for subdivision, and thus the geometry image is transformed from "color" image (the 3D coordinates of a vertex (x, y, z)) to a series of "gray" images (a displacement in the normal direction that is determined by the base mesh). The 3D coordinates of a vertex can be encoded as a float-point number, called *vertex code* of the node.

To facilitate the transmission, the structure code P and vertex codes array VA are encapsulated into a communication packet as a transmission unit of data

stream. Apart from the 3D coordinates on the surface in object space, the image position of each vertex must be available when P-Quadtree and surfaces are constructed on the client.

In fact, the image position a vertex in the node (u, v) can be computed from the image position (u_0, v_0) and the image size d of the node, which can be decoded from structure code P of the node. For example, the image coordinates of the center pixel can be computed as $(u_0+d/2, v_0+d/2)$, and the other pixels could be computed in similar way.

It seems unnecessary to transmit the image coordinates, if we transmit the vertices in the order of their relative position. However, the vertices of the node may be shared with its neighbors. To avoid transmitting the vertices redundantly, we mark the vertices transmitted, and remove them from the vertex array in the following communication packets. So the number of vertices in the vertex array is variable.

To determine whether a vertex has been transmitted, we employ a control code to encode the image coordinates efficiently. Each vertex maps a bit flag in the control code. If the vertex is not transmitted, the flag is set to 0, otherwise to 1. The center vertex of the node must be transmitted, and thus only the four corners need to be controlled by a 4-bits control code. If m is the number of pixels of the GIM, and n is the number of nodes in P-Quadtree, the size of transmitting data for vertices would be no more than $(32\ m + (4/5)n) < 33m$.

4 Streaming and Reconstruction

After the P-Quadtree is prepared, we begin to stream the nodes according to the clients' requests. The data stream can be transmitted through the network. The data streaming is generated on the server in real time, and the surfaces are reconstructed on the client progressively.

4.1 Data Streaming

The breadth-first tree-traversal could fit the data streaming well. Since the shape of charts could be arbitrary polygons, this may cause many gray nodes across the irregular boundary. The gray nodes must be subdivided until all the children are valid or null nodes. The valid nodes would be added to the data stream, while the null nodes are discarded.

Initial Nodes Streaming. To encode the 3D coordinates of vertices, the base meshes are necessary to construct the normal meshes. We firstly traverse the quadtree from the root node to get the valid nodes at the most top level. The obtained nodes are *the initial nodes* for transmission, which also serve as the base meshes of normal meshes.

Detail Texture Transmitting. The normal texture atlas could be utilized to improve the visual appearance of the partial transmitted surfaces. The texture can be transmitted progressively. since the texture coordinates of the texture

atlas are implicit, it is unnecessary to transmit them. The geometry data stream and normal map could be transmitted interactively in our approach.

Silhouette Nodes Streaming. The quality of the rendering image on the silhouette is still poor after normal details texture is mapped. The nodes on the silhouette would be subdivided and the information of sub-nodes be sent to the client to enhance the silhouette of the object. To transmit the nodes near the silhouette in the highest priority, the view-dependently continuous LOD technique could be employed to refine the surfaces on the silhouette. We perform the silhouette test using the algorithms in [10].

The nodes that pass the silhouette test are called *the silhouette nodes*. They are subdivided and packed into the transmitting stream with the highest priority. The front-face nodes toward the viewer, which could be enhanced by normal mapping, and are assigned with the mediate priority. Back-face nodes are culled in current viewpoint.

4.2 Data Transmission

The generated data stream would be transmitted in the client-server mode through a TCP/IP network. The following is the detailed steps:

Session Creation. Once the clients send the request commands and viewing parameters, a session is created. The data setting is prepared to begin the transmission. A flag buffer for each client is created on the server. It is used to avoid redundant nodes transmission.

Initial Nodes Transmission. The initial nodes stream is sent to the clients, and then the grasp of the models can be visualized in a short time. The initial nodes stream can also be served as the base mesh for decoding the coordinates of vertices from the data package received on the clients.

Normal Map Atlas Transmission. After the initial nodes transmission, the client can request for the normal map optionally. This can improve the realism of the rendering image on the client. The normal texture atlas has image structure, and can be transmitted in progressively by the 2D image transmission techniques.

Silhouette Nodes Transmission. To enhance the coarse silhouette, the silhouette nodes stream is requested. The data structure of the silhouette nodes is similar to that of the initial nodes, however the coordinates of vertices for the silhouette nodes are encoded and compressed to a scalar value.

When the viewpoint on the client is changed, the silhouette nodes stream must be generated again. Because the silhouette nodes are highly adaptive, the data transmission is minimal for each frame, and the high flexibility and interactivity are obtained.

Session close. If the transmission has been finished, or the user interrupts the transmitting process, the session is closed, and the flag buffer is released.

4.3 Surface Reconstruction and Rendering

When a communication packet in the data stream is received on the client, the corresponding node's structure could be reconstructed by package decoding.

Since the transmitted nodes are in a sequential structure, we must reconstruct their hierarchical structure firstly. By decoding the structure code, we could reconstruct the topological connectivity and hierarchical structure; the vertices or pixels of the geometry image can be obtained from decoding of the vertex codes. The algorithm to decode the structure code is similar to converting a decimal digit into a quinary code, which is the reversed encoding process mentioned in Section 3.2.

After the data are decoded, the quadtree on the client can be re-constructed. We firstly traverse the quadtree from the root node to locate the position of the coming node in the quadtree according to the decoded quinary digits, and then the point of this position in quadtree is assigned to this node.

Once the quadtree has been reconstructed, we could select the appropriate nodes to reconstruct the surface by using the quadtree traversal. One problem during rendering is that crack may come up when two adjacent nodes have different levels. Since the vertices are transmitted progressively, we check the transmitted vertices on the edge, and then triangulate the patch to prevent from the possible crack. The geometry image can be served as flags map of the received pixels. Once a vertex code is decoded, it is assigned to the corresponding pixel.

A node can be decomposed as a triangle fan. Once the rendering triangles are pre-pared, the data are sent to the rendering pipeline. The normal mapping in our approach can be implemented on GPU by using the per-pixel lighting technique.

5 Results and Discussion

We firstly illustrate the effect of our interactive transmission in Fig.3, where (a) is the mesh of the initial nodes, and (b) is the result of its geometry rendering. Once the normal texture atlas is applied to the initial meshes, the visual effect improved greatly(see Fig.3(c)). However, if the viewpoint is changed, the coarse silhouette is shown (see Fig.3(d)). When the requesting silhouette nodes are transmitted, the silhouette is refined (Fig.3(e)).

The main advantage of our transmission scheme is the interactivity and flexibility. The nodes can be transmitted in arbitrary order, because only the nodes near the silhouette need to be refined. Since the normal texture atlas can enhance the realism of rendering result, the data require to be transmitted is minimal. Fig. 4 shows the meshes for the silhouette refinement effect. Table 1 shows the number of the silhouette nodes per frame. Due to the regular structure of P-Quadtree, our streaming implementation has lower run-time overhead on the server, whose performance results are shown in Table 1. All of the test cases run on a PC with a 2.6GHz Pentium 4 processor with 1G DRAM under windows2000 operating system.

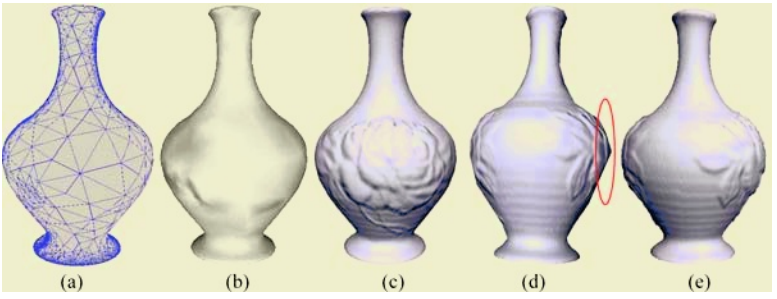


Fig. 3. The interactive transmission of "Vase"

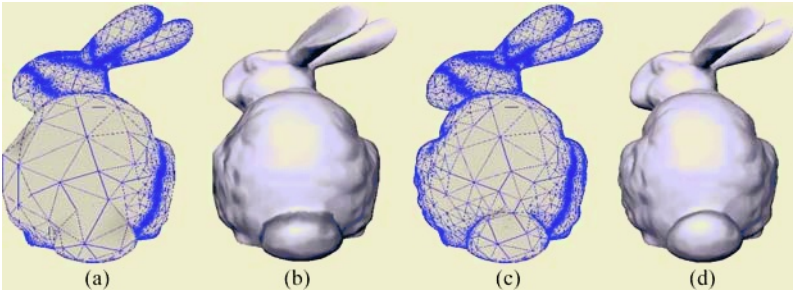


Fig. 4. The meshes in the transmitting process

Comparing with the streaming schemes based on progressive meshes[11][12][7], the polygon number of the initial mesh in our approach is about thousands(see Table 1). However, our approach has some advantages over them. Firstly, more

Table 1. Statistics of P-Quadtree streaming

Model	Igea	Vase	Bunny
# Charts	2	2	9
# Total nodes	99,975	40,559	140,753
# Initial nodes	2352	1647	6613
# Silhouette nodes	1,233	538	2,035
Streaming rate for the P-Quadtree on server(fps)	38	80	21

features can be shown and visual effect is much more better, and the transmitting rate of modern network can be enough to download meshes with thousands triangles quickly. Secondly, although the normal map atlas may be used to enhance the realism of the rendering images in other approaches, the texture coordinates are implicit in our approach because of the "texture-like" structure of GIM. Lastly, the node encoding in our approach is more compact and flexible due to the regular structure of P-Quadtree.

6 Conclusions

We have demonstrated a new interactive transmission approach to highly detailed surfaces. Our approach combines the normal mapping and view-dependent continuous LOD technique. Because only the nodes near the silhouette need to be refined progressively and interactively, only a few nodes are required to be transmitted each viewpoint. Since an efficient encoding scheme is used to compress the geometry data, the transmitting efficiency and the visual effect are improved.

References

1. V. Abadjev, M. del Rosario, A. Lebedev, A. Migdal and V. Paskhaver: MetaStream. Proceedings of VRML. (1999)
2. XF. Gu, S. J. Gortler and H. Hoppe: Geometry Images. Proceedings of SIGGRAPH 2002. (2002) 355–361
3. I. Guskov, K. Vidimce, W. Sweldens, and P. Schröder: Normal meshes. Proceedings of SIGGRAPH 2000. (2002) 95–102
4. H. Hoppe: Progressive Meshes. Proceedings of SIGGRAPH 1996. (1996) 99–108
5. H. Hoppe: View-dependent Refinement of Progressive Meshes. Proceedings of SIGGRAPH 1997. (1997) 189–198
6. I. Khodakovsky, P. Schröder, and W. Sweldens: Progressive Geometry Compression. Proceedings of SIGGRAPH, 2000. (2000)
7. J. Kim, S. Lee, and L. Kobbelt: View-dependent Streaming of Progressive Meshes. To appear at Shape Modeling International 2004
8. U. Labsik, L.Kobbelt, R. Schneider, and H.-P.Seidel: Progressive transmission of subdivision surfaces. Computational Geometry, 15(2000) 25–39
9. A. Lee, H. Moreton, and H. Hoppe: Displaced Subdivision Surfaces. Proceedings of SIGGRAPH 2000. (2000) 85–94
10. D. Luebke, and C. Erikson: View-Dependent Simplification of Arbitrary Polygonal Environments. Proceedings of SIGGRAPH 1997. (1997) 199–208
11. R. Pajarola, and J. Rossignac: Compressed Progressive Meshes. IEEE Transactions on Visualization and Computer Graphics. vol. 6, no. 1 (2000)
12. C. Prince: Compressed Progressive Meshes. Progressive Meshes for Large Models of Arbitrary Topology, M. S. Dissertation, University of Washington. (2000)
13. H. Samet: The Quadtree and Related Hierarchical Data Structures. ACM Computting Surveys, vol. 16, no. 2 (2000) 187–260
14. P. Sander, J. Snyder, S. Gortler, and H. Hoppe: Texture Mapping Progressive Meshes. Proceedings of SIGGRAPH 2001. (2001) 409–416
15. P. Sander, Z. Wood, S. Gortler, J. Snyder, and H. Hoppe: Multi-Chart Geometry Images. Eurographics Symposium on Geometry Processing (2003) 146–155
16. M. Tarini, P. Cignoni, C. Rocchini and R. Scopigno: Real Time, Accurate, Multi-Featured Rendering of Bump Mapped Surfaces. Computer Graphics Forum (Eurographics 2000 Conference Issue),vol. 19(3), (2000) 119–130

Contour-Based Terrain Model Reconstruction Using Distance Information

Byeong-Seok Shin and Hoe Sang Jung

Inha University, Department of Computer Science and Information Engineering
253 Yonghyeon-Dong, Nam-Gu, Inchon, 402-751, Korea
bsshin@inha.ac.kr, g2031372@inhavision.inha.ac.kr

Abstract. In order to create three-dimensional terrain models, we reconstruct geometric models from contour lines on two-dimensional map. Previous methods divide a set of contour lines into simple matching regions and clefts. Since long processing time is taken for reconstructing clefts, performance might be degraded while manipulating complicated models. We propose a fast reconstruction method, which generates triangle strips by computing distance of corresponding vertex pairs in adjacent slices for simple matching region. If there are some branches or dissimilarities, it computes midpoints of corresponding vertices and reconstructs geometry of those areas by tiling the midpoints and remaining vertices. Experimental results show that our method reconstructs geometric models fairly well and it is faster than the previous method.

1 Introduction

Terrain modeling and rendering is an important technology in interactive computer games, geographic information system (GIS) and flight simulation. Commonly used source data for terrain modeling is height field produced by scanning a region with laser range scanners in satellites or airplanes. Since they inherently contain huge amount of sample data it is not easy to manipulate them in consumer PC's. Furthermore, most terrain scanning is performed by irregular sampling such as LIDAR (light detection and ranging) scanning and reconstructed to irregular triangle meshes such as TIN (triangulated irregular network). Considerably long time is required to reconstruct TIN's and it is very hard to simplify that TIN models [1],[2],[3].

Another approach to reconstruct terrain model is to use contour lines on a 2D map. Since a contour line is obtained by sampling the points that have the same height value with regular interval, we can make 3D terrain models by restoring original geometry in-between two consecutive slices. Triangle strips are generated from two contours, which are optimal representation for graphics hardware.

In this paper, we take into account automatic reconstruction methods for 3D terrain models from 2D contour lines. It identifies corresponding vertices on two adjacent contour lines, and reconstructs geometry by connecting the vertices with edges. A commonly used surface reconstruction method proposed by Barequet *et al.* divides a pair of corresponding contour lines into simple matching regions and clefts [4],[5]. It

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1177 – 1186, 2005.

reconstructs triangle meshes by applying partial curve match technique to simple regions, and triangulates cleft regions with dynamic programming method [8]. Although it has an advantage of generating considerably accurate models, it is not efficient to manipulate contour lines containing a lot of clefts. We propose a fast reconstruction method of 3D terrain model. It generates a triangle strip by considering the distance of corresponding vertices for a simple matching region. If there are branches and dissimilarities, our method computes midpoints of corresponding vertices and reconstructs original geometry by connecting the midpoints to the vertices in the cleft region.

In Section 2, we briefly review the related work. In Section 3 and 4, we explain our method and show experimental results. Lastly, we conclude our work.

2 Related Work

Several methods to reconstruct original geometry from contour lines have been proposed [4],[5],[6],[7]. Reconstruction process is composed of three components: correspondence determination, tiling and branch processing [9]. Correspondence determination is to identify the contour lines that should be connected to a specific contour. Tiling is to produce triangle strips from a pair of contour lines. Branch processing is to determine the correspondence when a contour line faces to multiple contour lines.

Assume that contour lines are defined on slices parallel to XY-plane. When a contour is sufficiently close to its adjacent contour(s), it is easy to determine the exact correspondence. Otherwise, we have to exploit prior knowledge or global information about overall shape of target objects [10]. Most of reconstruction methods assume that a pair of corresponding contours should be overlapped while projecting a contour onto its corresponding one in perpendicular direction [11].

In order to cover the area between two adjacent contours with triangle strips, we define *slice chords*. A slice chord is a connecting edge between a vertex of a contour and another vertex on its corresponding one. Bajaj *et al*. defined three constrains for surface to be constructed and derived exact correspondence and tiling rules from those constraints [9]. Since it uses multi-pass tiling algorithm, complicated regions are reconstructed after processing all the other regions. Kline *et al*. proposed a method to generate accurate surface model using 2D distance transformation and medial axes [12]. It takes a long time since it exploits image-space distance transformation and it derive medial axes for all pairs of contours to determine correspondence of vertices.

Branches occur when a slice has N contours and its corresponding slice contains M contours ($M \neq N$, $M, N > 0$). Meyers *et al*. simplified multiple-branch problem into one-to-one correspondence by defining composition of individual contour lines [11]. They proposed a special method for canyons between two neighboring contours of the same slices. Barequet *et al*. [4],[5] and Bajaj *et al*. [9] presented methods for tiling branches that has complex joint topology and shape. In the previous research, we proposed a branch processing method that partitions the contour lines into several sub-contours by considering the number of vertices and spatial distribution and tiles those sub-contours [13],[14].

3 Terrain Model Reconstruction Method

We can separate contour lines on a map into consecutive slices arranged with regular interval. Let a pair of consecutive slices be $<S_n, S_{n+1}>$ and sets of contours belong the slices be C_i^n (i=0,…,M-1) and C_j^{n+1} (j=0,…,N-1). A contour C_i^n is an ordered set of vertices $\{v(x,y,n)\}$ denoted as V_i^n. Fig. 1 shows overview of our reconstruction algorithm. It determines correspondence of contours and checks similarity of them. Basic tiling method is applied to similar region. In case of handling dissimilar regions and clefts, it computes medial axes of the structures and tiles the vertices.

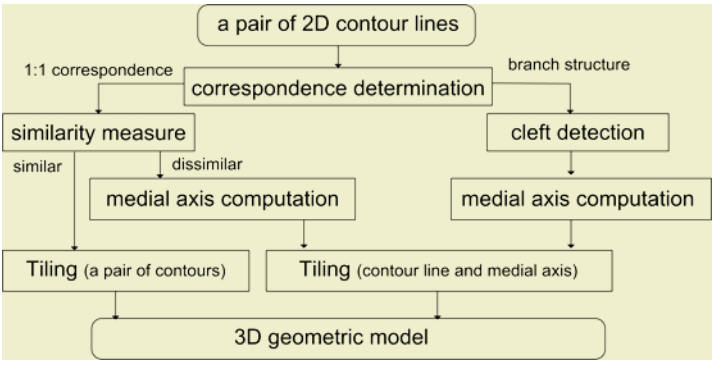


Fig. 1. Overview of the proposed method

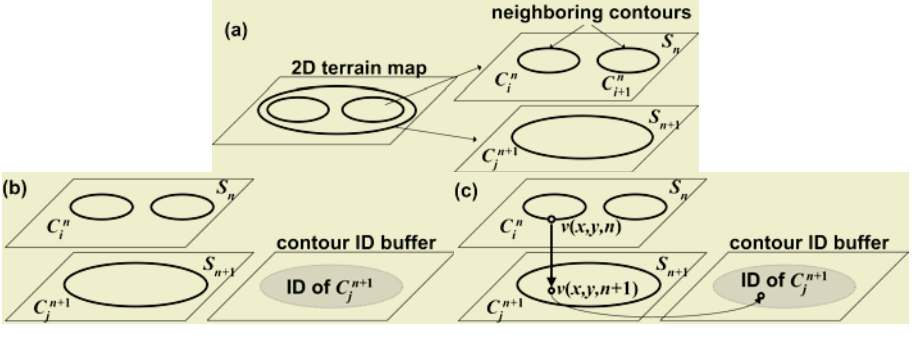


Fig. 2. Correspondence determination between two adjacent contours (a) extracts contour lines from a 2D map (b) fills up interior of C_j^{n+1}as its ID's in contour ID buffer (c) projects vertices of C_i^n onto S_{n+1} and checks the ID of corresponding points

3.1 Correspondence Determination

When a vertex of a contour is projected onto inside of another contour on its adjacent slice, two contour lines are regarded as corresponding to each other. Our method allocates a 2D-buffer that has the same size of a slice and fills up interior of C_j^{n+1} as its identifier with boundary-fill algorithm. Then it projects vertices $v(x,y,n)$ of C_i^n onto a slice S_{n+1} and checks whether the buffer value is identical to the contour ID or not at

the projected position $\{v(x,y,n+1)\}$ (Fig. 2). When the vertices of more than one contours correspond to single contour, the contour lines may produce branches.

3.2 Tiling

Vertices on the same contour maintain connectivity with neighboring ones. However, there is no topological information between two vertices on different slices. Therefore we have to determine the corresponding vertex on the adjacent contour for a specific vertex and produce triangle strips by connecting vertex pairs with slice chords.

In general, topologically adjacent vertices have the smallest distance in comparison to the other vertices. Let $V_i^n(k)$ be the k-th vertex of C_i^n and the number of vertices of C_i^n and C_j^{n+1} are K and L respectively. We define a function $\psi(v)$ that returns the nearest vertex on the corresponding contours as Eq.(1), where $|(v_1,v_2)|$ computes Euclidian distance between v_1 and v_2. Applying $\psi(v)$ to all the vertices of C_i^n and connecting pairs of corresponding vertices with slice chords produces a triangle strip for the contour lines (Fig. 3).

$$V_j^{n+1}(l) = \psi(V_i^n(k)) = \min_{t=0}^{L-1} \{ | (V_i^n(k), V_j^{n+1}(t)) | \} . \tag{1}$$

We need to perform distance computation $O(KL)$ times. Computation cost for the previous methods that exploits *shift value* to determine corresponding vertex is also $O(KL)$ [4],[5],[13],[14]. Since Klein's method calculates 2D distance field for each slice to determine medial axes, it requires more time to determine correspondence between two vertices [12].

When the most recently produced edge of the strip is $e(v_k, v_l)$, a vertex to be added is one of the candidates v_{k+1} and v_{l+1}. So it compares the length of two edges generated by connecting the candidate vertices, and chooses a vertex that produces shorter edge as the next vertex of the strip. When $|(v_k,v_{l+1})| > |(v_{k+1},v_l)|$, v_{k+1} becomes the next vertex, otherwise v_{l+1} is selected as the next vertex.

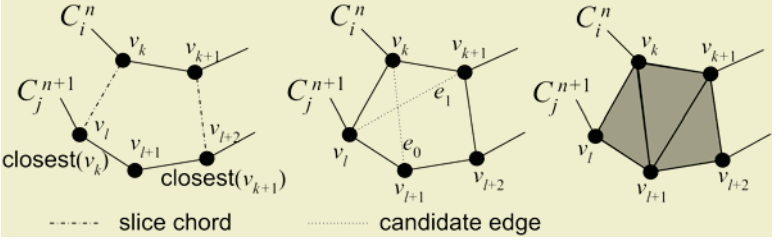


Fig. 3. Basic tiling operation (left) determining the corresponding vertices in consecutive slice (middle) connecting vertices that have the shortest edge (right) resulting triangle mesh

In most cases, corresponding contours have different shape. When the shape of a contour is far different from its corresponding one, dissimilar area may occur as shown in Fig. 4(b). Applying the basic tiling method to dissimilar region may produce incorrect geometric model (Fig. 4 (d)). So we have to devise a method to detect the occurrence of dissimilarity and to reconstruct correct geometry.

The first step is to detect dissimilar region. A function $\mathrm{hop}(v_1,v_2)$ returns the difference of indices between v_1 and v_2. For example, $\mathrm{hop}(V_i^n(k), V_i^n(k+s))$ is s (positive integer) since $V_i^n(k)$ is an ordered set. Even though $V_i^n(k+1)$ is the neighboring vertex of $V_i^n(k)$, $\psi(V_i^n(k+1))$ cannot be a neighbor of $\psi(V_i^n(k))$. When the intervals of indices are regular for consecutive vertices (Fig. 4 (a)), we can reconstruct geometric models applying only the basic tiling rule. However, when an interval of two vertices is different from that of another vertex pair, dissimilarity occurs (Fig. 4 (e)). We can detect dissimilarities on vertex lists by considering the standard deviation of intervals for vertex pairs as in Eq. (2) where δ_{th} is predefined threshold of standard deviation.

After we determine that adjacent contours are dissimilar, we have to find exact dissimilar region using Eq.(3). It is regarded as a dissimilar region when the hop value of an interval from $\psi(V_i^n(k))$ to $\psi(V_i^n(k+1))$ is greater than the average of hop values for all pairs of corresponding vertices.

$$\begin{cases} \text{dissimilar obect} & \text{if } \delta\{\mathrm{hop}(\psi(V_i^n(p)),\psi(V_i^n(p+1)))\} \geq \delta_{th} \\ \text{similar object} & \text{otherwise} \end{cases} . \tag{2}$$

$$\mathrm{hop}(\psi(V_i^n(k)),\psi(V_i^n(k+1)) > avg\{\mathrm{hop}(\psi(V_i^n(p)),\psi(V_i^n(p+1)))\} . \tag{3}$$

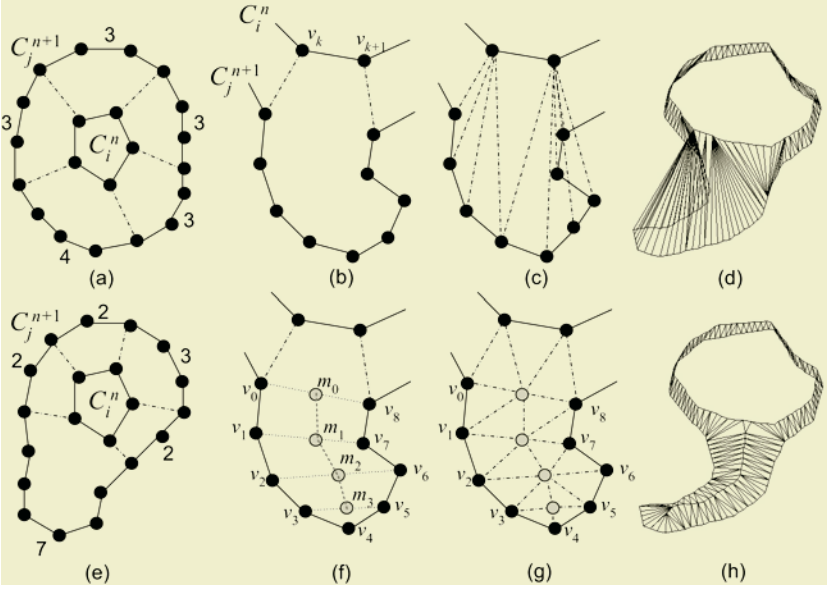


Fig. 4. An example of reconstructing dissimilar region (a) similar contour pair (b) detecting dissimilar region (c)-(d) incorrect triangulation when applying the basic tiling method (e) dissimilar contour pair (f) computing midpoints and medial axis (g) reconstructing dissimilar region with our method (h) reconstructed model with our method

The next step is to generate a medial axis of dissimilar region by computing midpoints of two corresponding vertices on the same contour and connecting them.

Unlike correspondence of simple matching region, index of corresponding vertex on dissimilar region is symmetric along the medial axis. For example, in Fig. 4 (f), v_8 is the corresponding vertex of v_0 and v_7 is for v_1, and so on.

A medial axis is generated from midpoint m_0 to m_i. We can make a triangle strip by connecting the medial axis and partial contour of dissimilar region as shown in Fig. 4 (g). In this case, midpoints are evaluated twice for completing triangle strip, that is, two sub-contours $\{v_0, v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8\}$ and $\{m_0, m_1, m_2, m_3, m_2, m_1, m_0\}$ are used to reconstruct the dissimilar region.

3.3 Branch Processing

A branch occurs when a contour C_j^{n+1} corresponds to N contours $\{ C_i^n, \cdots, C_{i+N-1}^n \}$. First step of branch processing is to identify the clefts in branch area. From a vertex of C_j^{n+1} it starts to find corresponding vertices on adjacent contours sequentially. In the case of 1:1 correspondence of contour pair, $\psi(V_j^{n+1}(l))$ is always located on C_i^n. However, in branch area, $\psi(V_j^{n+1}(l))$ and $\psi(V_j^{n+1}(l+1))$ can belong to the different contour as shown in Fig. 5. This implies that a cleft occurs on that area. In order to distinguish the start and end point of a cleft, we assign tags to the vertices of which the ownership is changed. For example, when corresponding vertices of $\{V_j^{n+1}(p)\}$ $(p=0,..,k)$ are located in C_i^n and $\psi(V_j^{n+1}(l+1))$ belong to its neighboring contour C_{i+1}^n, the vertex $\psi(V_j^{n+1}(l))$ is tagged as E (means 'end of current matched region') and the vertex $\psi(V_j^{n+1}(l+1))$ is tagged as S (means 'start of new matched region'). Regions bounded by C_j^{n+1} and partial contours $C_i^n(S,E)$ are regarded as matched region and they can be reconstructed by applying basic tiling algorithm explained in section 3.2. Remaining areas are clefts and specialized tiling method is used for them.

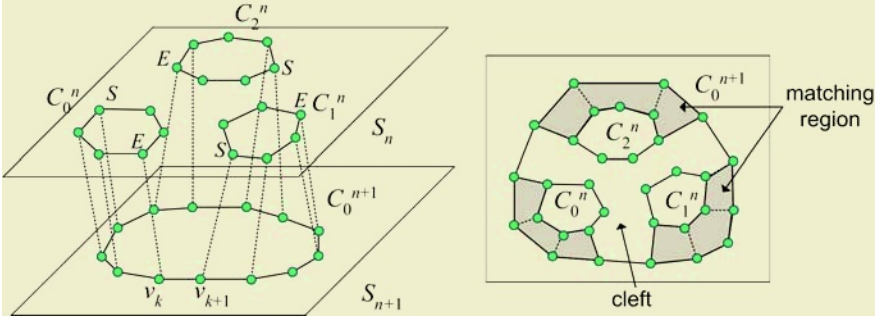


Fig. 5. classification of simple matching region and clefts in branch structure

At first, we define another function $\Phi(v)$ that provides the nearest vertex its neighboring contour (not corresponding contours in adjacent slice) as in Eq.(4).

$$V_j^n(l) = \Phi(V_i^n(k)) = \min_{t=0}^{L-1} \{ | (V_i^n(k), V_j^n(t)) | \} . \tag{3}$$

Branch reconstruction procedure can be described as follows:

(1) For the vertices $V_i^n(k)$ on a partial contour $C_i^n(S,E)$, it computes midpoints of vertex pair $\{V_i^n(k), \Phi(V_i^n(k))\}$ to produce medial axes of a cleft (Fig. 6 (a)-(c)).
(2) It inserts a center vertex v^* for all medial axes to eliminate holes in cleft center. The position of v^* is the center of gravity of the closest end-points of all medial axes (Fig. 6 (d)).
(3) It merges individual axes into single axis that passes through v^* (Fig. 6 (e)).
(4) It produces a triangle mesh from the cleft and the medial axis using the tiling method presented in Section 3.2. In Fig. 7, the cleft contour is composed of vertices $\{v_0', v_0^0, v_0^1, v_0^2, v_0^3, v_1', v_2', v_1^0, v_1^1, v_1^2, v_1^3, v_3', v_4', v_2^0, v_2^1, v_2^2, v_2^3, v_5'\}$ and the medial axis contour is composed of vertices $\{m_0, m_1, v^*, m_2, m_3, m_2, v^*, m_4, m_5, m_4, v^*, m_1\}$. Resulting triangle strip is shown in Fig. 6(f).

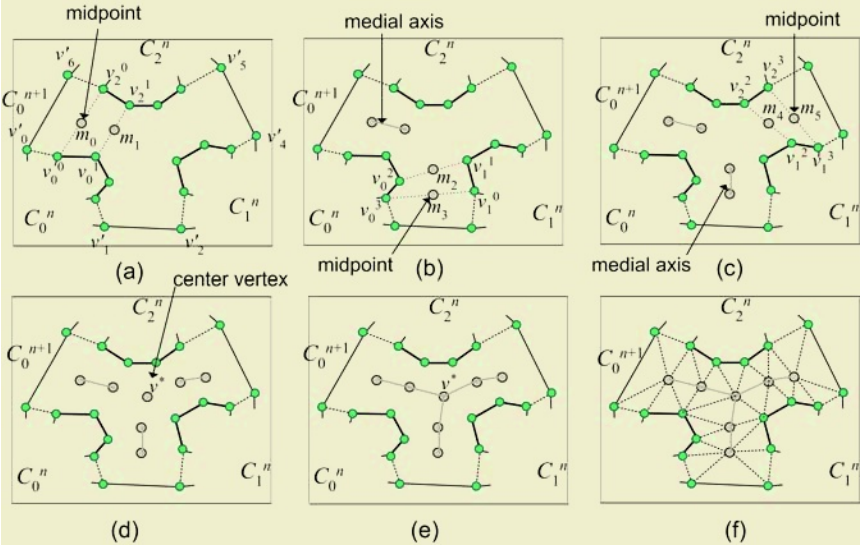


Fig. 6. Tiling of cleft region (a)-(c) computing medial axes of the cleft region (d)-(e) merging all the axes into single medial axis using center vertex (f) resulting triangle mesh

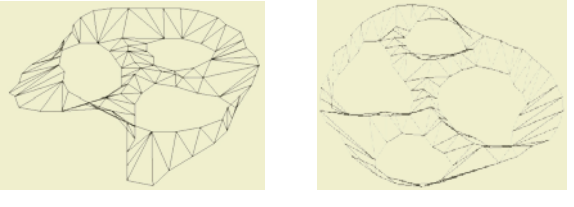


Fig. 7. An example of tiling a contour pair that contains 1:3 (left) and 1:4 (right) branches

It can be applied to any type of 1:*N* branches. Combining triangle meshes for simple matching region and clefts region can produce complete geometric model for branch structure. Fig. 8 shows reconstructed models for 1:3 and 1:4 branches.

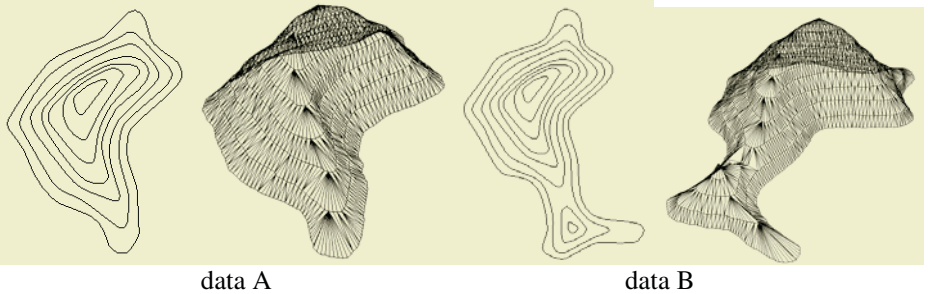
4 Experimental Results

In order to show how efficient our method, we implement terrain model reconstruction system which extracts vertex lists for contour lines from 2D map, and produces three dimensional models. It is implemented on a PC equipped with Pentium IV 2.2GHz CPU, and 1GB main memory. We use a set of contours of 2D maps for several areas.

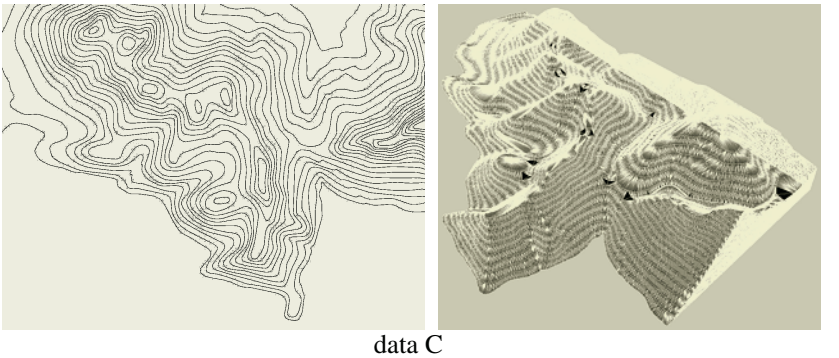
Fig. 8 depicts images produced by rendering 3D terrain model with our method. It can reconstruct accurate models from contour lines even when we deal with considerably complex regions. It produces geometric models fairly well not only for simple matching area but also for dissimilarities and branches.

Table 1 shows the time to be taken for reconstruction of geometry according to the size of data (number of slices and the number of vertices on a contour) and complexity (shape of contour lines and existence of branches). Experimental results show that our method can reconstruct arbitrary shaped models using contour lines on a 2D map within short time.

Fig. 9 shows a comparison of reconstructed models using Berequet's method and ours. Table 2 represents how much time is taken while applying these two methods. Our method can produce geometric models fairly well and take less time in comparison to Berequet's method since it exploits simple distance computation.



data A data B



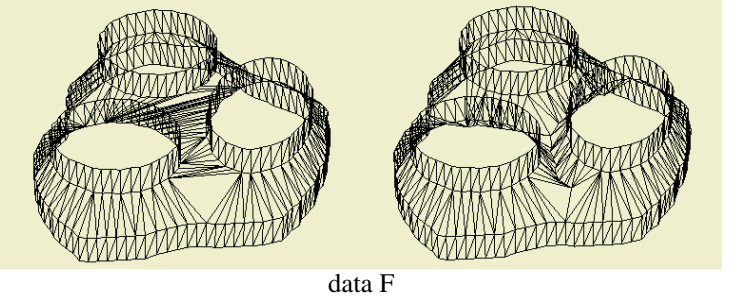
data C

Fig. 8. Examples of reconstructed models. Left image of each pair presents 2D contour lines for specific regions and right one is wire-frame of reconstructed model

Table 1. The time to be taken for reconstructing several geometric models using our method

	Existence of branches	Number of contous	Number of polygons	Reconstruction time (sec)
data A	no	7	1393	0.67
data B	yes	10	2136	1.09
data C	yes	43	43185	50.74

data E



data F

Fig. 9. A comparison of reconstructed models using Barequet's method (left column) and our method (right column)

Table 2. A comparison of reconstruction time for Barequet's method and ours

	methods	Number of contous	Number of polygons	Reconstruction time (msec)
data E	Barequet's	6	1336	47
	ours	6	1437	15
data F	Barequet's	8	704	31
	ours	8	768	8

5 Conclusion

Terrain modeling and rendering is an important technology in large-scale virtual environment rendering. We propose a simple and fast reconstruction method of 3D terrain

model from contour lines. It generates triangle strips by computing distance of corresponding vertex pairs in adjacent slices for simple matching region. If there are some branches or dissimilarities, it computes midpoints of corresponding vertices and reconstructs geometry of those areas by tiling the midpoints and remaining vertices. Experimental results show that our method reconstructs geometric models fairly well and it is faster than the previous method.

Acknowledgement

This work was supported by the Ministry of Information & Communications, Korea, under the Information Technology Research Center (ITRC) Support Program.

References

1. Cohen-Steiner, D., Colin de Verdiere, E., Yvinec, M.: Conforming Delaunay triangulations in 3D. Proceedings of 18th Symposium on Computational Geometry (2002) 199-208
2. Attali, D., Boissonnat, J.D., Lieuter, A.: Complexity of the Delaunay Triangulation of points on surfaces. Proc. of 19th Annual ACM Symposium on Comput. Geometry (2003) 201-210
3. Erickson, J.: Dense point sets have sparse Delaunay triangulations. Proceedings of 13th Annual ACM-SIAM Symposium on Discrete Algorithms (2002) 125-134
4. Barequet, G., Sharir, M.: Piecewise-Linear Interpolation between Polygonal Slices. Computer Vision and Image Understanding, Vol. 63, No. 2 (1996) 251-272
5. Barequet, G., Shapiro, D., Tal, A.: Multilevel Sensitive Reconstruction of Polyhedral Surfaces from Parallel Slices. The Visual Computer, Vol. 16, No. 2 (2000) 116-133
6. Fuchs, H., Kedem, Z., Uselton, S.: Optimal Surface Reconstruction from Planar Contours. Communications of the ACM, Vol. 20, No. 10 (1977) 693-702
7. Kehtarnavaz. N., Simar, L., Figueiredo, R.: A Syntactic/Semantic Technique for Surface Reconstruction from Cross-sectional Contours. CVGIP, Vol. 42 (1988) 399-409
8. Klincsek, G.: Minimal Triangulations of Polygonal Domains. Annals of Discrete Mathematics, Vol. 9 (1980) 121-123
9. Bajaj, C., Coyle E., Lin, K.: Arbitrary Topology Shape Reconstruction for Planar Cross Sections. Graphical Models and Image Processing, Vol. 58, No. 6 (1996) 524-543
10. Soroka, B.: Generalized Cones from Serial Sections. Computer Graphics and Image Processing, Vol. 15, No. 2 (1981) 154-166
11. Meyers, D., Skinner, S., Sloan, K.: Surfaces from Contours. ACM Trans. on Graphics, Vol. 11, No. 3 (1992) 228-258
12. Klein, R., Schilling, A., Strasser, W.: Reconstruction and Simplification of Surface from Contours. 7th Pacific Conference on Computer Graphics and Applications (1999) 198-207
13. Shin, B., Jung, H.: Fast Reconstruction of 3D Terrain Model from Contour Lines of 2D Maps. Lecture Notes on Artificial Intelligence, Vol. 3398 (2005) 230-239
14. Shin, B. : Reconstruction of Human Anatomical Models from Segmented Contour Lines. Lecture Notes on Computer Science, Vol. 3314 (2004) 619-624

An Efficient Point Rendering Using Octree and Texture Lookup

Yun-Mo Koo[1] and Byeong-Seok Shin[2]

[1] Myongji College, Department of Computer Science and Information
356-1 Hongeun3-Dong Seodaemun-Gu, Seoul, 120-776, Korea
yunmo@mail.mjc.ac.kr
[2] Inha University, Department of Computer Science and Information Engineering
253 Yonghyeon-Dong, Nam-Gu, Inchon, 402-751, Korea
bsshin@inha.ac.kr

Abstract. As modern 3D scanning devices handle an enormous amount of point data, generation of triangle mesh becomes time-consuming job. Furthermore, projected triangles are smaller than pixel size, thereby increasing overhead for rasterization. In recent years point-based rendering has become an efficient method for the rendering of complex models. We propose an acceleration method for rendering of point-based geometry. It solves the visibility of point samples by taking advantages of octree structure constructed directly from a point cloud. This enables graphics hardware to render a scene in a single pass thus avoiding additional pass for visibility computation. In addition, we also present an efficient splatting technique to use a lookup table of alpha textures, resulting in alleviation of the load of pixel processing. It achieves better performance over the other methods.

1 Introduction

Point-based rendering has recently become a promising alternative to traditional triangle-based rendering [1],[2]. Rapid advances in 3D scanning system have increased the number of samples of an object over hundreds of millions. In case of generating triangle mesh from the point samples, several triangles could be projected on a single pixel during rendering, leading to redundant computation for triangle rasterization. On the contrary, in point-based rendering, the preprocessing time is relatively short. Therefore, point-based rendering would be a better choice for rendering complex models.

The point-rendering pipeline consists of several steps: forward mapping of points, per-point shading, visibility computation and blending splatting. To support correct visibility, Levoy and Whitted proposed a modified A-buffer algorithm which gathers all contributions for a given pixel, sorts the contributions depending on their depths, and finally performs bending operation for the contributions belonging to the same surface [3],[4]. In hardware-based methods, Z-buffer is commonly used for visibility calculation [1],[5],[6]. However, due to the conflict between Z-buffering and blending in hardware processing, most hardware-based approaches use two-pass rendering scheme. In the first pass, called *visibility splatting*, they render the scene to Z-buffer

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1187 – 1196, 2005.

with all z-values having an offset added to them. The second pass performs *blending splatting* of visible points of which the depth values are less than the z-values.

Coconu and Hege introduced a single-pass algorithm [7]. This algorithm ensures visibility order by using their octree representation similar to an LDI (layered depth image) tree and by warping the hierarchy in back-to-front order. Construction of the LDI tree is time consuming since LDIs are obtained by running a ray tracer several times or projecting images [6],[7]. In this paper, we propose an efficient single-pass algorithm. Although the algorithm is also an octree-based method, it constructs the octree directly from a point cloud without conversion to LDIs. Therefore, our method reduces preprocessing time.

Recent hardware-based methods support fuzzy splat [1],[5],[6],[7] with elliptical Gaussian filter which improves image quality. In the context of GPU-based splatting, the elliptical Gaussian filter is implemented as an alpha texture with its alpha channel which is computed with a radially decreasing Gaussian. The algorithm in [6] generates a rectangle for each point sample and renders the textured rectangle, thereby increasing the number of processed vertices by four times. Other schemes splat a point sample using a feature of the graphics hardware called *point sprites* [5],[7]. They apply the alpha texture of Gaussian filter to a square of a point sprite by modifying the pixel-processing step. Although these present efficient implementations without increasing the number of vertices, bottleneck may occur due to heavy computations in pixel processing. We solve this problem by alleviating the loads of pixel processing stage. Our method stores precomputed splat kernels in the texture memory, and performs lookup operation to get its alpha during pixel processing.

Related work is summarized in Section 2. In the next section, we explain our octree manipulation and GPU-based splatting in detail. Experimental results are shown in Section 4 and finally we present the conclusions of our work.

2 Related Work

Grossman and Dally have presented efficient algorithms for the generation and rendering of point sets [8],[9]. They developed QSplat to interactively render the large datasets of the Digital Michelangelo Project [1]. It uses hardware acceleration and combines a hierarchy of bounding spheres with splatting.

Pfister et al. defined surface elements as *surfels* [2]. Their method consists of sampling surfels, rendering the surfels using the hierarchy of LDI tree, visibility computation with Z-buffer and image construction. Zwicker et al. introduced surface splatting by screen space EWA (elliptical weighted average) filtering [10] originated from the EWA filter introduced by Heckbert [11]. The approach focuses on the improvement of image quality. Ren et al. reformulate the screen space EWA filter to object-space filtering to use graphics hardware [6]. However, it also uses two-pass scheme.

Botsch and Kobbelt presented a hardware acceleration method based on two-pass rendering [5]. It computes the weight of a pixel with the function of the distance of its corresponding point to the splat's center in pixel processing stage. Coconu and Hege proposed to use an octree-based spatial data structure containing points and triangles to do visibility calculations without Z-buffer. The approach deforms an alpha texture according to camera space normals in pixel processing stage [7].

Most recently, Dachsbacher et al. introduce sequential point trees, a data structure that allows adaptive rendering of point sets completely on the graphics hardware [12]. Although sequential point trees are based on a hierarchical representation, the traversal is replaced by sequential processing on the graphics processor. Since GPU performs most computation, the approach achieves high speed. However, while the splat rate is generally high for sequential point trees, they only work efficiently in frame rate if the object fits in cache and they are observed at a distance [16].

3 Point-Based Rendering Algorithm

Fig. 1 shows the overview of our method. In the preprocessing stage, the algorithm builds up an octree directly from a point cloud without converting to any other structure. Section 3.1 describes our octree in detail. Limited number of splat shapes are stored in the texture memory before rendering.

In the rendering stage, the algorithm traverses the octree in back-to-front order recursively until it reaches a splat whose projected size is smaller than the user-defined threshold. A set of visible splats is delivered to vertex buffer and processed by vertex shader and pixel shader.

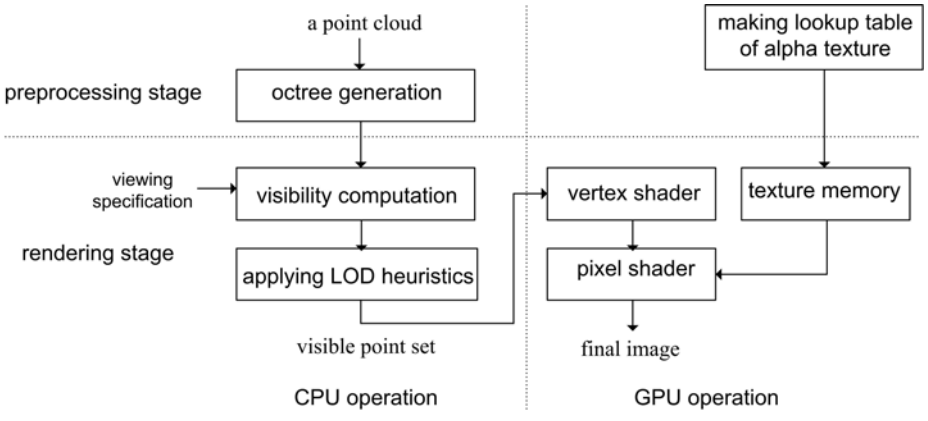


Fig. 1. Overview of our method. It is composed of preprocessing and rendering stages. In preprocessing stage, it generates an octree. After specifying viewing condition, it computes visibility and determines LOD level. The visible point set is supplied to vertex shader to compute projected position, normal and color. Pixel shader draws each pixel with the weight from a lookup table of splat textures

3.1 Octree Representation of Bounding Spheres

Octree in our method is similar to that of QSplat as a hierarchy of bounding spheres except that its nodes have eight children. A leaf node contains geometric information for a point sample such as position, radius and normal. Intermediate nodes have average values of these properties for the subtrees.

The preprocessing stage starts to build up an octree by splitting a bounding box containing point samples along medial axes located on the middle. It recursively computes eight subtrees until each leaf node contains a single point.

3.2 Visibility Computation

Our method determines the visibility of all the point samples by traversing the octree in back-to-front order. Since the octree presented in [7] is composed of LDI blocks, warping operation at children blocks should be performed. To guarantee correct painting order Coconu et al. extend the warping algorithm introduced in [13],[14]. Since octree used in our method is composed of bounding spheres, our algorithm simply splats visible nodes in correct display order.

There are three different ways of how children nodes can be accessed: *view order*, *distance order* and *direction order* [15]. View order is to determine the access order according to the principal viewing direction. However, in perspective projection, erroneous regions may appear. Distance order is to determine the access order according to the distance from the eye position to each node. Direction order is to use the vector from eye position to split point of subspace represented by currently processed parent node. Direction order is an equivalent approximation of distance order and is faster than the distance-order traversal. This is because it selects traversal order with single vector.

Our method exploits direction order. For best efficiency, it computes a new direction vector only in the node requiring modification. The view point determines eight octants separated by three planes intersecting the view point. If the region of visiting node is fully confined in an octant, directions of viewing rays at all positions in the node are identical. Thus any child of the current node can be accessed in the same order without recomputing the access order. On the contrary, if the current node intersects the planes, the access order for its children has to be recomputed. Fig. 2 showsan example of determining the access order. The number written in each block

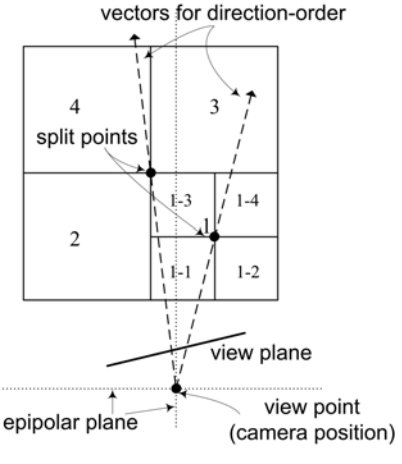


Fig. 2. How to determine the access order at each node of octree

represents access order. In the cases of nodes 2, 4, 1-2 and 1-4, the access order can be reused for all the subtrees. However, at nodes 1, 3, 1-1 and 1-3, the access order has to be recomputed for each viewing ray. Fig. 3 explains demonstrate traversing of octree in correct access order.

Our method provides LOD selection. If a large data set is rendered at low resolution, several points will be projected to a region smaller than the pixel size. In order to avoid unnecessary splatting, selection of appropriate LOD level is required. For the LOD selection, the algorithm uses the heuristic which relies on the projected size. The algorithm traverses the octree recursively up to the splat of which the projected size becomes smaller than the pixel size.

```
TraverseOctree (node, order)
{
    if (node is visible and leaf )        splat a point
    else
        if (the size of splat is below a user-defined threshold)        splat the node
        else {
            if(current node intersects the octants separated by the planes)
                establish a new access order (new_order) by computing
                    direction vector as one from camera position to a split point
            else   new_order = order
            TraverseOctree(child_node, new_order)
        }
}
```

Fig. 3. Visibility computation algorithm with octree

3.3 Splatting

We present a new method exploiting a lookup table of precomputed splat kernels to accelerate splatting. For the implementation, a vertex shader computes the projected position of each splat, a normal in camera space and lighting. A pixel shader determines each pixel with the weight derived from a lookup table of splat kernels.

3.3.1 Splatting Using Precomputed Splat Kernel

A point sample on the surface represents a small disc in object space. A circular Gaussian of this small disc is projected onto an elliptical Gaussian on the image plane. The shape of the ellipse depends on the corresponding normal in camera space.

In order to accelerate splatting, we use a precomputed lookup table of alpha textures. This idea is originated from Heckbert's work called *image pyramid* which is generated and indexed by quantizing the five parameters of an ellipse: x and y position, minor and major radii, and orientation [11]. Although this is suitable for software implementation, it is difficult to implement on current graphics hardware. Hardware does not support quantizing all the ellipse parameters and indexing the image pyramid to access the corresponding resampling filter. Therefore, we propose a method to utilize the pre-computed resampling filter. By storing several splat kernels as a lookup

as a lookup table in the texture memory of graphics hardware, the pixel-processing stage only gets an alpha texture from the lookup table.

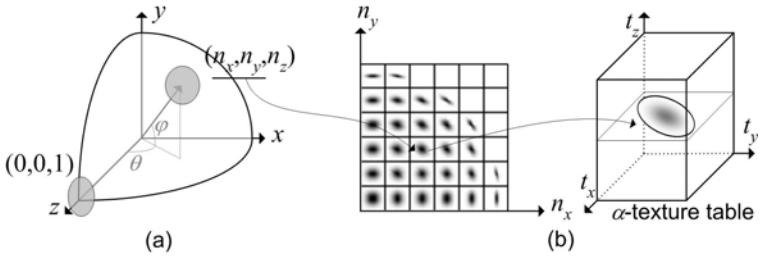


Fig. 4. An example of storing alpha textures into a lookup table (a) how to generate alpha textures according to normals (b) table of alpha textures

In order to compute elliptical Gaussian, we first make an initial alpha texture with a circular Gaussian in preprocessing stage. Let the splat kernel be a circle of which the normal be (0,0,1). Each elliptical alpha texture is computed by applying a rotation matrix that transforms the initial normal to that of current splat. Fig. 4(a) illustrates how to generate a set of alpha textures. Assuming that the normal is given as (n_x, n_y, n_z), we can then compute the angles (θ, ϕ) and the corresponding rotation matrix. The algorithm builds elliptical alpha textures for quantized normals, and stores them into lookup table. Once a normal vector is normalized, only two components n_x and n_y are needed for identifying it. Fig. 4(b) illustrates the table of alpha textures corresponding to the quantized normals.

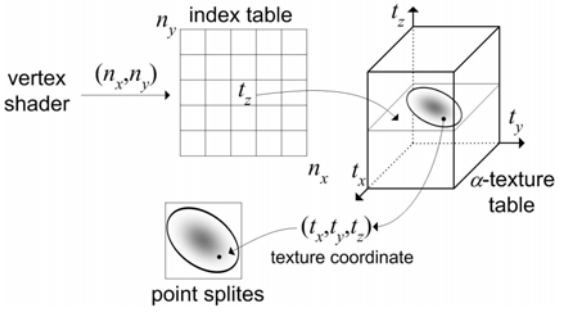


Fig. 5. How to apply an alpha texture to the region of a point sprite in pixel shader

Once the lookup table has been constructed, it is treated like a constant array. Our method stores the alpha textures into the texture memory of graphics hardware before rendering. Each point sample is projected into a square determined by point sprites, and the region is textured with a corresponding alpha texture. To refer the corresponding texture, the third component of the texture coordinates (t_z) should have the information for the normal. However, since a channel can store only two components of a

normal, the algorithm uses another two-dimensional texture containing an index corresponding to the normal n_x and n_y. Fig. 5 demonstrates the projection of each alpha texture to the region determined by point sprites.

In [17], pre-computed splat kernels are used. The authors implemented it in their software-based algorithm. They should have built several resolutions of splat mask for magnified image. However, in our method one resolution of splat mask is sufficient by taking an advantage of the texture sampling operation which enables GPU to linearly interpolate.

3.3.2 Vertex and Pixel Processing

The operation of vertex shader is performed per point splat. The vertex shader computes a position and a normal in camera space with a viewing matrix as a composition of world-view-projection matrix. A point size on the fly is computed in CPU according to the viewing matrix and passed to the vertex shader. This value is used for determining the square size of a point sprite. The vertex shader passes the position, normal and color down to the pixel shader using output registers. While passing a normal, the sign and absolute values of each component are separately transferred. When a normal points to the opposite direction, alpha texture should be reversely addressed during the pixel processing.

The pixel shader modifies pixels in the region of point sprites. The operation of pixel shader is only to fetch an alpha texture in the lookup table. The pixel shader use the texture coordinate (t_x, t_y, t_z) to access the texture, where the t_z is the index for the table containing the alpha textures. Finally the pixel shader outputs the color received from vertex shader with the new fetched alpha. Since our pixel shader is very short, we can achieve improvement in speed.

4 Experimental Results

Experiments were made on a 2.8GHz Pentium4 PC with 1GB memory and an ATI Radeon 9800 graphics card. Our method was applied to four different models for evaluating rendering time and image quality.

Table 1 shows octree construction time for each model. While the method of [7] took a couple of minutes in building an LDI tree, our method took only a few seconds since it constructs octree directly from a point cloud.

Table 2 shows rendering times for test models. The number of visible splats indicates the number of actually selected, visible and drawn points per frame, not including any culled points. The #splat/s also indicates the pure splat rate as the number of visible splats multiplied by fps(frames per second). While in [16] the splat rate is measured as the number of actual splats divided by the splatting time excluding LOD selection time, in [5],[7] it is measured as the number of the whole points representing each model multiplied by fps. Therefore, it is difficult to compare rendering performance only with the splat rate. In addition, the performance varies according to the portion the objects occupy. In the experiments, the size of viewport is 512 × 512 and test models are displayed to occupy 50% of the viewport. Our method achieves a splat rate of about 4M filtered splats. However if it is measured in the same way as [5],[7], it implies about 14M, thus overcoming the others.

Table 1. Octree construction time for four different models

Model	Number of points	octree generation time (sec)
Bunny	35130	0.172
Venus	45367	0.218
Burner	561916	2.782
Budda	543652	2.938
Dragon	422525	2.016

Table 2. Rendering time for four different models

model	number of visible splats	FILTER fps (#splat/s)	NOFILTER fps (#splat/s)	Overhead due to the filtering
Bunny	17487	216.16(3.8M)	225.34(3.9M)	4.0%
Venus	22795	171.33(3.9M)	179.9(4.1M)	4.7%
Burner	255884	26.18(4.8M)	28.07(5.1M)	6.7%
Budda	245525	28.54(4.7M)	30.32(5.0M)	5.8%
Dragon	207193	21.24(4.4M)	22.19(4.6M)	4.3%

Fig. 6. Rendered images applying our method to Budda, Burner, Venus (from left to right in top) , Bunny and Dragon (bottom) models

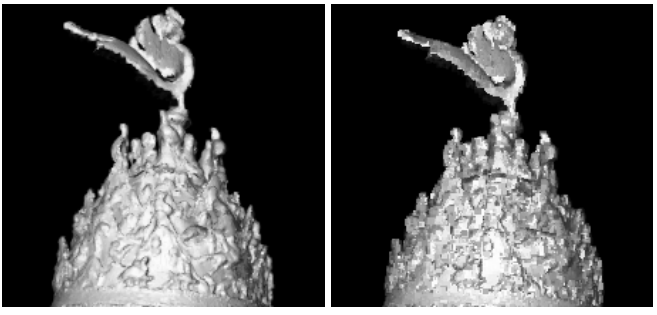


Fig. 7. Comparison between rendered images with Gaussian splat (left) and quad splat (right)

First we measure the rendering time in the application of resampling filter (FILTER) and then estimate the time for rendering a scene without filtering (NOFILTER). The difference of the rendering times for each case shows how much time is taken in filtering during pixel shader stage. In our method, about 4~6% of rendering time is spent in filtering. However, in [5] additional time in filtering is about 57%~63% and 38%~56% of rendering time in [7].

Fig. 6 depicts the images rendered with our method for the models. Fig. 7 shows the magnified images using Gaussian filter (left) and quad splat (right) respectively. We can see that rendering with Gaussian filter can produce high quality images in comparison with using quad splat. Since the alpha texture is set to be interpolated linearly in GPU, the quality of images can be maintained even in magnified images. The resolution of each alpha texture would not affect the quality owing to the texture operations of linear interpolation. This allows us to make the alpha textures in low resolution, resulting in less consumption of texture memory. Our implementation uses alpha textures with resolution of 16×16, for 256 different normals. Therefore, the lookup table of alpha textures requires a small amount of the GPU memory.

5 Conclusion

In this paper, we presented an efficient point-based rendering algorithm that takes advantages of GPU. The algorithm renders a scene in a single pass avoiding additional pass for visibility splatting by utilizing an octree of bounding spheres. We intended to devise a method for rendering point clouds directly without converting to LDI structure. If a point cloud could be easily achieved from 3D scanners, it is convenient to display the point cloud itself. Furthermore, our algorithm accelerated the computation of pixel processing by using a lookup table of alpha textures. Since fill-rate is a bottleneck, it is important to reduce the rendering time of pixel processing for the best efficiency.

Acknowledgement

This work was partly supported by grant No. R05-2004-000-12187-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

1. Rusinkiewicz, S., Levoy, M.: QSplat: A Multiresolution Point Rendering System for Large Meshes. ACM SIGGRAPH (2000) 343-352
2. Pfister, H., Zwicker, M., van Baar, J., Gross, M.: Surfels: Surface Elements as Rendering Primitives. ACM SIGGRAPH (2000) 335-342
3. Carpenter, L.: The A-buffer, an Antialiased Hidden Surface Method. ACM SIGGRAPH (1984) 103-108
4. Levoy, M., Whitted, T.: The Use of Points as a Display Primitive. TR 85-022, Computer Science Department, University of North Carolina at Chapel Hill (1985)
5. Botsch, M., Kobbelt, L.: High-Quality Point-Based Rendering on Modern GPUs. Proc. 11th Pacific Conference on Computer Graphics and Applications (2003) 335-343
6. Ren, L., Pfister, H., Zwicker, M.: Object Space EWA Surface Splatting: A Hardware Accelerated Approach to High Quality Point Rendering. Computer Graphics Forum Proceeding of Eurographics (2002) 461-470
7. Coconu, L., Hege, H.: Hardware-Accelerated Point Based Rendering of Complex Scenes. Proceeding of Eurographics Workshop on Rendering (2002) 41-51
8. Grossman, J.P.: Point Sample Rendering. Master's thesis, Department of Electrical Engineering and Compute Science, MIT (1998)
9. Grossman, J.P., Dally, W.J.: Point Sample Rendering. Rendering Techniques '98, Eurographics (1998) 181-192
10. Zwicker, M., Pfister, H., van Baar J., Gross, M.: Surface splatting. ACM SIGGRAPH (2001) 371-378
11. Heckbert, P.: Fundamentals of Texture Mapping and Image Warping. Master's thesis, UC Berkeley, Department of Elelctrical Engineering and Computer Science (1989)
12. Dachsbacher, C., Vogelgsang, C., Stamminger, M.: Sequentail Point Trees. ACM SIGGRAPH (2003) 657-662
13. McMillan, L.: Computing Visibility Without Depth. Technical Report. TR 95-047. University of North Carolina at Chapel Hill (1995)
14. McMillan, L.: A List-Priority Rendering Algorithm for Redisplaying Projected Surfaces. Technical Report. TR 95-005. University of North Carolina at Chapel Hill (1995)
15. Neubauer André: Cell-Based First Hit Ray Casting. Diploma Thesis. http://www.vrvis.at/vis/resources/DA-ANeubauer/node17.html (2001)
16. Sainz, M., Pajarola, R., Lario, R.: Points Reloaded : Point-Based Rendering Revisited. Proceedings of Eurographics Symposium on Point-Based Graphics (2004) 121-128
17. Botsch, M., Wiratanaya, A., Kobbelt, L.: Efficient High Quality Rendering of Point Sampled Geometry. 13th Eurographics Workshop on Rendering (2002) 53-64

Faces Alive: Reconstruction of Animated 3D Human Faces

Yu Zhang, Terence Sim, and Chew Lim Tan

School of Computing, National University of Singapore, Singapore 117543

Abstract. This paper presents a new method for reconstructing anatomy-based, animatable facial models with minimal manual intervention. The technique is based on deforming a multi-layered prototype model to the acquired surface data in an "outside-in" manner: deformation applied to the skin layer is propagated, with the final effect of deforming the underlying muscles. In the skin layer deformation, the generic skin mesh is represented as a dynamic deformable model which is subjected to internal force stemming from the elastic properties of the surface and external forces generated by input data points and features. A fully automated approach has been developed for deforming the muscle layer that includes three types of muscle models. Our method generates animatable models from incomplete input data and reconstructed facial models can be animated directly to synthesize various expressions.

1 Introduction

As virtual faces appear more frequently in various fields such as virtual reality, entertainment, cosmetic and surgical operation simulation, it becomes increasingly important to reconstruct animatable, individualized faces. In this task we are usually confronted with two conflicting goals: one is the requirement for accurate reproduction of face shape, the other is the demand for an efficient representation which can be animated easily and quickly. Current dense surface measurement techniques allow us to generate precise 3D shapes of faces by using 3D shape acquisition systems such as a range scanner, a stereo photogrammetry system, or an active light striper. However, using the range data for face reconstruction is widely known to suffer from several key problems:

- absence of functional structure for animation;
- irregular and dense surface data that can not be used for optimal animatable model construction and realtime animation;
- incomplete data due to shadowing effects or bad reflective properties of the surface.

To address these problems, we propose a new approach to efficient reconstruction of animatable 3D faces of real human individuals. The technique is based on deforming a prototype model to the acquired surface data in an "outside-in" manner: deformation applied to the external skin layer is propagated, with the final effect of deforming the underlying muscles. A global alignment is first carried out to align the prototype model

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1197–1208, 2005.

with the scanned data based on measurements between a subset of specified anthropometric landmarks. The generic skin mesh is then modeled as a dynamic deformable surface. Deformation of the mesh results from the action of internal force which imposes surface continuity constraints and external forces which attract the surface such that it fits the data. We automatically deform the underlying muscle layer that includes three types of muscle models. The reconstructed facial models can be animated directly to synthesize various expressions.

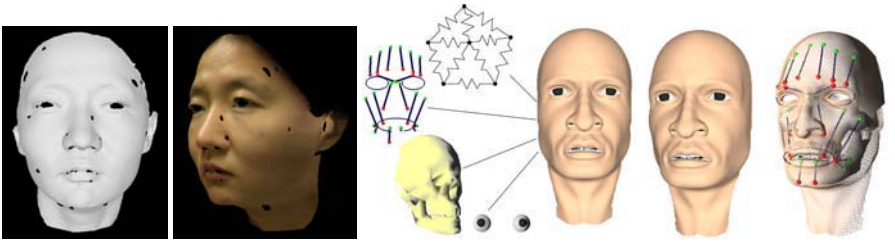
This paper is structured as follows. Section 2 reviews the previous work. Section 3 presents the acquired surface data and our prototype model. Section 4 and 5 describe automatic deformation of the skin and muscle layers, respectively. We show our results in Section 6 and conclude in 7.

2 Previous Work

Techniques for animating faces have been an active area of research over the last 30 years [15, 16]. See [14] for a good survey. Particularly, a few approaches have been proposed for individualized face modeling. The early parametric models attempt to generate different face shapes by modifying conformation parameters of a generic model [15]. However, developing a parameterization flexible enough to create any possible face is challenging and manual parameter tuning is time-consuming.

In more efficient image-based method, a generic face model is taken and facial features extracted from photos are used to modify the generic model using a geometric deformation [1, 6, 7, 10, 12]. Although this kind of technique can provide reconstructed facial models easily, it does have drawbacks such as too few features to guarantee accurate face shape reconstruction and too much loose automatic method used to modify non-feature points. Pighin et al. [18] combine 2D morphing with 3D transformations of the geometric model to produce photorealistic facial animation. Animation is limited, though, since each expression has to be captured in advance and morphing expressions is achieved by linear blends. Blanz et al. [2] are able to generate a head model from only a single photograph. Animation in images and video is possible by transferring the learned expressions in the database to the reconstructed 3D face, and rendering it back to the original image or video.

Decarlo et al. [3] construct a range of static facial models with realistic proportions using a variational constrained optimization technique based on anthropometric measurements. Some approaches morph a generic facial mesh into specific shapes with scattered data interpolation technique [9, 20]. This technique can smoothly interpolate the desired change in shape defined by the vector-valued offset between the manually defined features on the generic mesh and those on the face geometry of a specific person. Waters and Terzopoulos [22] and Lee et al. [11] conform a generic face mesh to a specific person's face by locating the individuals facial features in a cylindrical range image. However, their conformation method in the 2D image domain is not robust in case of nonsalient facial features and/or unfavorable lighting conditions. And it uses the geometric scaling for facial feature fitting. Since 3D positions of the mesh nodes are recovered by sampling the range image, interpolating missing data of the range image (filling holes) is necessary. Furthermore, in their face model, only one kind of facial muscles (linear muscle) is modeled and adapted.



(a) (b) (c) (d) (e)

Fig. 1. (a) Scanned surface (122,752 points). (b) Texture-mapped geometry. (c) Structural components of the prototype model. (d) Face geometry. (e) Layered anatomical structure of the skin, muscles, and skull

3 Data Acquisition

The surface data of human faces is acquired using a Minolta VIVID-900 laser range scanner [13]. The result is a dense point cloud of range data (about 120k points), frequently with holes due to missing data (see Fig. 1 (a)). No further post-processing of hole-filling is performed on it. The acquired 640×480 reflectance (RGB) image is registered automatically against the range data and used for texture-mapping (see Fig. 1 (b)).

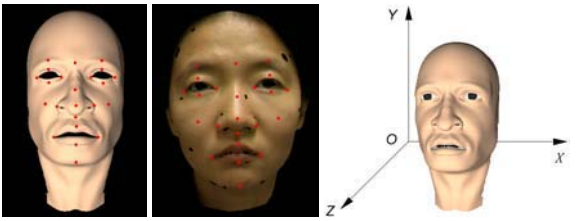
We have developed an anatomy-based facial model for use in our physically-based facial animation system [23]. This prototype model encapsulates four structural components: skin, muscles, skull, and eyes (see Fig. 1). The skin surface is represented as a triangular mesh, consisting of 2,848 vertices and 5,277 triangles. Basically, the edges and vertices of the skin mesh are converted to springs and point masses to simulate dynamic deformation of the soft tissue. A layer of muscles is attached to the skin mesh to control facial movement. The muscles can contract in a linear or circular fashion to drive the deformation of surrounding tissue. We have modeled 23 of the major muscles responsible for facial expressions. A synthesis of facial expressions is carried out by a deformation of the skin mesh resulting from the combined contraction of a set of muscles based on Action Units (AUs) of the Facial Action Coding System (FACS) [4]. The embedded skull is represented as a triangular mesh and is used only during initialization of the structure for constructing muscles at anatomically correct positions. The eyes are separately modeled using geometric models.

4 Skin Layer Deformation

Let $\mathcal{F}$ and $\mathcal{F}^*$ denote the generic prototype model and scanned data, respectively. By adapting $\mathcal{F}$ to $\mathcal{F}^*$, $\mathcal{F}$ takes on the shape and texture of a specific person and can be animated with predictability.

4.1 Global Adaptation

The global adaptation adapts the position, size, and orientation of $\mathcal{F}$ to align it with $\mathcal{F}^*$ in the 3D space. We have interactively specified a set of 23 anthropometric landmarks [5] on both $\mathcal{F}$ and $\mathcal{F}^*$ (see Fig. 2). Given two sets of N_p landmarks $\mathbf{p}_i$ and $\mathbf{p}_i^*$



(a) (b) (c)

Fig. 2. Anthropometric landmarks specified on the prototype model (a) and scanned data (b). (c) World coordinate system

($i = 1, \ldots, N_p$) specified on $\mathcal{F}$ and $\mathcal{F}^*$ respectively, we use a subset consisting of 6 landmarks (corners of the eyes and mouth) for the global adaptation. For $\mathcal{F}$, the eye center positions $\mathbf{p}^{le}$ and $\mathbf{p}^{re}$ are obtained as the midpoint of two corner positions of the left and right eye, respectively. The mouth center position $\mathbf{p}^{m}$ is computed as the midpoint of two corner positions of the mouth. The plane on which 3D center points of the eyes and mouth lie is called *eye-mouth plane*. The model center $\mathbf{p}^{c}$ is defined as the midpoint between $\mathbf{p}^{le}$ and $\mathbf{p}^{re}$. The corresponding key points of $\mathcal{F}^*$, $\mathbf{p}^{*le}$, $\mathbf{p}^{*re}$, $\mathbf{p}^{*m}$ and $\mathbf{p}^{*c}$, are calculated in the same way.

The global adaptation is carried out in two steps. In the first step, $\mathcal{F}^*$ is transformed such that the line through the estimated eye center positions is parallel to the x-axis of the world coordinate system and that the sagittal plane (vertical plane cutting through the center of the face) coincides with the y-z plane. In principle, six parameters must be estimated, three rotation angles (r_x^*, r_y^*, r_z^*) around the x-, y-, and z-axes, and three translation components along the x-, y-, and z-axes. Using the reference eye center positions ($\mathbf{p}^{*le}$ and $\mathbf{p}^{*re}$) and their orthographic projections on the x-z plane ($\mathbf{p}^{*le}_{|xz}$ and $\mathbf{p}^{*re}_{|xz}$), r_y^* is estimated as the angle between vector $\overrightarrow{\mathbf{p}^{*le}_{|xz}\mathbf{p}^{*re}_{|xz}}$ and the x-axis, and r_z^* is estimated as the angle between vectors $\overrightarrow{\mathbf{p}^{*le}\mathbf{p}^{*re}}$ and $\overrightarrow{\mathbf{p}^{*le}_{|xz}\mathbf{p}^{*re}_{|xz}}$. r_x^*, on the other hand, is obtained by calculating the angle between the normal of the eye-mouth plane and the z-axis. The translation components are estimated such that the y-z plane coincides with the sagittal plane.

In the second step, the generic model $\mathcal{F}$ is to be scaled for matching the size of $\mathcal{F}^*$ and to be rotated and translated for matching the position of $\mathcal{F}^*$. Three rotation angles (r_x, r_y, r_z) are estimated in the same way as to determine the rotation parameters of $\mathcal{F}^*$. The scaling factors (s_x, s_y, s_z) are estimated from the ratio of lengths between a pair of key points as measured both in $\mathcal{F}$ and $\mathcal{F}^*$:

$$s_x = \frac{\| \mathbf{p}^{*le} - \mathbf{p}^{*re} \|}{\| \mathbf{p}^{le} - \mathbf{p}^{re} \|}, \quad s_y = \frac{\| \mathbf{p}^{*c} - \mathbf{p}^{*m} \|}{\| \mathbf{p}^{c} - \mathbf{p}^{m} \|}, \quad s_z = \frac{1}{2}(s_x + s_y) \tag{1}$$

where $\| \cdot \|$ denotes the Euclidean norm. The translation parameters are estimated by matching the model center of $\mathcal{F}$ with that of $\mathcal{F}^*$. The same transformation is applied to the underlying muscle layer and geometric component of eyes to get their initial layout for further adaptation.

4.2 A Dynamic Deformable Model

After the global adaptation, we can perform skin mesh adaptation to fit $\mathcal{F}$ to $\mathcal{F}^*$. To accomplish the fitting, a physical model based on the mass-spring-damper (MSD) system is created from $\mathcal{F}$: each vertex in the surface of $\mathcal{F}$ is characterized by a mass m and each mass point is linked to its neighbors by damped massless springs of natural length greater than zero. The discrete Lagrangian equations of motion for the MSD mesh can be expressed in 3D vector form as:

$$\mathbf{M}\frac{d^2\mathbf{x}}{dt^2} + \mathbf{D}\frac{d\mathbf{x}}{dt} + \mathbf{F}_{elastic}(\mathbf{x}, \mathbf{K}) = \mathbf{F}_{data}(\mathbf{x}) + \mathbf{F}_{feature}(\mathbf{x}) \tag{2}$$

Given $N_{\mathcal{F}}$ vertices on $\mathcal{F}$, $\mathbf{x}$ represents a $3N_{\mathcal{F}}$ vector of nodal displacement, $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$ are $3N_{\mathcal{F}} \times 3N_{\mathcal{F}}$ matrices describing the mass, damping, and stiffness between vertices in the mesh, respectively. $\mathbf{F}_{elastic}$, $\mathbf{F}_{data}$, and $\mathbf{F}_{feature}$ are vectors of dimension $3N_{\mathcal{F}}$ and represent the elastic, data, and feature forces, respectively. Give an initial mesh surface $\mathbf{x}(t = 0)$, the new position of each skin vertex is obtained by calculating the energy equilibrium state of the entire system $d\mathbf{x}/dt = d^2\mathbf{x}/dt^2 \approx 0$.

4.3 Elastic Force

Suppose an arbitrary vertex $\mathbf{x}_i$ in $\mathcal{F}$ is connected to one of its neighbors $\mathbf{x}_j$ by a spring. The elastic force produced by the springs over $\mathbf{x}_i$ is

$$\mathbf{F}_{elastic}(\mathbf{x}_i) = -\sum_{j \in \Omega_i} k_{ij} \frac{(\|\mathbf{x}_i - \mathbf{x}_j\| - d_{ij})}{\|\mathbf{x}_i - \mathbf{x}_j\|}(\mathbf{x}_i - \mathbf{x}_j) \tag{3}$$

Ω_i Set regrouping all neighboring mass points that are linked by springs to $\mathbf{x}_i$
k_{ij} Stiffness of the spring
d_{ij} Natural length of the spring

4.4 Data Force

A criterion of a good fitting is that the generic mesh $\mathcal{F}$ should be as close as possible to the scanned data $\mathcal{F}^*$. To this end, we apply data forces $\mathbf{F}_{data}$ which are generated

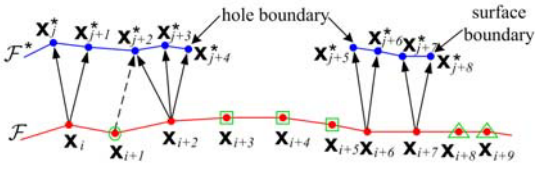


Fig. 3. Four types of vertices on the generic mesh. The *fully-influenced vertices* (red points) are affected by the data forces (solid black arrows) from the close data points. The *partially-influenced vertices* (red points with a circle) subject the data force (dashed black arrow) from the nearest data point. The *hole vertices* (red points with a square) are vertices corresponding to the areas of $\mathcal{F}$ that are not covered by the scan data (holes). The *outer vertices* (red points with a triangle) have no corresponding part of $\mathcal{F}$

from $\mathcal{F}^*$ and attract the vertices in $\mathcal{F}$ to minimize distance between them. Since the influence of $\mathbf{F}_{data}$ is local in order to reconstruct the details of the face shape, we take into account only the 'close' data points. For the data force generated from each scanned data point $\mathbf{x}_j^*$, the vertex of its influence on the mesh surface, $\mathbf{x}(\Pi\mathbf{x}_j^*)$, is the one of minimal distance to it:

$$\|\mathbf{x}_j^* - \mathbf{x}(\Pi\mathbf{x}_j^*)\| = \min_{i=1}^{N_{\mathcal{F}}} \|\mathbf{x}_j^* - \mathbf{x}_i\| \quad j = 1, \ldots, N_{\mathcal{F}^*} \tag{4}$$

where $N_{\mathcal{F}^*}$ is the number of scanned data points. The set of close data points of each mesh vertex $\mathbf{x}_i$, C_i^{data}, is obtained by searching for data points whose data force influence vertex in the generic mesh is $\mathbf{x}_i$, i.e. $C_i^{data} = \{j | \mathbf{x}(\Pi\mathbf{x}_j^*) = \mathbf{x}_i\}$.

With the obtained C_i^{data}, we divide the vertex set V of generic mesh into four subsets: V_f, V_p, V_h, and V_o, as shown in Fig. 3. *Fully-influenced vertex set* V_f represents the set of vertices that have nonempty C_i^{data}. A vertex $\mathbf{x}_i$ in V_f therefore receives data forces from data points whose point indexes are stored in C_i^{data}. Since scanned data is used to attract the generic mesh and deform it into a consistent shape, we use a type of springs with natural length of zero for this kind of deformation:

$$\mathbf{F}_{data}(\mathbf{x}_i) = \sum_{j \in C_i^{data}} \eta_j \frac{\|\mathbf{x}_i - \mathbf{x}_j^*\|}{H} \overrightarrow{\mathbf{x}_i\mathbf{x}_j^*} \quad \forall \mathbf{x}_i \in V_f \tag{5}$$

where η_j controls the strength of the forces, H is a normalizing constant, and $\overrightarrow{\mathbf{x}_i\mathbf{x}_j^*}$ denotes the force direction. Intuitively, H represents the range of the influence of the scanned data on the mesh surface. In our implementation, H is chosen as the maximum of distances between the data points and their influence vertices, i.e. $H = \max_{j=1}^{N_{\mathcal{F}^*}} \|\mathbf{x}_j^* - \mathbf{x}(\Pi\mathbf{x}_j^*)\|$.

For a mesh vertex $\mathbf{x}_i$ whose close data point set C_i^{data} is empty, we find the nearest data point on $\mathcal{F}^*$ to it. If the nearest data point $\mathbf{x}_{n_i}^*$ is not on the boundary edge of $\mathcal{F}^*$, the vertex $\mathbf{x}_i$ is categorized into the *partially-influenced vertex set* V_p. The data force applied to a partially influenced vertex $\mathbf{x}_i$ is generated from $\mathbf{x}_{n_i}^*$:

$$\mathbf{F}_{data}(\mathbf{x}_i) = \eta_{n_i} \frac{\|\mathbf{x}_i - \mathbf{x}_{n_i}^*\|}{H} \overrightarrow{\mathbf{x}_i\mathbf{x}_{n_i}^*} \quad \forall \mathbf{x}_i \in V_p \tag{6}$$

Hole vertex set V_h is a collection of vertices that have empty C_i^{data} and their nearest data point is on a hole boundary edge of $\mathcal{F}^*$. For such vertices the system automatically sets the data force to zero, so that they will only be affected by the internal force $\mathbf{F}_{elastic}$. As a result, holes in the scanned data will be filled in by seamlessly deformed parts of the generic mesh. We also set the data force scaling factor η_j as a function to ensure decrease of the data forces with the decrease of the distance from the hole boundary on $\mathcal{F}^*$. η_j is ramped linearly towards 0 within five-ring neighborhood of the hole boundary. Because the data forces taper gradually to zero near holes, we obtain a smooth blend between regions with good data and regions with no data.

There are number of mesh vertices that have empty C_i^{data} and their nearest data point is on the outermost surface boundary of $\mathcal{F}^*$. They are grouped into the *outer vertex set* V_o. Such vertices are 'out of' the boundary of $\mathcal{F}^*$ and have no appropriate corresponding part on $\mathcal{F}^*$ (e.g., the part of $\mathcal{F}$ representing the neck might be longer than the corresponding part of $\mathcal{F}^*$). To avoid incorrect correspondence, these vertices are not involved in the dynamic fitting and final surface rendering.

4.5 Feature Force

To avoid undesirable minima and to fit facial features, we use the set of feature points shown in Fig. 2 as constraints. For each feature point $\mathbf{p}_k^*$ ($k = 1, \ldots, N_p$) on the scanned data $\mathcal{F}^*$, it has a corresponding feature point $\mathbf{p}_k$ on the prototype model $\mathcal{F}$. $\mathbf{p}_k$ is located in a triangle element T_{n_k}, where n_k is the triangle index. We call T_{n_k} the *influence core* of the feature point $\mathbf{p}_k^*$. The force generated by $\mathbf{p}_k^*$ is applied to the influence core and its local neighborhood.

The two types of deformations, global and local deformations driven by the feature and data forces respectively, should be balanced. We change the influence of both types of deformations over time: initially, $\mathcal{F}$ is mostly influenced by the feature force, therefore moving toward its global shape; then the feature force decreases and the data force is allowed to dominate so that $\mathcal{F}$ is deformed to fit $\mathcal{F}^*$ closely. To achieve this shift during the evolution, the feature force from a feature point $\mathbf{p}_k^*$ applied to a skin mesh vertex $\mathbf{x}_i$ is defined as:

$$\mathbf{F}^k_{feature}(\mathbf{x}_i) = \begin{cases} \frac{\delta_k \|\mathbf{p}_k^* - \mathbf{x}_i\|}{L_k} \overrightarrow{\mathbf{x}_i \mathbf{p}_k^*} & \text{if} \quad q_{ik} \le Q^k_{ref}(t) \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

and total feature force applied on $\mathbf{x}_i$ is the sum of forces generated from all feature points of $\mathcal{F}^*$. In Eq. 7, δ_k is a scaling parameter, $\overrightarrow{\mathbf{x}_i \mathbf{p}_k^*}$ indicates the force direction, q_{ik} denotes the topological distance of the mesh vertex $\mathbf{x}_i$ from the influence core of kth feature point measured in terms of the smallest number of edges between them, $Q^k_{ref}(t)$ denotes the reference topological distance of kth feature point at time t, and L_k is a normalizing constant. Intuitively, $Q^k_{ref}(t)$ and L_k reflect the range of the influence of feature points on the generic mesh. They are defined as:

$$Q^k_{ref}(t) = q_k^0 \cos(\frac{t}{t_0} \cdot \frac{\pi}{2}), \quad L_k = \max_{\{i | q_{ik} \le q_k^0\}} \|\mathbf{p}_k^* - \mathbf{x}_i\| \tag{8}$$

where q_k^0 is the topological distance threshold that defines the region on $\mathcal{F}$ influenced by kth feature point when $\mathcal{F}$ is in its initial position (we use a threshold of 5 in our experiments), and t_0 is the time threshold that controls the influence of feature force. The smaller t_0, the smaller the deformation due to feature attractions. $Q^k_{ref}(t)$ decreases from q_k^0 when t=0 to zero when t=t_0 which means that mesh vertices attracted by the feature point get closer to the influence core with iterations. Therefore, the influence of feature force becomes smaller during mesh deformation.

We integrate Eq. 2 forward through time using the explicit second-order Verlet-Leapfrog method [19]. Since forces $\mathbf{F}_{elastic}$, $\mathbf{F}_{data}$, and $\mathbf{F}_{feature}$ are independent from each other, we parallel their computation on our 2 processor Intergraph Zx10 workstation to accelerate the simulation.

5 Muscle Layer Deformation

There are three types of muscles incorporated in our prototype model: the *linear muscles*, *sheet muscles*, and *sphincter muscles* [23], as shown in Fig. 4. The linear muscles that contract in a linear fashion are suspended between the skin and skull layers. The sheet muscles consist of a series of almost-parallel fibers spread over a rectangular area.

The central muscle fiber of these two kinds of muscles has one end attached to the skull which is called the *muscle attachment point* and the other end inserted to the skin which is called the *muscle insertion point*. Since the positions of the muscle attachment and insertion points completely define the location of a muscle, the adaptation of these two types of muscles is only a problem of how to determine new positions of their attachment and insertion points with the skin mesh adaptation.

For a muscle insertion point $\mathbf{m}^I$, it is located on the surface of an indexed triangle of $\mathcal{F}$ with its barycentric coordinates to be obtained. After $\mathcal{F}$ is adapted, the new position of $\mathbf{m}^I$, $\tilde{\mathbf{m}}^I$, is obtained as the linear interpolation of positions of the triangle vertices using the obtained barycentric coordinates as the interpolation coefficients. To determine new position of each muscle attachment point $\mathbf{m}^A$, we first compute its corresponding skin point $\mathbf{m}^{A(skin)}$ on the un-adapted $\mathcal{F}$ by casting a ray along the normal of $\mathbf{m}^A$. The place where the ray pierces the skin mesh surface is the point $\mathbf{m}^{A(skin)}$. Each muscle attachment point on the skull surface is related to its corresponding point on the skin by a *tissue depth vector* $\mathbf{d}$, giving an offset from $\mathbf{m}^A$ to $\mathbf{m}^{A(skin)}$ (see Fig. 5). We obtain the updated position of the skin point on the adapted skin mesh, $\tilde{\mathbf{m}}^{A(skin)}$, in the same way as to calculate $\tilde{\mathbf{m}}^I$. Then, the new position of muscle attachment point, $\tilde{\mathbf{m}}^A$, is obtained by offsetting $\tilde{\mathbf{m}}^{A(skin)}$ along the negated vector $\mathbf{d}$.

The sphincter muscles consist of fibers that loop around the eyes and mouth and can draw towards a virtual center; the examples are the *Orbicularis Oculi* around the eyes and *Orbicularis Oris* around the mouth. They are modelled as a parametric ellipse and are placed between the skin and skull layers (see Fig. 4). Each sphincter muscle is associated with a set of three location parameters $(\mathbf{o}, a, b)$ where $\mathbf{o}$ is the epicenter, a and

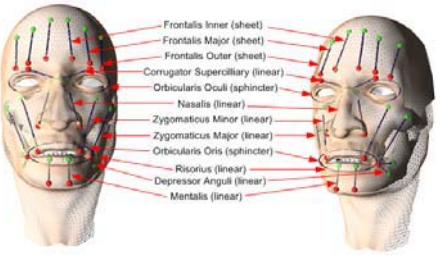


Fig. 4. Muscle structure of the prototype model. The attachment and insertion points of the linear and sheet muscles are represented by green and red dots, respectively

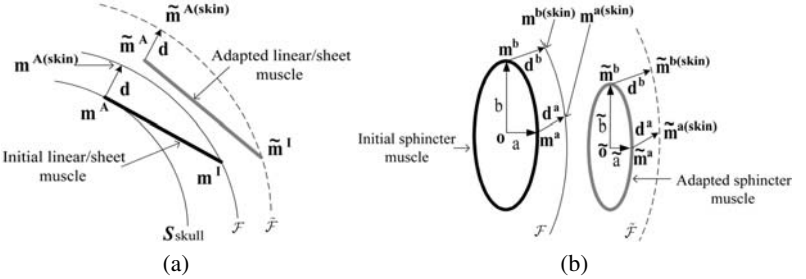


Fig. 5. Adaptation of the linear/sheet muscle (a) and sphincter muscle (b)

b are the half-lengths of its two object axes. To adapt a sphincter muscle, one end of its two axes, $\mathbf{m}^a$ and $\mathbf{m}^b$, are employed (see Fig. 5). The positions of their corresponding points on the skin, $\mathbf{m}^{a(skin)}$ and $\mathbf{m}^{b(skin)}$, are obtained in the same way as to adapt the linear/sheet muscle. The offset vector from sphincter muscle points to their corresponding skin points, $\mathbf{d}^a$ and $\mathbf{d}^b$, are computed. The new positions of the corresponding skin points after the skin mesh adaptation, $\tilde{\mathbf{m}}^{a(skin)}$ and $\tilde{\mathbf{m}}^{b(skin)}$, are obtained using transformed positions of the skin surface triangles that contain $\mathbf{m}^{a(skin)}$ and $\mathbf{m}^{b(skin)}$, respectively. New positions of the sphincter muscle points, $\tilde{\mathbf{m}}^a$ and $\tilde{\mathbf{m}}^b$, are obtained by offsetting $\tilde{\mathbf{m}}^{a(skin)}$ and $\tilde{\mathbf{m}}^{b(skin)}$ along the negated vectors $\mathbf{d}^a$ and $\mathbf{d}^b$, respectively. The new parameter set, $(\tilde{\mathbf{o}}, \tilde{a}, \tilde{b})$, is then determined as:

$$\tilde{\mathbf{o}}_{(x)} = \tilde{\mathbf{m}}^b_{(x)}, \quad \tilde{\mathbf{o}}_{(y)} = \tilde{\mathbf{m}}^a_{(y)}, \quad \tilde{\mathbf{o}}_{(z)} = \frac{1}{2}(\tilde{\mathbf{m}}^a_{(z)} + \tilde{\mathbf{m}}^b_{(z)}),$$
$$\tilde{a} = |\tilde{\mathbf{m}}^a_{(x)} - \tilde{\mathbf{o}}_{(x)}|, \quad \tilde{b} = |\tilde{\mathbf{m}}^b_{(y)} - \tilde{\mathbf{o}}_{(y)}|. \tag{9}$$

where subscripts (x), (y), and (z) denote three coordinates in Euclidean 3-space. With obtained new positions of all facial muscles, the muscle layer is adapted automatically.

6 Results

The presented method has been tested on the scanned data shown in Fig. 1. Fig. 6 (a) and (b) depict the adapted facial model with smooth shading and texture mapping, respectively. The animatable face structure of adapted muscle layer is shown in Fig. 6 (c). We have used our system to reconstruct facial models for various people. Fig. 7 shows two more reconstructed models of male and female individuals with different face shapes.

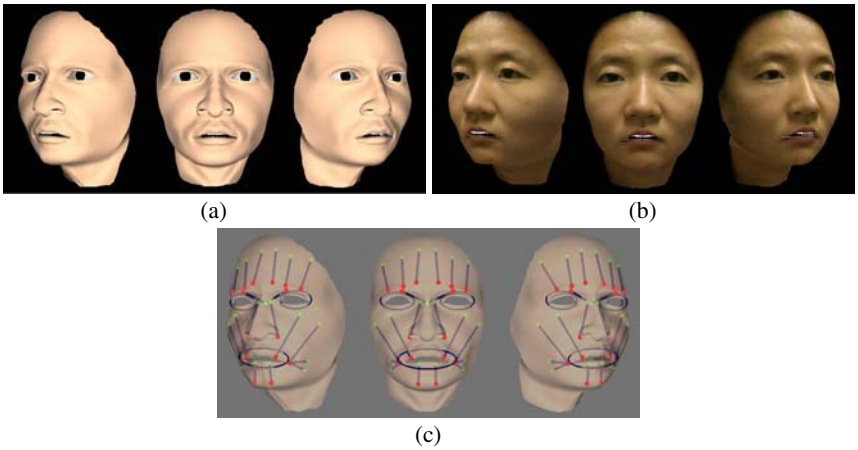


Fig. 6. Reconstructed face: (a) rendered with Gouraud shading, (b) texture mapping, (c) muscle structure

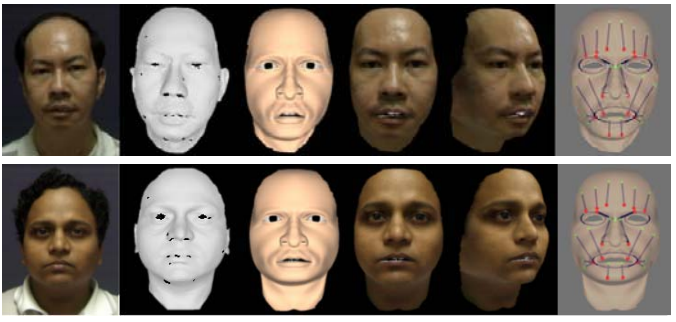


Fig. 7. Reconstructed facial models of various people. In each example, from left to right: original photograph; scanned data; reconstructed model with Gouraud shading and texture mapping; adapted muscle structure

The reconstructed facial model has a well-defined anatomical structure for physically-based animation. Animation can be controlled on different levels. At the lowest level, mesh vertices in the region influenced by an individual muscle are driven by a field of muscle force vectors to be displaced to their new positions. At the highest level, a synthesis of different facial expressions is carried out by a deformation of the skin mesh resulting from the combined contraction of a particular set of facial muscles based on the FACS [4]. Fig. 8 shows some facial expressions animated on the reconstructed models with the given muscle parameters.

In our method, the only initial requirement is that the anthropometric landmarks are specified. This process takes about 2-3 minutes. The automated multi-layer deformation is then executed. We arrive at about 6 minutes total run time for the skin layer

Fig. 8. Various expressions synthesized on the prototype model (first row) and reconstructed models (rest rows). The muscle parameters are set the same for each expression

deformation on a PIII 750MHz Intergraph Zx10 with a 120k point scan data-set. About two seconds are needed in the muscle layer deformation. This time is dominated by ray/mesh intersection tests and point/point distance computations, which are done in a brute force manner. We expect a speed-up from optimization of these intersection tests. For the eyes and teeth, they are automatically transformed with the skin mesh in the global alignment step. To fit them exactly into the adapted model, some fine-tuning is taken by using an interactive editing tool. Given the scan data, the whole reconstruction process including the tuning of eyes and teeth positions takes about 15 minutes.

7 Conclusion

We have presented a new method for reconstruction of an animatable, anatomy-based facial model of a specific person from the captured surface data. The contributions of our method are:

- An efficient individualized face reconstruction technique with minimal manual intervention.
- A physically-based skin layer deformation using a dynamic deformable model which has not been utilized by any previous face reconstruction method.
- Automatic deformation of the muscle structure which includes three kinds of muscle models.
- The ability to generate useful models from incomplete input data, with no requirement for further processing of hole filling.

For future work we would like to automate the landmark specification process by using the vision face recognition approach to detect facial features and find correspondence from image inputs. For the skin layer deformation, alternative surface models, such as a finite element model, can also be used so that the implementation of the system can be made more precisely. For representing high-resolution surface detail, we will use displacement map with triangular mesh subdivision to generate surface models at multiple levels-of-detail.

References

1. T. Akimoto, Y. Suenaga, and R. S. Wallace. "Automatic creation of 3D facial models." *IEEE Computer Graphics and Application,* 13(5):16-22, September 1993.
2. V. Blanz and T. Vetter. "A morphable model for the synthesis of 3D faces." *Proc. SIGGRAPH'99,* pp. 187-194, August 1999.
3. D. DeCarlo, D. Metaxas, and M. Stone. "An anthropometric face model using variational techniques." *Proc. SIGGRAPH'98,* pp. 67-74, July 1998.
4. P. Ekman and W. V. Friesen. *Facial Action Coding System.* Consulting Psychologists Press Inc., California, 1978.
5. L. G. Farkas. *Anthropometry of the Head and Face,* 2nd ed. Raven Press, 1994.
6. T. Goto, S. Kshirsagar, and N. Magnenat-Thalmann. "Automatic face cloning and animation." *IEEE Signal Processing Magazine,* 18(3): 17-25, May 2001.
7. Horace H. S. Ip, Lijun Yin. "Constructing a 3D individualized head model from two orthogonal views." *The Visual Computer,* 12:254-266, 1996.
8. I. T. Jollife. *Principal Component Analysis.* Springer Verlag, New York, 1986.

9. K. Kähler, J. Haber, H. Yamauchi, and H. P. Seidel. “Head shop: Generating animated head models with anatomical structure.” *Proc. ACM SIGGRAPH Symposium on Computer Animation,* pp. 55-64, 2002.
10. T. Kurihara and K. Arai. “A transformation method for modeling and animation of the human face from photographs.” *Proc. Computer Animation’91,* Springer-Verlag Tokyo, pp. 45-58, 1991.
11. Y. Lee, D. Terzopoulis, and K. Waters. “Realistic modeling for facial animation.” *Proc. SIGGRAPH’95,* pp. 55-62, August 1995.
12. W. S. Lee and N. Magnenat-Thalmann. “Fast head modeling for animation.” *Journal Image and Vision Computing,* 18(4): 355-364, March 2000.
13. Minolta VIVID-900, http://www.minolta-3d.com/products/
14. F. I. Parke and K. Waters. *Computer Facial Animation.* AK Peters, Wellesley, MA, 1996.
15. F. I. Parke. “Parameterized models for facial animation.” *IEEE Computer Graphics and Application,* 2(9): 61-68, November 1982.
16. F. I. Parke. *Computer generated animation of faces.* Master’s thesis, University of Utah, Salt Lake City, 1972.
17. A. Pentland, B. Moghaddam, and T. Starner. “View-based and modular eigenspaces for face recognition.” *Proc. CVPR’94,* pp. 84-91, 1994.
18. F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. H. Salesin. “Synthesizing realistic facial expressions from photographs.” *Proc. SIGGRAPH’98,* pp. 75-84, July 1998.
19. W. H. Press, B. P. Fannery, S. A. Teukolsky, and W. T. Verrerling. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, Cambridge, UK, 1986.
20. F. Ulgen. “A step toward universal facial animation via volume morphing.” *Proc. 6th IEEE Interantional Workshop on Robot and Human Communication*, pp. 358-363, 1997.
21. M. Turk and A. Pentland. “Eigenfaces for recognition” *Journal of Cognitive Neuroscience 3*, 1: 71-86, 1991.
22. K. Waters and D. Terzopoulus. “Modeling and animating faces using scanned data.” *Journal of Visualization and Computer Animation*, vol.2, pp. 123-128, 1991.
23. Y. Zhang, E. C. Prakash and E. Sung. “A new physical model with multi-layer architecture for facial expression animation using dynamic adaptive mesh.” *IEEE Transactions on Visualization and Computer Graphics*, 10(3):339-352, May 2004.

Quasi-interpolants Based Multilevel B-Spline Surface Reconstruction from Scattered Data*

Byung-Gook Lee[1], Joon-Jae Lee[1], and Ki-Ryoung Kwon[2]

[1] Division of Internet Engineering, Dongseo University, Busan 617-716, Korea
lbg, jjlee@dongseo.ac.kr
[2] Division of Computer and Electronic Engineering,
Pusan University of Foreign Studies, Busan 608-738, Korea
krkwon@pufs.ac.kr

Abstract. This paper presents a new fast and local method of 3D surface reconstruction for scattered data. The algorithm makes use of quasi-interpolants to compute the control points from a coarse to fine hierarchy to generate a sequence of bicubic B-spline functions whose sum approaches to the desired interpolation function. Quasi-interpolants gives a procedure for deriving local spline approximation methods where a B-spline coefficient only depends on data points taken from the neighborhood of the support corresponding B-spline. Experimental results demonstrate that high-fidelity reconstruction is possible from a selected set of irregular samples.

1 Introduction

The problem of recovering a surface from scattered data is one of those interesting problems that is simple in concept but tricky when get into the detail. As we know, the real world is made up of continuous surfaces, not discrete points. So, we want to create a continuous surface from the unorganized data points. The ultimate goal of this paper is a surface reconstruction method as getting a smooth and high fidelity of 3D surface from scattered data points. In particular, the description should be sufficiently completed to reconstruct the 3D surface within a certain tolerance error, given their relative locations and expected noise.

There exist many techniques for surface approximation to improve the approximate continuity and smoothness in handling scattered data[1,6,7,8]. Tensor product of B-splines surfaces is widely used to approximate rather than to work with other types of approximation because of the advantages inherent in working with tensor products. Tensor product guarantee internal continuity if the knot vectors are set properly.

This paper is based on the multilevel B-splines approximation techniques presented by the publication of Lee, Wolberg and Shin[11]. In the previous works,

* This work was supported by grant No.R01-2004-000-10851-0, R05-2004-000-10968-0 from Ministry of Science & Technology.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1209–1218, 2005.

Forsey and Bartels[5] developed a surface fitting method which is adaptive on hierarchical spline functions. However, this method cannot deal with scattered data. Lee presented a multilevel B-spline algorithm to fit a uniform bicubic B-spline surface to scatterd data where multilevel or hierarchy is used to reduce the approximation errors. Although the previous methods are processed locally, they can not only be computationally expensive when they applied to the large number of points sets, but also not guarantee a reasonable global approximation at initial level.

The splines approximation technique used in this paper is quasi-interpolants, first developed by de Boor and Fix[3]. The quasi-interpolants operators were later generalized by Lyche and Schumaker[9], and it is their version that used in the alternative surface approximation technique. A quasi-interpolants operator approximates a curve by calculating coefficients that are used to weight samplings of the curve to be approximated. The Lyche and Schumaker quasi-interpolants operator uses coefficients that are inexpensive to calculate and samplings that are relatively expensive to calculate. It turns out to produce splines approximation with the required accuracy.

We introduce a new algorithm using quasi-interpolants to implement the multilevel B-spline approximation and apply to scattered data. The proposed method converges in a few iterates while maintaining the accuracy. This algorithm achieved C^2-continuous interpolation function from arbitrary scattered data with numerically stable. The algorithm is described in section 2. Section 3 gives the explanation on how to reconstruct the quasi-interpolants. Then, section 4 shows the experimental results for numerical examples and finally, conclusions are given in section 5.

2 Multilevel B-Spline Approximation

The methods explored in this paper take a set of scattered data as input and produce tensor product B-spline surfaces as output. The algorithms run in a multiresolutional setting over uniform partitions such that the final surface f is composed of a sequence of surfaces at dyadic scales,

$$f = f_0 + f_1 + \ldots + f_k,$$

where $f_i \in S_i, i = 0, 1, \ldots, k$, and $S_0, S_1, \ldots, S_k$ is a nested sequence of subspaces of S_k,

$$S_0 \subset S_1 \subset \ldots \subset S_k.$$

The basic algorithms used for the results presented in this paper were published in 1997 by Lee, Wolberg and Shin. They called the schemes *Multilevel B-splines*. Our interest is mainly scattered data interpolation and approximation, which is also the main focus in [11].

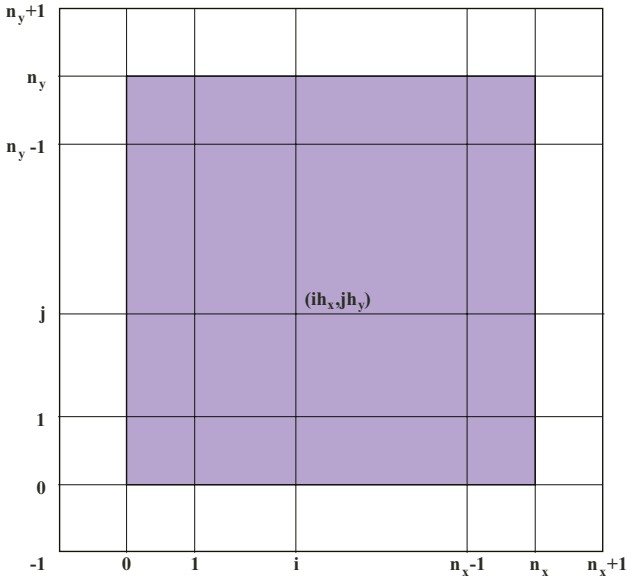


Fig. 1. The configuration of control lattice Ω

2.1 The Basic Schemes of B-Spline Approximation

Given a set of scattered points $P = \{P_i\}_{i=1}^{n}$, $P_i = (x_i, y_i, z_i) \in R^3$ and let $\Omega = \{(x, y)|0 \leq x < m_x, 0 \leq y < m_y\}$ be a rectangular domain in the xy-plane such that (x_i, y_i) is a point in Ω. Let Φ be a control lattice overlaid on a domain Ω. The control lattice Φ is an uniform tensor product grids over Ω.

To approximate scattered data points P, we formulate initial approximation function f as a uniform bicubic B-spline function, which is defined by a control lattice Φ. Let the initial number of control points on the lattice as $n_x = m_x/h_x$ in x-axis, and $n_y = m_y/h_y$ in y-axis. The knot intervals are uniform interval defined as h_x in x-axis and h_y in y-axis. So, for uniform cubic B-spline case, degree $d = 3$ and the set of knot vectors are defined as below:

$$\tau_x = \{-dh_x, \ldots, 0, h_x, \ldots, n_x h_x, \ldots, (n_x + d)h_x\}$$
$$\tau_y = \{-dh_y, \ldots, 0, h_y, \ldots, n_y h_y, \ldots, (n_y + d)h_y\}.$$

Let c_{ij} be the value of the ij-th control point on lattice Φ, located at position (ih_x, jh_y) of the grid defined by Φ, for $i = -1, 0, 1, \ldots, n_x + 1$ and $j = -1, 0, 1, \ldots, n_y + 1$. The approximation function f is defined in terms of these control points at position $(x, y) \in \Omega$ is given as

$$f(x, y) = \sum_{i=1}^{n_x+1} \sum_{j=1}^{n_y+1} c_{ij} B_{i,d}(x) B_{j,d}(y) \tag{1}$$

where $B_{i,d}$ and $B_{j,d}$ are uniform cubic B-spline basis functions, $d = 3$ and knot vector for cubic B-spline basis are below:

$$\{(i-2)h_x, (i-1)h_x, ih_x, (i+1)h_x, (i+2)h_x\}$$
$$\{(j-2)h_y, (j-1)h_y, jh_y, (j+1)h_y, (j+2)h_y\}.$$

2.2 Multilevel B-Spline Approximation

B-spline approximation(BA) algorithm generates a tradeoff exiting between the shape smoothness and accuracy of the approximation function. To overcome this tradeoff, multilevel B-splines approximation(MBA) algorithm is introduced [11]. The algorithm makes use of a hierarchy of control lattices to generate a sequence of function f_i and the final approximation function f is defined as the sum of functions f_i,

$$f = \sum_{i=1}^{k} f_i. \tag{2}$$

To optimize this process, B-spline refinement is used to reduce the sum of these functions into one equivalent B-spline function. The MBA algorithm serves result as smooth initial approximation f_0 to $\Delta^0 P = P$ defined on the coarsest control lattice $\Phi_0 = \Phi$, by applying the BA algorithm. To continue to the finer levels, below explanation are quoted from [11]: The first approximation possibly leaves large discrepancies at the data points in P. In particular, f_0 leaves a deviation

$$\Delta^1 z_i = z_i - f_0(x_i, y_i) \text{ for } i = 0, \ldots, n. \tag{3}$$

The next finer control lattice Φ_1 is then used to obtain function f_1 that approximates the difference $\Delta^1 P = \{(x_i, y_i, \Delta^1 z_i)\}$.

Then, the sum of $f_0 + f_1$ yields a smaller deviation (3) for each (x_i, y_i) in Ω.

$$\Delta^2 z_i = z_i - f_0(x_i, y_i) - f_1(x_i, y_i) \text{ for } i = 0, \ldots, n.$$

In general, for each level k in the hierarchy, the point set $\Delta^k P = \{(x_i, y_i, \Delta^k z_i)\}$ is approximated by a function f_k defined over the control lattices Φ_k, where

$$\Delta^k z_i = z_i - \sum_{l=0}^{k-1} f_l(x_i, y_i) = \Delta^{k-1} z_i - f_{k-1}(x_i, y_i)$$

and $\Delta^0 z_i = z_i$. This process starts from the coarsest lattice Φ_0 and continue incrementally to the finest lattice Φ_k with the set of knot vectors are defined as below:

$$\tau_x^k = \{-d\frac{h_x}{2^k}, \ldots, 0, \frac{h_x}{2^k}, \ldots, 2^k n_x \frac{h_x}{2^k}, \ldots, (2^k n_x + d)\frac{h_x}{2^k}\}$$
$$\tau_y^k = \{-d\frac{h_y}{2^k}, \ldots, 0, \frac{h_y}{2^k}, \ldots, 2^k n_y \frac{h_y}{2^k}, \ldots, (2^k n_y + d)\frac{h_y}{2^k}\}.$$

The final approximation function f is defined as the sum of the functions (2). They are many methods for refining a control lattice into another so that they generate the same B-spline functions. In this paper, B-spline refinement of an $(n_x + 3) \times (n_y + 3)$ control lattice $\Phi_0 = \Phi$ is always refined to a $(2^k n_x + 3) \times (2^k n_y + 3)$ control lattice Φ_k whose the control point spacing is half.

3 Quasi-interpolants

Many applications of splines make use of some approximation method to produce a spline function from given discrete data. Popular methods include interpolation and least squares approximation. However, both of these methods require solution of a linear system of equations with as many unknowns as the dimension of the spline space, and are therefore not suitable for real-time processing of large streams of data. For this purpose local methods, which determine spline coefficients by using only local information, are more suitable. To ensure good approximation properties it is important that the methods reproduce polynomials and preferably the functions in the given spline space. A method based on derivative information was constructed in [3], while a more general class was studied in [9]. In order to reproduce the spline space, the local information of the methods in [9] was restricted to lie in one knot interval. In this paper we remove this restriction. We then discuss some specific approximation methods for quadratic and cubic splines. We use B-splines as a basis for splines and denote the i^{th} B-spline of degree d with knots τ by $B_{i,d} = B_{i,d,\tau}$, and the linear space spanned by these B-splines by $S_{d,\tau}$.

Given a function f, the basic problem of spline approximation is to determine B-spline coefficients $(c_i)_{i=1}^n$ such that

$$Pf = \sum_{i=1}^{n} c_i B_{i,d}$$

is a reasonable approximation to f. The basic challenge is therefore to devise a procedure for determining the B-spline coefficients. We assume that f is defined on an interval $[a, b]$, and that we have selected a space of splines $S_{d,\tau}$ defined on $[a, b]$ (i.e., so that $\tau = (t_j)_{j=1}^{n+d+1}$ is nondecreasing with $t_{d+1} = a$ and $t_{n+1} = b$).

When determining c_k, this procedure gives us the freedom to restrict our attention to a local subinterval $I = [t_\mu, t_\nu]$ of our choice. By doing this we may reduce the complexity of the problem. Secondly, we have the freedom to choose the local approximation method P_I. Typical choices will be interpolation, least squares approximation, or a smoothing spline.

A general class of approximation methods are obtained by letting P_I be given as point functionals of the form

$$\lambda_{k,j} f = f(x_{k,j}) \text{ for } j = 1, \ldots, m_k,$$

where $m_k = \nu - \mu + d$ and $x_{k,1}, \ldots, x_{k,m_k}$ are given points. With this choice, it is well known (see page 200 of [1]) that if

$$B_{\mu-d-1+j,d}(x_{k,j}) > 0 \text{ for } j = 1, \ldots, m_k,$$

then we obtain c_k in the form

$$c_k = \lambda_k f = \sum_{j=1}^{m_k} w_{k,j} f(x_{k,j}), \tag{4}$$

for some vector $w_k = (w_{k,j})$. Equivalently, we can find w_k by solving the linear system

$$\delta_{i,k} = \lambda_k(B_{i,d}) = \sum_{j=1}^{m_k} w_{k,j} B_{i,d}(x_{k,j}), \tag{5}$$

for $i = \mu - d, \ldots, \nu - 1$ where $\delta_{i,k} = 1$ if $i = k$ and $\delta_{i,k} = 0$ otherwise, as usual. In practice one would usually determine c_k numerically, either from (4), or (5), except in special cases where the formulas are particularly simple. We consider some examples in the case where the knots and the degree of the spline are given.

Example 1. In the cubic spline case ($d = 3$). To determine coefficient c_k, we choose the interval $I = [t_k, t_{k+4}]$ which means that the local spline space has dimension 7,

$$S_{d,\tau,I} = \text{span}\{B_{k-3,d}, B_{k-2,d}, \ldots, B_{k+3,d}\}.$$

Here, the data points $\{P_{k,i}\}_{i=1}^{m_k}, P_{k,i} = (x_{k,i}, y_{k,i}) \in R^2$ are restricted to lie in the interval $I = [t_k, t_{k+4}]$.

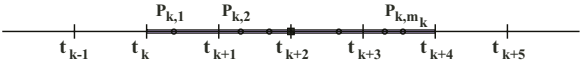


Fig. 2. A cubic spline, $I = [t_\mu, t_\nu] = [t_k, t_{k+4}]$

Coefficient matrix :

$$\begin{bmatrix} B_{k-3,3}(x_{k,1}) & B_{k-3,3}(x_{k,2}) & \ldots & B_{k-3,3}(x_{k,m_k}) \\ B_{k-2,3}(x_{k,1}) & B_{k-2,3}(x_{k,2}) & \ldots & B_{k-2,3}(x_{k,m_k}) \\ \vdots & \vdots & \ddots & \vdots \\ B_{k+3,3}(x_{k,1}) & B_{k+3,3}(x_{k,2}) & \ldots & B_{k+3,3}(x_{k,m_k}) \end{bmatrix}$$

The tensor product of the two spline spaces is defined as a family of all functions of the form

$$(Pf)(x,y) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} c_{ij} B_{i,d}(x) B_{j,d}(y)$$

where $B_{i,d}$ and $B_{j,d}$ are the B-splines on $\tau_x = (t_j)_{j=1}^{n_x+d+1}$ and $\tau_y = (s_j)_{j=1}^{n_y+d+1}$ respectively.

Example 2. In the tensor product cubic spline case ($d = 3$). To determine coefficient c_{ij}, we choose the interval $I = [t_\mu, t_\nu] \times [s_\mu, s_\nu] = [t_i, t_{i+4}] \times [s_j, s_{j+4}]$ which means that the local spline space has dimension 49,

$$S_{d,\tau,I} = \text{span}\{B_{i-3,d}B_{j-3,d}, \ldots, B_{i+3,d}B_{j+3,d}\}.$$

Here, the data points $\{P_{ij,k}\}_{k=1}^{m_{ij}}, P_{ij,k} = (x_{ij,k}, y_{ij,k}, z_{ij,k}) \in R^3$ are restricted to lie in the interval $I = [t_i, t_{i+4}] \times [s_j, s_{j+4}]$

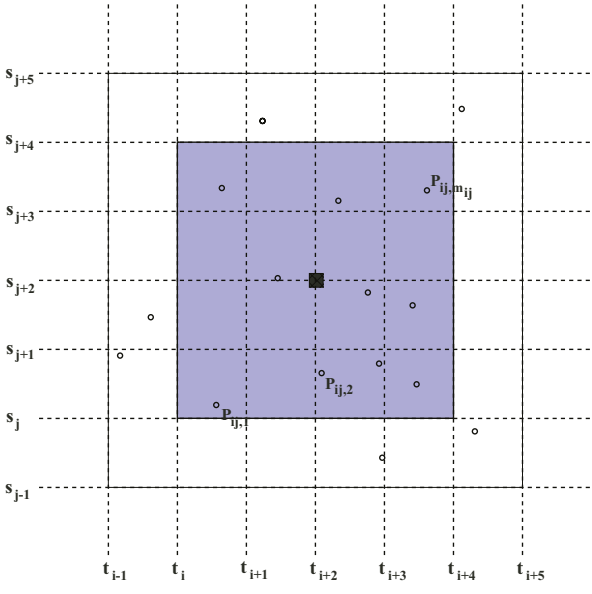


Fig. 3. A tensor product cubic spline, $I = [t_\mu, t_\nu] \times [s_\mu, s_\nu] = [t_i, t_{i+4}] \times [s_j, s_{j+4}]$

Coefficient matrix :

$$\begin{bmatrix} B_{i-3,3}(x_{ij,1})B_{j-3,3}(y_{ij,1}) & \ldots & B_{i-3,3}(x_{ij,m_{ij}})B_{j-3,3}(y_{ij,m_{ij}}) \\ B_{i-3,3}(x_{ij,1})B_{j-2,3}(y_{ij,1}) & \ldots & B_{i-3,3}(x_{ij,m_{ij}})B_{j-2,3}(y_{ij,m_{ij}}) \\ \vdots & \ddots & \vdots \\ B_{i+3,3}(x_{ij,1})B_{j+3,3}(y_{ij,1}) & \ldots & B_{i+3,3}(x_{ij,m_{ij}})B_{j+3,3}(y_{ij,m_{ij}}) \end{bmatrix}$$

4 Experimental Results

To demonstrate the accuracy of reconstruction by the proposed algorithm, we performed experiments with the same test functions used in [11]. Given a test function $g(x, y)$, we first sampled data points from it and applied the algorithm

to obtain an approximation function f. The difference between g and f is then measured by computing the normalized RMS(root mean square) error which is divided the RMS error by the difference of maximum and minimum values of g between the function values on a dense grid. That is,

$$RMS = \sqrt{\frac{\sum_{i=0}^{M}\sum_{j=0}^{N}(g(x_i,y_j)-f(x_i,y_j))^2}{(M+1)(N+1)}}$$

where $x_i = i/M$, $y_j = j/N$, and $M = N = 50$.

The test functions are

$$\begin{aligned}
g_1(x,y) &= 0.75\exp\left[-\frac{(9x-2)^2+(9y-2)^2}{4}\right] \\
&\quad + 0.75\exp\left[-\frac{(9x+1)^2}{49}-\frac{9y+1}{10}\right] \\
&\quad + 0.5\exp\left[-\frac{(9x-7)^2+(9y-3)^2}{4}\right] \\
&\quad - 0.2\exp\left[-(9x-4)^2-(9y-1)^2\right] \\
g_2(x,y) &= (\tanh(9-9x-9y)+1)/9 \\
g_3(x,y) &= (1.25+\cos(5.4y))/(6+6(3x-1)^2) \\
g_4(x,y) &= \exp\left[-\frac{81}{4}((x-0.5)^2+(y-0.5)^2)\right]/3 \\
g_5(x,y) &= \sqrt{64-81((x-0.5)^2+(y-0.5)^2)}/9-0.5
\end{aligned}$$

where the domain is $\{(x,y)|0 \le x \le 1, 0 \le y \le 1\}$

For each test function, we used three data sets, M100, M500 and R500, where M100 and M500 are small and large data sets, which consist of 100 and 500

Table 1. Normalized RMS errors between test functions and their approximations

M100	g_1	g_2	g_3	g_4	g_5
1	.01451	.03020	.00693	.00215	.00025
2	.01174	.03345	.00515	.00830	.00024
3	.00902	.02800	.00447	.00540	.00022
M500	g_1	g_2	g_3	g_4	g_5
1	.01176	.02747	.00567	.00163	.00016
2	.00416	.00776	.00046	.00148	.00002
3	.00047	.00115	.00017	.00024	.00002
R500	g_1	g_2	g_3	g_4	g_5
1	.01328	.03098	.00627	.00189	.00021
2	.00689	.01367	.00079	.00166	.00006
3	.00244	.00801	.00034	.00053	.00005

points, respectively. We uniformly sampled 7×7 and 15×15 data points, respectively, while the others was randomly sampled. And R500 points were totally randomly sampled.

Fig. 4 shows 5 tested functions used in the experiment. Fig. 5 shows approximation surface of g_1 and error surface at initial level, respectively, where the circle represent the sampled data. The second level approximation result is obtained from sum of A) and B). Table 1 demonstrates that the proposed method reconstructs test functions very accurately regardless of type of the function within a few level. We started from the number of control points of 7x7 at initial level to three levels. Particularly, it generates good approximation corresponding to reasonable global approximation at initial level for smooth functional surface.

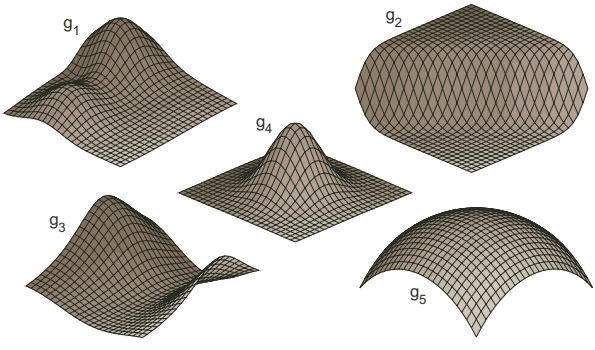


Fig. 4. Test functions

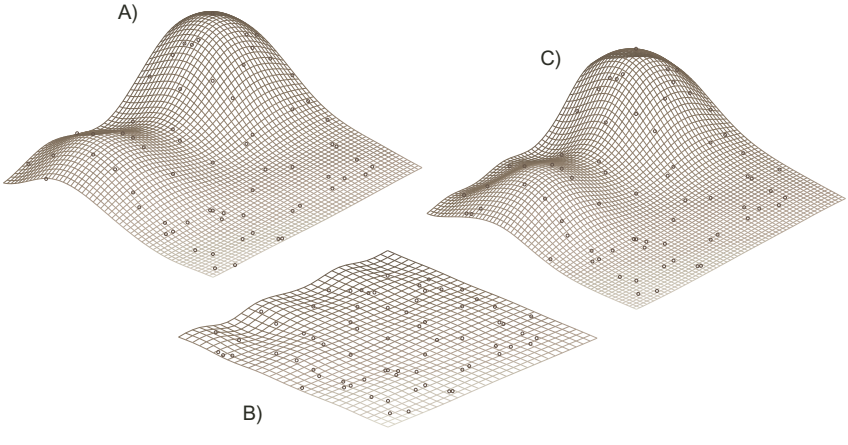


Fig. 5. A) Initial approximation surface from g_1 B) Error surface with $\Delta^1 z = z_i - f_0(x_i, y_i)$ C) Multilevel B-spline approximation with level 2, $f_0 + f_1$

5 Conclusion

This paper focuses on multilevel B-spline approximation based on quasi-interpolants for scattered data approximation and interpolation. The algorithm is fast and generates a C^2-continuous surface through a set of unevenly spaced points. Experimental results reveal that smooth 3D object reconstruction is possible from scattered data and irregular samples. Multilevel B-spline approximation was presented to circumvent the tradeoff which exists between the shape smoothness and approximation accuracy of the function, depending on the control lattice density. It is effectively gain in large performance. The quasi-interpolants is a special case of more general constructions and performs better approximation to reduces error results.

References

1. J. Barhak, A. Fischer, Parameterization and reconstruction from 3D scattered points based on neural network and PDE techniques, IEEE Trans. Visualization Comput. Graph. **7**(1) (2001) 1-16
2. de Boor, C., A Practical Guide to Splines, Springer Verlag, New York 1978
3. de Boor, C. and G.J.Fix, Spline approximation by quisiinterpolants, J. Approx. Theory **8** (1973) 19-45
4. Byung-Gook Lee, Lyche, T. and Knut Mørken, Some Examples of Quasi-Interpolants Constructed from Local Spline Projectors, Mathematical Methods for Curves and Surfaces, Oslo 2000, Lyche, T. and Schumaker L.L (eds), Vanderblit Press, Nashville (2001) 243-252
5. David Forsey and Richard Bartels. Surface fitting with hierarchical splines, ACM Trans. on Graphics **14** (1995) 134-161
6. P. Gu, X. Yan, Neural network approach to the reconstruction of free-form surfaces for reverse engineering, Comput. Aided Des. **27**(1) (1995) 59-64
7. M. Hoffmann, L. Varady, Free-form surfaces for scattered data by neural networks, J. Geometry Graph. **2** (1998) 1-6
8. A. Iglesias, G. Echevarria, A. Galvez, Functional networks for B-spline surface reconstruction, Future Generation Computer Systems **20** (2004) 1337-1353
9. Lyche, T. and L.L.Schumaker, Local spline approximation methods, J. Approx. Theory **25** (1979) 266-279
10. Øyvind Hjelle, Approximation of Scattered Data with Multilevel B-splines, Technical Report STF42 A01011, SINTEF 2001
11. S. Y. Lee, G. Wolberg, and S. Y. Shin. Scattered Data Interpolation with Multilevel B-Splines. IEEE Transactions on Visualization and Computer Graphics **3**(3) (1997) 229-244

Efficient Mapping Rule of IDEF for UMM Application

Kitae Shin[1], Chankwon Park[2], Hyoung-Gon Lee[3,†], and Jinwoo Park[3]

[1] Department of Industrial and Systems Engineering, Daejin University,
San 11-1, Sundan-dong, Poch'on-si, Kyonggi-do, Korea
ktshin@daejin.ac.kr
[2] Department of e-business, Hanyang Cyber University,
17 Haengdang-dong, Seongdong-gu, Seoul, Korea
chankwon@hycu.ac.kr
[3] Department of Industrial Engineering, Seoul National University,
San 56-1, Shinlim-dong, Kwanak-gu, Seoul, Korea
hklee@cybernet.snu.ac.kr, autofact@snu.ac.kr

Abstract. Various methodologies for business process analysis and design have been developed by many organizations. However, these methodologies are all uniquely developed to fit each organizations in their own ways and thuds are not compatible with each other. This poor compatibility between these methodologies incurs unnecessary analysis cost. In order to reduce this analysis cost, methods to map between the different methodologies need to be developed. UMM (UN/CEFACT Modeling Methodology) that has an object-oriented point of view can resolve the limits of the existing bottom-up approaches and make it more reasonable. It also simplifies the business and administrative procedures. IDEF (Integrated Definition Language) that has a structural point of view and is widely used as a system analysis and design method, needs to be mapped to UMM so that the existing IDEF models can be re-used. In this study, we present a guideline for procedures utilizing IDEF models from which the UMM models can be derived to develop an electronic commerce system that includes electronic documents exchange. By comparing IDEF and UMM, we analyze the differences between those two methodologies. Based on these differences, we propose basic strategies for mapping from IDEF to UMM. We also propose a mapping guideline which can be used to make suppositions of UMM results based on the modeling results of IDEF. The existing IDEF analysis/design results can be used to adopt UMM methodology for electronic business system. Therefore, with this, analysts who are familiar with the IDEF methodology can develop UMM work-flow by utilizing their existing results and skills.

1 Introduction

Even though there have been extensive advancements in computer hardware and certain software technologies in past several years, there is still ceaseless efforts put into researches related to huge information systems intended to find an effective and comprehensible methodologies that can be utilized throughout the whole system life cycle.

† Corresponding author. Tel.:+82-2-880-7180; fax.+82-2-873-7146

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1219 – 1228, 2005.

A methodology ensures that users have access to coherent and reliable information. Basically, its primary purpose is to improve communication between shareholders of the company concerned to the system. This is a well-made processes itself, which strives to capture the best practice and know-how, and it holds a symbolical mark that can communicate these procedure efficiently. IDEF (Integration DEFinition) and UMM (UN/CEFACT Modeling Methodology) are currently issues among the various existing methodologies related to electronic commerce that are being discussed today. In this study, we wish to present a standard for analyzing both IDEF & UMM methodologies and to provide a guideline so as to produce an output described by UMM workflow based on the IDEF output. By comparing IDEF and UMM, we analyzed the differences between the two methodologies. Based on these differences, we will suggest the basic strategies for mapping from IDEF to UMM.

2 Information System Modeling Methodologies

The main purpose for conducting a business process analysis on electronic commerce system is to subdivide high level processes used in electronic transactions. By analyzing business information and documents, information related to business transactions can be prescribed and business process specification and business document definition type outputs can be obtained. A formalized system of these procedures is what we call methodology.

A methodology can be defined as a set of compulsory procedure to guarantee the coherent production and delivery of outputs, and to prescribe activities to be executed, the role of participants, methods to be used and outputs to be achieved.[1]

2.1 IDEF Methodology

IDEF is a system analysis and design methodology which extracts an existing model of an enterprise or organization (AS-IS model) and develops a better and improved enterprise model which can extract problems through a hierarchical analysis process (TO-BE model). This methodology was developed in the mid 1970s with the support of the US Department of Defense, and has been recently designated a standard methodology for CALS (Commerce At Light Speed) activity derived in USA.

Initially, the IDEF suite was composed of an activity (function) modeling method, called IDEF0, a conceptual modeling method called IDEF1 and a simulation model specification method - the IDEF2. But new additions were made to a IDEF1, giving birth to IDEF1x. In addition, the process of developing a process modeling method known as IDEF3 commenced in the 1980s. IDEF3 has now taken over a major portion of the functions of IDEF2 as it can specify preliminary simulation models and has an object-state component that can be used to model how objects undergo change in a process.[3] These IDEF methodologies are not an interconnected sequences but independent methodologies themselves applied to achieve difference needs. [2], [5], [6]

In this study, we aim to investigate the functional model, information model and process modeling methodologies which has been applied for structural aspect of paradigm for converting. IDEF4 which basically belongs to the object-oriented model

methodology will not be considered here, because it shares many similarities with the UML object-oriented mechanism. The methodologies investigated in this research shall be IDEF0, IDEF1X and IDEF3.

2.2 UMM Methodology

The UMM is a formal description technique for describing any Open-edi scenario as defined in ISO/IEC 14662, Open-edi reference model, and is depicted in Fig. 1. An Open-edi scenario is a formal means to specify a class of business transactions having the same business goal, such as, purchasing or inventory management. The primary scope of UMM is the Business Operations View (BOV) and not the Functional Service View (FSV) as defined in ISO/IEC IS 14662. The BOV is defined as "a perspective of business transactions limited to those aspects regarding the making of business decisions and commitments among organizations", while the FSV is focused on implementation specific, technological aspects of Open-edi. As such, UMM provides a procedure for specifying (modeling), in a technology-neutral, implementation-independent manner business processes involving information exchange.

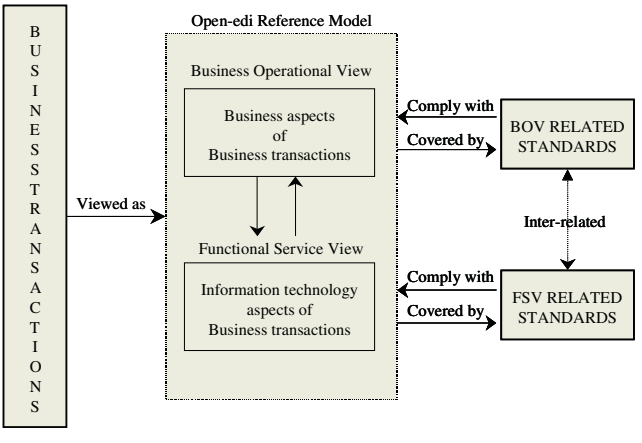


Fig. 1. open-edi reference model

The Unified Process (UP) methodology used in UMM recognizes that all software development projects pass through a series of general phases over the course of time. The phases are: [7]

· Inception
· Elaboration
· Construction
· Transition

The historic view of software engineering projects is a series of sequential steps, or workflows, moving from technology independent business process modeling to technology dependent deployment. The Unified Process methodology used by UMM is tied to these historic project steps as illustrated in Fig. 2.

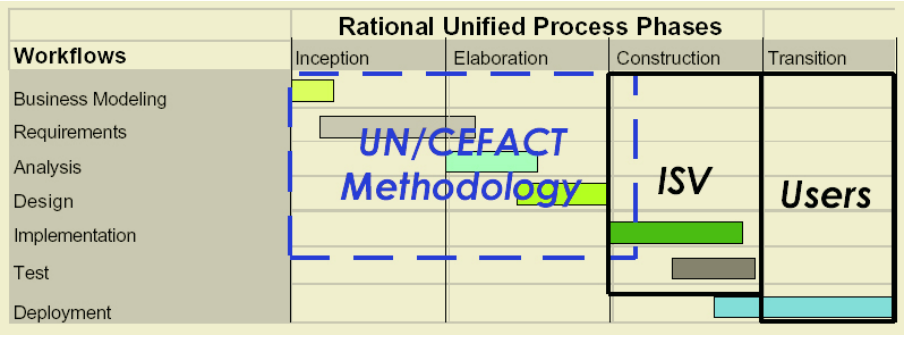


Fig. 2. Phases and Workflows

The focus of the UMM developed by UN/CEFACT is predominately the technology neutral intersection of the UP phases of Inception and Elaboration and the Software Engineering project workflows of Business Modeling, Requirements, Analysis and Design. This intersection coincides well with the UN/CEFACT TMWG (Techniques and Methodologies Work Group) charter for defining a methodology to support the BOV of ISO/IEC IS 14662 standard for the technology neutral definition of Open-edi scenarios.

3 IDEF Mapping on UMM

3.1 Comparison Between UMM Workflow and IDEF Methodology

In the unified process methodology which is applied in UMM, it assumes that every software development projects have a succession of general steps. Among these steps, UMM is focused on 'Inception' and 'Elaboration'. It utilizes workflow to understand business requirements and based on which it generates business scenario, business object, and business collaboration. The workflow consists of business modeling, requirements, analysis, and design.

A comparing of the primary functions of IDEF methodology with UMM workflow is shown in Fig. 3. As can be seen from this figure, IDEF0 focuses on inception step of business requirements while IDEF3 focuses on the elaboration step from integrated process. In addition, IDEF1X performs analysis and design of business objects at the elaboration step, while IDEF4 performs design of business objects and also its collaborative relationships. Package producing for classifying main concepts and business models is included. Requirement workflow applies business model as input to identify the requirements for target B2B solution. Detailed requirement analysis and UMM use case diagram are also included here. Analysis workflow embodies requirements by detailing occurring activities, collaboration between partners, and use case diagram in the former step. Design workflow defines precisely the dynamic view of collaborative relationships according to the data structure that is exchanged between business partners. Each workflow produces outputs which are lately being used as input for the next stage. Therefore, to reorganize these outputs considering the workflow area which is included as target of UMM and IDEF methodology as shown in Fig. 3, the following diagram in Fig. 4 is obtained.

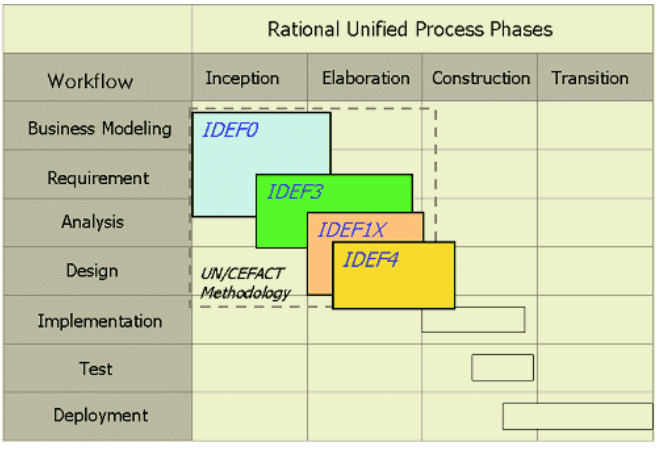


Fig. 3. Comparison between UMM with IDEF methodology on workflow

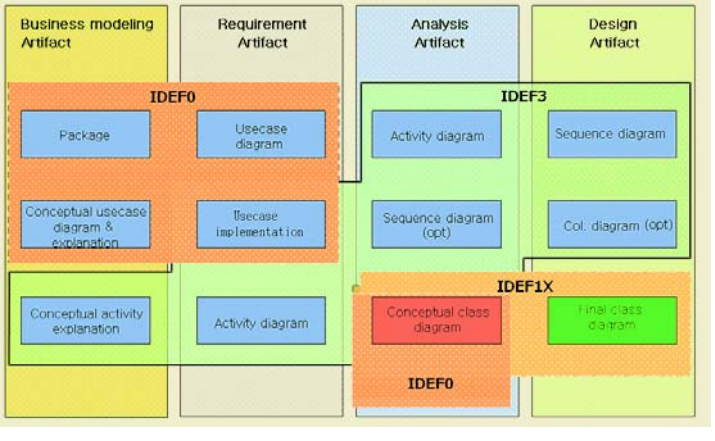


Fig. 4. UMM artifacts on workflow and IDEF methodologies

Basically, IDEF0 develops business processes and information based on business modelling and requirement workflow, provides information for producing package for business model classification, and producing and detailing former use case diagram. IDEF1 can be applied to produce the data structure which is exchanged between business partners. IDEF3 produces information that is needed to describe dynamic viewpoint of business partner's collaborative relationships from business modeling to design. Especially, as IDEF4 provides an object-oriented design methodology, it can be applied totally to produce UMM outputs throughout the entire workflow.

3.2 Outlines of IDEF Methodology Mapping

UMM workflow has to be able to be applied on newly targeted developing business domain, so development process should be applied to produce all the outputs gradu-

ally through the workflow step by step, as shown in Fig. 4. In this specification, the process of mapping from IDEF to UMM output does not utilize UMM workflow, but receives IDEF outputs to be used as inputs for producing UMM outputs. Therefore, with the final version of the UMM output, the mapping relationships among them can be easily described. Basically, by reorganizing UMM outputs, we will demonstrate the possibility of producing final outputs from IDEF0 outputs (See Fig. 5).

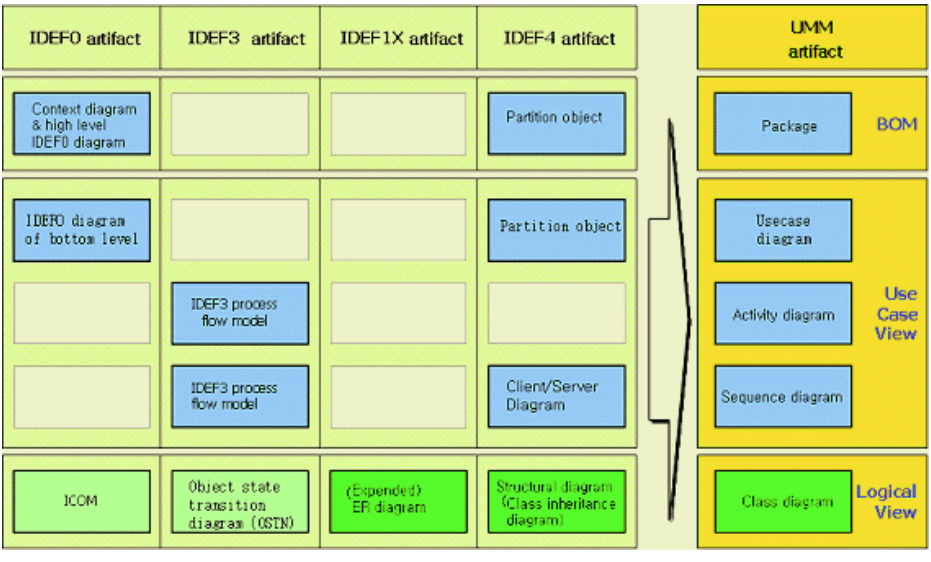


Fig. 5. UMM artifacts which are able to be produced from IDEF outputs

Consequently, the mapping rule to UMM outputs from prescribed guides of mapping can be summarized in detail as Table 1.

Table 1. A summary of mapping rule from IDEF artifact to yield UMM artifact

UMM artifact		IDEF artifact	
Diagram	Main elements to show	Target Model	Mapping Rule
Package	Business area, Process area, Business process	[IDEF0]	Produce suitable business area and process area with corresponding processes by analyzing node tree
		[IDEF0]	Confirm possibility of matching level0 diagram to business area, level1 to process area, and level2 and the lower to business process
		[IDEF0]	Examine the relationships among packages by interpreting connections among activities of diagram

Table 1. (*continued*)…

Use case diagram	Usecase	[IDEF0]	Produce the usecase from corresponding business process by analyzing BOM structure on the basis of node tree
		[IDEF0]	Produce concepts of inclusion and expansion by analyzing semantic of IDEF0 model
		[IDEF3PF]	Analyze the possible assumption that a process flow model belongs to an usecase.
	Actor	[IDEF0]	Produce actor by analyzing ICOM related to IDEF0 activity yielded as an usecase. Analysis is done in relation to the structure of mechanism, and also, by grasping input & output, whether they will participate as a major role can be determined.
	Usecase Scenario	[IDEF0]	Sketch usecase scenario referring description that explains activity
		[IDEF3]	Support scenario information from UOB elaboration which corresponds to referred IDEF0 function
Activity diagram	Activity	[IDEF3PF]	Match UOB to activity
	Trigger	[IDEF3PF]	Define trigger concerning simple precedence link information
	Syn bar & Decision activity	[IDEF3PF]	Convert to synchronization bar and decision activity according to the logical type of junction
	Activity Decomposition	[IDEF3PF]	Analyze process decomposition and activity decomposition from the same viewpoint
Sequence diagram	Object	[IDEF3PF]	Define basic object from information type referent
		[IDEF3OS]	Produce object concerning object state transition diagram
	Message flow	[IDEF3PF]	Estimate possible message low information from UOB elaboration and referent
Class diagram	Class	[IDEF0]	Produce basic object information with classifying ICOM into suitable object information and attribute information concerning bottom level function
		[IDEF3PF]	Define basic object from information type referent
		[IDEF3OS]	Produce object concerning object state transition diagram
		[IDEF1X]	Define class from entity
	Attribute	[IDEF0]	Produce basic object information with classifying ICOM into suitable object information and attribute information concerning bottom level function
		[IDEF1X]	Define class attribute from entity attribute
	Association	[IDEF1X]	Define class association from relationship
	Method	[IDEF3PF]	Extract method related information from UOB elaboration
		[IDEF3PF]	Produce method information from referent defined for each UOB
		[IDEF3OS]	Extract method information from referent added on the object state transition arrow

4 Case Application for Mapping

In Chapter 3, we discussed about the basic concepts that produce UMM output from modeling output of IDEF methodology. In this chapter, we will present apply our mapping rule on certain outputs as a case model of 'Buy & Sell Corn'. The model which implements a virtual enterprise for supply chain of corn using IDEF0 methodology is shown in Fig. 6.

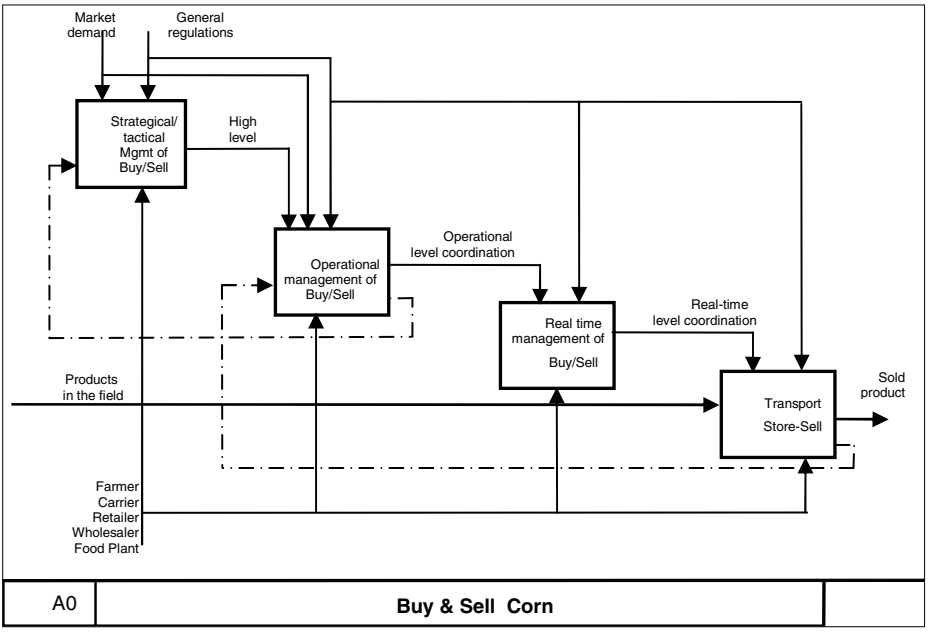


Fig. 6. IDEF0 diagram of level A0

Glancing at A0 diagram of the former case, the relationships among activities can be easily identified. Connected relations lies on activities which are marked from upper left of diagram to bottom right. The relationship that connects IDEF0 activities represents information of constraint proffer, function as input, and feedback. Package relationships are either dependency relation or transmission relation, but it is difficult to know where the IDEF0 relationships belong to. Thus, it is necessary to establish relationships which can discriminate semantics of IDEF0 interconnections. Package diagram generated by applying our guideline is shown in Fig. 7.

Applying usecase and actor produced in the former guidance, to sketch a usecase diagram results the following Fig. 8. This figure illustrates only a part of produced usecase, showing only the actors that belong to the farmer and wholesaler. We can see that activities (which is modeled by usecase) called "products availability, determination of quality range" are produced as actors in that they are performed by wholesaler management staff or performed with the support of primary information from farmer as inputs.

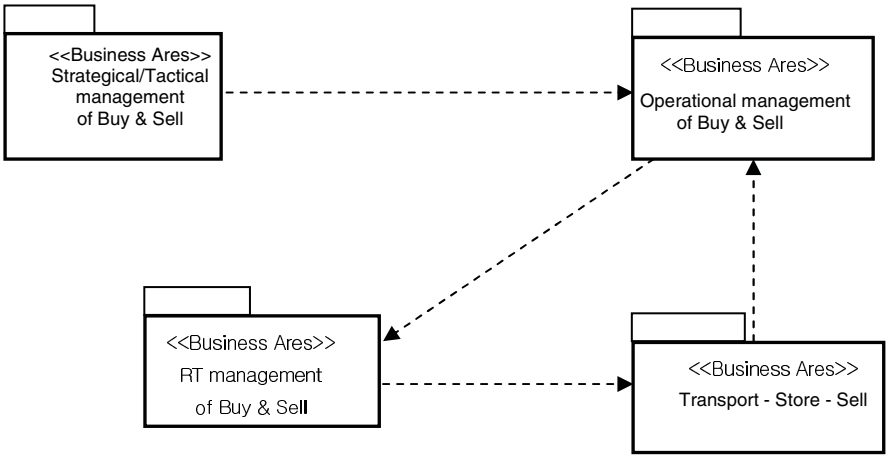


Fig. 7. Package diagram generated from IDEF0 model

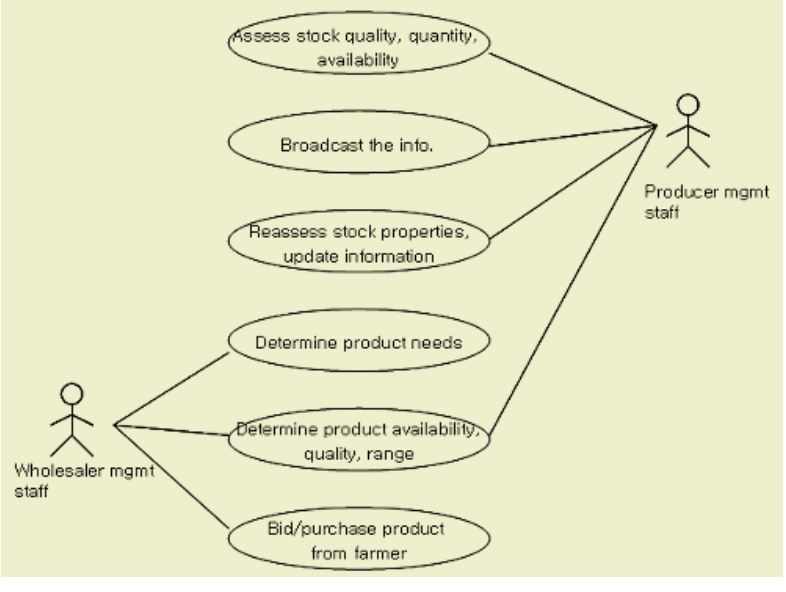


Fig. 8. A part of usecase diagram yielded

5 Conclusion

The UMM(UN/CEFACT Modeling Methodology), which is a modeling methodology of ebXML framework, is recommended as an effective business analysis model for electronic documents development in Korea. However, due to the fact that there are various different business process analysis/design methods being used by different enterprises and organizations, each outputs are is most cases not compatible. The

specification proposed herein provides a guideline on how to apply IDEF (Integrated Definition Language), which is the CALS standard, on UMM.

In this paper, we presented a guideline on the procedural steps by inducing IDEF model to apply UMM or system development of electronic commerce, including electronic document transmission. Assuming that IDEF modeling output is given, the result of this study designates how to make an output described by workflow of UMM. IDEF0, IDEF1X, and IDEF3 should be given as IDEF modeling output. If any one of these models are missing, this will mean that the person who performs the analysis and design activities will have more work.

References

1. Business Process and Business Information Analysis Overview, Version 1.0.
2. Information Integration for Concurrent Engineering (IICE) IDEF3 Process Description Capture Method Report, Knowledge Based Systems.
3. Information Integration for Concurrent Engineering (IICE) IDEF4 Object-Oriented Design Method Report, Knowledge Based Systems.
4. Information Technologies – Open-EDI Reference Model. ISO/IEC 14662 : 1997(E). International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), 1997.
5. Integration Definition for Function Modeling (IDEF0), FIPS 183.
6. Integration Definition for Information Modeling (IDEF1X), FIPS 184.
7. UN/CEFACT Modeling Methodology, CEFACT/TMWG/N090R10. UN/CEFACT.

A Case Study on the Development of Employee Internet Management System

Sangkyun Kim and Ilhoon Choi

Somansa, Woorim e-Biz Center, Yangpyeongdong 3-ga, Yeongdeungpogu,
Seoul, Korea
{saviour, acechoi}@somansa.com

Abstract. As enterprise computing environments become more Internet-oriented, the importance of Internet traffic monitoring and filtering increases. The Internet monitoring and filtering system was introduced as a simple utility for analysis and blocking access to objectionable Web sites. And, it has evolved into a major corporate application. This paper provides our experience to design and develop Web traffic monitoring and filtering system called WebKeeper. In this paper, we present requirements, design, implementation and operational statistics of WebKeeper. With the use of this system including advanced databases, it now will provide enhanced security, productivity and increased bandwidth in the form of the Internet access management.

1 Introduction

Web traffic monitoring and filtering system is an application by which users have been blocked from accessing inappropriate Web sites and content. Controlling access to the Internet by means of filtering software has become a growth industry in many countries [1]. The motivation to block certain sites is driven by business goals to reduce legal liability and nonproductive Web surfing, and to lower bandwidth cost [2]. The primary technology used has been a negative database. The database of URL addresses for objectionable sites has been built and maintained by the vendor, and provided with regular updates to the employer. The database resides on a filtering server on the employer network, and all employee Web traffic is first checked against the forbidden list of sites.

This paper provides our experience of requirement analysis, design, implementation and operational statistics which were learned from the development of WebKeeper: Web traffic monitoring and filtering system. Major function and impacts of WebKeeper are as follows: 1. Increase productivity by blocking non-business sites with categorized URL database and rule-based polices, 2. Make provision against information leakage by archiving and searching all information through Web-hosted mail, Web-board messages and its attached files, 3. Reports network usage patterns and conditions with itemized charts and tables.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1229 – 1238, 2005.

2 Requirements

There are several researches which address requirements of the security system or software. We reviewed researches on an architecture, function and performance of the security system or software.

Scott provided comparison charts of network security and protection solutions. He used comparison factors: function, minimum RAM, hardware, operating system, source language, source code available, network operating systems, network protocols supported, networking environment supported, network monitoring protocols supported, Web services supported, component model supported, pricing, special marketing programs, first installed, maintenance, support services, additional requirements, product options, and product description [3]. Security functions may be classified as preventive, detective, deterrent, recovery or corrective. Preventive functions attempt to avoid the occurrence of unwanted events. Detective functions attempt to identify unwanted events during they are occurring or after they have occurred. Deterrent controls are intended to discourage individuals from intentionally violating information security policies or procedures. Recovery controls restore lost computing resources or capabilities and help the organization recover monetary losses caused by a security violation. Corrective controls either remedy the circumstances that allowed the unauthorized activity or return conditions to what they were before the violation [4, 5]. ISO9126 offers a framework for appraisal of software on the basis several areas of performance. The areas of investigation illustrated in ISO9126 and there are as follows: functionality, reliability, usability, efficiency, maintainability, portability [6]. Table 1 shows the most important requirements that should be satisfied in general security systems.

Table 1. Requirements of the security system

Architecture	Function	Performance
optimal hardware requirement, common OS supported, verified source language, common NOS supported, common protocols supported	preventive, detective, deterrent, recovery, corrective	functionality, reliability, usability, efficiency, maintainability, portability

In this paper, we take an architecture, functionality and performance shown in table 1 into consideration to describe requirements of the Web traffic monitoring and filtering system as a security system.

2.1 Architecture

Web traffic monitoring and filtering system should be easy to install and use on any network segment. Every administrator who wants to monitor and filter their network segment should be able to install this system on a notebook or desktop and connect it to his main network segment with ease. So, we took Microsoft Windows family as a

main OS(Windows NT/98/95 with Internet Explorer 5.0 or higher installed) and Intel compatible PCs as a hardware platform.

2.2 Function

Robert suggested functionalities which represent specific activities associated with common security objectives: hiding, encrypting, decrypting, locking, limiting, reserving, filtering, shielding, containing, authenticating, access controlling, enforcing, tunneling, obliterating, eradicating, replicating, mirroring, preserving, restoring, probing, scanning, monitoring, logging, inspecting, auditing, integrity checking, notifying, reporting, patching, substituting, proxying, inoculating(vaccinating), and retreating [7]. Kim suggested functional criteria of firewall, IDS, anti-virus, and VPN [8]. The relationship of Krause's model, Firth's model and Kim's model is shown in table 2 [4, 7, 8]. As shown in table 2, we can describe functional criteria of certain security control using Krause's model and Firth's model.

As described in table 1 and 2, we took five functional criteria and several function list into consideration to generate functional requirements of Web traffic monitoring and filtering system. Krause's model was used to derive functional criteria. Firth's model was used to derive function list. Functional requirements of Web traffic monitoring and filtering system, WebKeeper, are summarized in table 3.

Table 2. Functional criteria of security controls

Criteria and function list of security controls		Firewall	IDS	Anti-virus	VPN
Preventive function	hiding	x	x	x	x
	encrypting & decrypting	x	x	x	o
	locking	x	x	x	o
	limiting & reserving	x	x	x	x
	filtering	o	x	x	x
	shielding	o	o	o	x
	containing	o	x	x	o
	authenticating & AC	o	x	x	x
	enforcing	x	o	x	x
	tunneling	x	x	x	o
	obliterating & eradicating	x	x	x	x
	substituting	x	x	x	x
	proxying	x	x	x	x
Detective function	probing & scanning	x	o	o	x
	monitoring, logging & auditing	o	o	o	o
	inspecting	x	o	o	o
	integrity checking	x	x	o	o
Deterrent function	notifying	o	x	x	o
Recovery function	reporting	o	o	o	o
	replicating & mirroring	x	x	x	x
	preserving & restoring	x	x	x	x
	patching	x	x	x	x
Corrective function	inoculating(vaccinating)	x	o	o	x
	retreating	x	o	o	x

Table 3. Functional requirements of Web traffic monitoring and filtering system

Criteria	Function List	Description
Preventive function	Filtering	✓ Blocking on objectionable Web sites predefined in filtering URL database.
	Authentication	✓ Access to Web traffic monitoring and filtering system must be controlled using strong access controls.
Detective function	Monitoring	✓ Keeps every screen (html and picture) and URL surfed.
	Web activity	✓ Keeps every uploaded message and its attached files through Web-hosted mail and Web board.
	Inspecting (Keyword searching)	✓ For every message sent through Web-hosted mail or Web board.
		✓ For every attached files (more than 200 formats including zip, rar compressed format).
Deterrent function	Notifying (Warning Banner)	✓ A warning banner which contain alerting messages that user's navigation will be monitored, recorded and analyzed must be shown when user starts to access Internet.
	Notifying	✓ Customizable deny messages that are displayed to end users.
Recovery function	Reporting	✓ Network usage reporting by category, day, hour, frequently visited sites, etc.
		✓ Network policy results by group, IP, specific category, etc.
Corrective function	Inoculating (Establishing network usage policy)	✓ Filtering Web activity by clients(IP), servers(URL, subURL or IP), and site categories (filtering URL DB).
		✓ Access controlling for clients(groups), time(business time or day time) or categories(user-defined site blocking).
	Retreating	✓ Filtering URL must be updated to keep up the latest list of objectionable Web site.

2.3 Performance

We used ISO9126 to derive performance characteristics. Performance requirements of Web traffic monitoring and filtering system, WebKeeper, are summarized in table 4.

Table 4. Performance requirements of Web traffic monitoring and filtering system

Performance factor of ISO9126	Related char-acteristics	Description
Functionality	Security of system itself	Should support promiscuous mode to intercept and read each network packet that arrives in its entirety without its own IP
Reliability	Fault tolerance	The error of other systems in same network does not cause any problem to Web traffic monitoring and filtering system
	Fail-soft	If Web traffic monitoring and filtering system crashed, all network traffics should be transferred without blocking
Usability	GUI	User interface based on graphics (icons and pictures and menus) instead of text should be supported
Efficiency	No lags	Should not cause additional burden on network (passive sniff-ing technology)
Maintainability	Remote mgmt.	Management on remote systems should be supported
Portability	No need for reconfiguring network	Should be easy to install and use on any network without reconfiguring network. Anyone wants to monitor and filter a specific network segment should be able to install this system on a notebook or desktop and connect it with ease

3 Design

Based on the requirements discussed in the previous section, we have designed a Web traffic monitoring and filtering system. The designed architecture of WebKeeper is illustrated in Fig. 1. The overall system consists of three parts: *Filtering Agent, Database* and *Administration Manager*. *Filtering Agent* captures Internet traffics, sends blocking packets if needed, and generates logs. *Database* contains data used in WebKeeper. *Administration Manager* provides management functions on *Filtering Agent* and *Database*. *Filtering Agent* consists of the packet capturer, blocking packet generator, packet analyzer, receiver and log processor. *Database* consists of the filtering policy, filtering URL, and log. *Administration Manager* consists of the rule administrator, update manager, viewer, finder, and reporter.

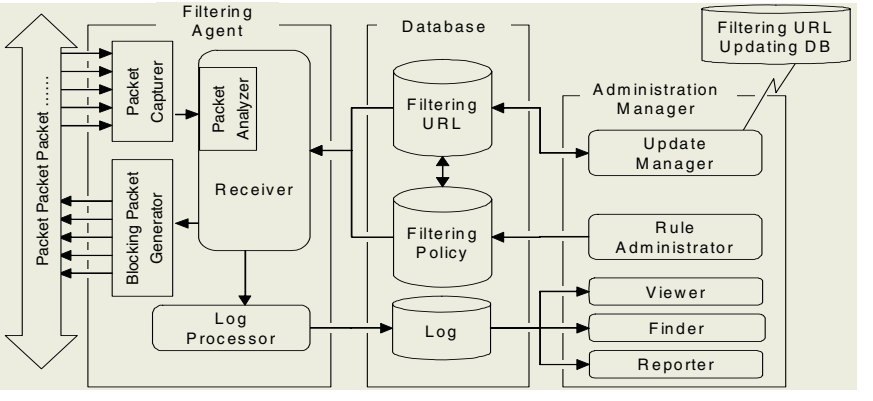


Fig. 1. Architecture of Web traffics monitoring and filtering system

3.1 Filtering Agent

It consists of the packet capturer, blocking packet generator, packet analyzer, receiver, and log processor. To capture all network packets, it acts in the promiscuous mode. This is due to the property of the Ethernet itself: packet broadcasting. Capturing all packets is the most important and basic operation of *Filtering Agent.* After capturing of

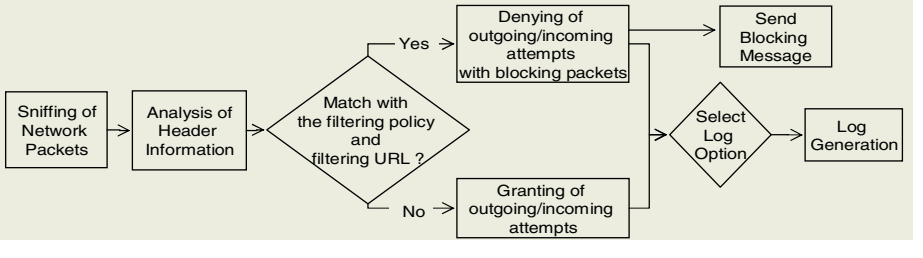


Fig. 2. Logic of the *Filtering Agent*

network packets, it operates according to the logic shown in Fig. 2. The packet capture device sniffs all network packets. The receiver analyzes each header of these packets and matches it with the filtering policy and filtering URLs.

The packet capturer captures packets from the network. The packet analyzer extracts the information of each packet. The receiver compares the information of each packet with filtering URL and filtering policy. The blocking packet generator sends blocking packet to block objectionable Web site. The log processor writes the information that the receiver analyzed and blocked to the log file.

3.2 Database

Database of WebKeeper consists of the filtering policy, filtering URL, and log. The filtering policy is managed by the rule administrator in *Administration Manager*.
There are four methods which have been widely adopted in construction of a filtering URL database, some in conjunction with one another, by many of the major filter vendors [9]. These are blacklists, keyword blocking, whitelists and rating systems [10]. In WebKeeper, we used *Blacklists* and *Rating Systems*. Table 5 shows non-business categories of the filtering URLs that should be blocked. Business categories of the filtering URLs that should be granted are as follows: news, education & research, H/W & S/W, enterprise, government & politics, reference, web mail, Internet phone & mobile, unclassified.

Table 5. Non-business categories of Web traffic monitoring and filtering system

Category	Description
Adults only	Material labeled by its author or publisher as being strictly for adults and porn site.
Illegal	Advocating, promoting, or giving advice on carrying out acts widely considered illegal national or international law. This includes lock-picking, bomb-making, fraud, breaching computer security (hacking), phone service theft, pirated software archives, or evading law enforcement.
Gambling	Gambling services, or information relevant primarily to gambling.
Investment	The investing category contains URLs that deal with personal investments and investment options. URLs in this category may contain stock quotes, money management forums, online publications, discount brokerage services, trading stock/mutual funds, portfolio set up, and more.
Portal sites	This category contains URLs that serve as a major starting point for users when they get connected to the Web.
Chat	Chat sites, services that allow short messages to be sent to others immediately in real time.
Entertainment	The entertainment category contains URLs devoted to movies, television, music, hobbies, clubs, and amusement parks. The sites can either be fan-club style or sponsored by a company.
Job search	This category contains URLs related to a job search. It encompasses sites concerned with resume writing, resume posting by individuals, interviewing, changing careers, classified advertising, and large job databases.
Non-business sales	This category contains URLS that include any form of merchandise sales or service that benefits the individual only, such as the sale of clothing, accessories, appliances, pets, and more.

The granted or denied attempts of incoming or outgoing sessions of the users were recorded in the log database. The viewer, finder and reporter of *Administration Manager* use the log database.

3.3 Administration Manager

It consists of the rule administrator, update manager, viewer, finder, and reporter. The rule administrator enables to create rules that will control the Internet access of the users to be monitored and blocked. These rules can govern who can access what areas of the Internet at what time of day. Rules can be positive (allowing access to sites or categories) or negative(denying access to sites or categories of sites). The update manager periodically monitors remote DB which contains up-to-date filtering URLs and takes updated URLs. The viewer provides a view of activity by user and sites requested. As traffic is generated on the network, information about user activity is recorded in the log database and then displayed by the viewer module. Also, it shows activity on the network as it is happening in real time. As traffic is generated on the network, administrator can see it appear in the viewer. The finder supports various search functions about server IP, client IP, keyword, filename and contents of file with time and date, and Boolean connectors. The reporter is a rule-based reporting which supports user-defined conditions of group, clients, date, time, etc. It shares policy with the viewer.

4 Implementation

According to the requirements and design specifications described above, we implemented Web traffic monitoring and filtering system. WebKeeper consists of three parts: *Filtering Agent*, *Database*, *Administration Manager*. WebKeeper can be installed as shown in Fig. 3.

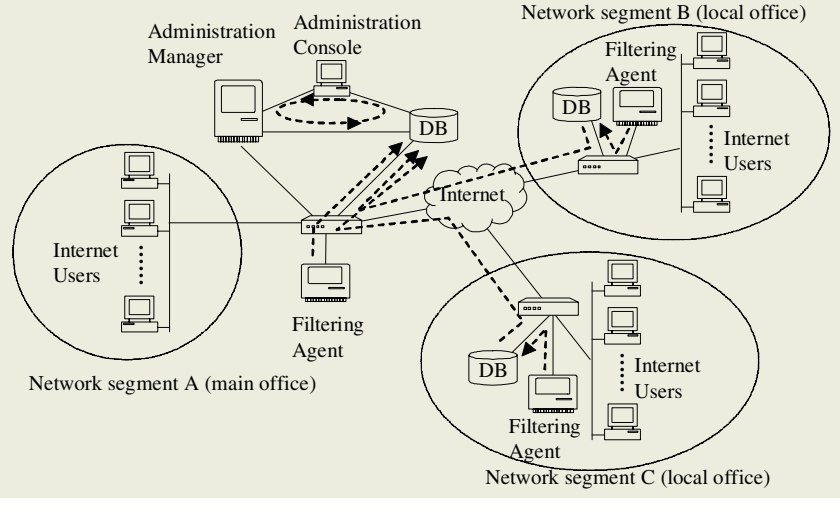


Fig. 3. Installation diagram of WebKeeper: Web traffic monitoring and filtering system

Filtering Agent may be installed in several points to monitor and filter traffics of each network segment. *Database* may be installed with each *Filtering Agent*. And, *Databases* installed in several network segments interoperate with central *Database* installed in the main office. With administration console located in the main office, administrator can connect to *Administration Manager* and its *Database*. Major functions of *Administration Manager* are as follows: rule administrator, viewer, reporter and finder. Fig. 4 shows main windows of manager system of WebKeeper.

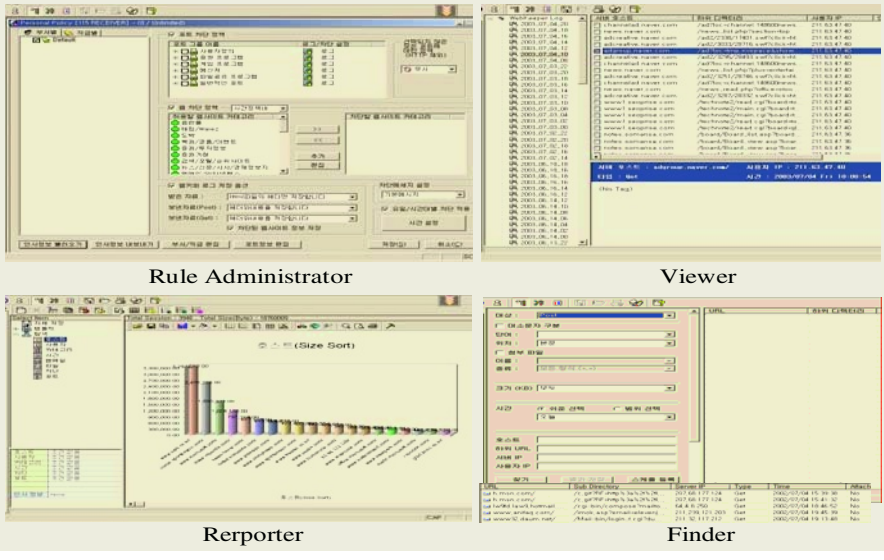


Fig. 4. Main windows of *Administration Manager* of WebKeeper

5 Operational Statistics

We performed a case study to show the operational statistics of this system. The survey system is implemented in ABC Co. Ltd.. Features of ABC Co., Ltd. are as follows: 7,000 employees, electronics manufacturer. Pattern of the sessions by daily-working time is shown in fig. 5.

The non-business traffic increases after lunch time and before the closing time. The business traffic is maximized in 11.00 am.

Operational statistics of employees' Internet access of ABC Co., Ltd. are summarized in table 6. Table 6 is average data of the business and non-business Internet access of one week.11.54% of one business day traffic is non-business traffic that The non-business traffic increases after lunch time and before the closing time. The business traffic is maximized in 11:00 am means non-working hours of enterprise employees. About half of non-business traffic is an access to the adults or illegal contents that causes legal liability of employee for negligent management on employees' illegal activities.

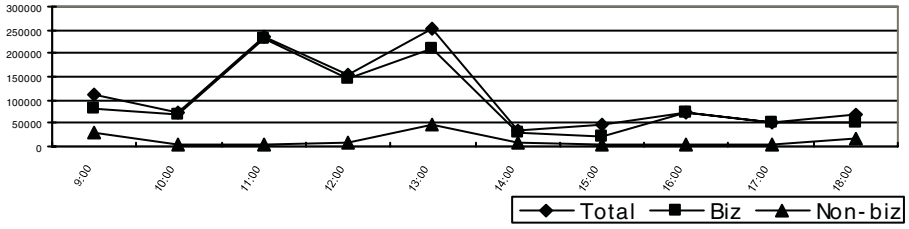


Fig. 5. Pattern of the sessions by daily-working time of ABC Co.,Ltd.

Table 6. Operational statistics of ABC Co.,Ltd

Category / Ranking	Business				Non-business			
	Category	Session	% per biz sites	% per total sites	Category	Session	% per non-biz sites	% per total sites
1	News	430,543	44.13	39.04	Adults only	30,284	23.80	2.75
2	Education & Research	295,438	30.28	26.79	Illegal	28,453	22.37	2.58
3	H/W & S/W	94,728	9.71	8.59	Gambling	17,632	13.86	1.60
4	Enterprise	52,945	5.43	4.80	Investment	16,273	12.79	1.48
5	Government & Politics	40,890	4.19	3.71	Portal sites	12,305	9.67	1.12
6	Reference	34,893	3.58	3.16	Chat	8,362	6.57	0.76
7	Web mail	19,282	1.98	1.75	Entertainment	5,092	4.00	0.46
8	Internet phone & Mobile	3,953	0.41	0.36	Job search	4,968	3.91	0.45
9	Unclassified	2,932	0.30	0.27	Non-biz sales	3,849	3.03	0.35
Sum		975,604	100.00	88.46		127,218	100.00	11.54

6 Conclusion

In this paper, we provided our experience of requirement analysis, design and implementation, and know-how which were learned from the development of WebKeeper: Web traffic monitoring and filtering system. Major function and impacts of WebKeeper are as follows: 1) Increase productivity by blocking non-business sites with categorized URL database and rule-based polices. 2) Make provision against information leakage by archiving and searching all information through Web-hosted mail, Web-board messages and its attached files. 3) Reports network usage patterns and conditions with itemized charts and tables.

The limitations of Web traffic monitoring and filtering system are as follows.

Insufficient interoperability with enterprise's HRM(Human Resource Management) system: Web traffic monitoring and filtering system should be interoperated with enterprise's HRM system for better authentication, policy management of filtering rule, and

utilization of log files. HRM system is a business application for the management of HR-related transactions, best practices and enterprise reporting. Functions typically include core HR tracking, payroll, benefits, recruiting, competency management, training, time management, performance management and self-service offerings. But, standard which provides communication interface between Web traffic monitoring and filtering system, and enterprise's HRM system is not defined yet. So, the communication interface of Web traffic monitoring and filtering system should be altered according to enterprise-specific HRM systems in every time of implementation.

Backlash effects due to the quality of filtering URL database: The major problem faced by the developers of filtering products is the sheer size of the Internet, and the difficulties this creates in building up lists of unacceptable sites. Most of vendors build their black lists by crawling the Web, and adding pages containing suspect material to black lists. These interim lists are supposed to be checked by real people before the sites make it onto the published black lists. Anti-censorship groups have expressed strong doubts about the effectiveness of this manual review, given the number of innocuous sites that have been blocked by major filter vendors [11]. Filtering errors will either result in undesirable content coming through, or in desirable content being blocked. So, Web traffic monitoring and filtering system may cause backlash effects.

The overall impacts of Web traffic monitoring and filtering system have a positive value because monitoring and filtering mechanisms cause users to lose a small number of URLs including useful information but save users from various risk points. With the use of Web traffic monitoring and filtering system including advanced databases, it now will provide enhanced security and productivity, and increased bandwidth in the form of the Internet access management to enhance the enterprise's competitiveness.

References

1. Rosenberg, R.S.: Controlling Access to the Internet: The Role of Filtering. Ethics and Information Technology **3** (2001) 35-54
2. Gassman: Internet Filtering and Reporting: Websense vs. SurfControl, Research Note. Gartner (2002)
3. Beall, S. and Hodges, R.: Protection & Security: Software Comparison Columns. Gartner (2002)
4. Krause, M. and Tipton, H.F.: Handbook of Information Security Management. CRC Press (1998)
5. Kim, S.: Security Consultant Training Handbook. Hyundai Information Technology (2002)
6. ISO: ISO9126-1: Software Engineering - Product Quality - Part 1: Quality Model, No. ISO/IEC 9126-1:2001. International Organization for Standardization (ISO), Geneva (2001)
7. Firth, R. et al.: An Approach for Selecting and Specifying Tools for Information Survivability. Software Engineering Institute, Carnegie Mellon University (1998)
8. Kim, S. and Kang, S.: A Study on the Security Vulnerabilities and Defense Mechanism for SET-based Electronic Commerce. The Journal of Korean Institute of CALS/EC **4** (1999)
9. Glazer, C.: Filtering & Pornography. http://cse.stanford.edu/class/cs201/projects-98-99/online-pornography/ (1999)
10. Hochheiser, H.: Computer Professionals for Social Responsibility Filtering FAQ. CPSR (2001)
11. Greenfield, P. et al.: Effectiveness of Internet Filtering Software Products. CSIRO (2001)

Cost-Benefit Analysis of Security Investments: Methodology and Case Study

Sangkyun Kim[1] and Hong Joo Lee[2]

[1] Somansa, Woorim e-Biz Center, Yangpyeongdong 3-ga, Yeongdeungpogu, Seoul, Korea
saviour@somansa.com
[2] P&M Research Team, Quality & Reliability Lab., DAEWOO Electronics Corp., 412 Cheong-chun 2 dong, ByuPyung, Incheon, Korea
phileo21@empal.com

Abstract. We live in an unsafe world in which we encounter threats against our safety and security every day. This is especially true in the information processing environment. Managements are engaging and facing difficult problems to manage information security issues. One of the most brain-teasing management issues is "How they could make a decision on security-related investment to maximize the economic balance?" To solve this problem the ROI of security investments must be measured and managed. This paper provides the integrated methodology which consists of a process model and analysis criteria of cost factors and benefit factors to support an economic justification of security investments. Also, a case study is provided to show practicality of this methodology.

1 Introduction

Some recent studies of Alpar (1990), Barua (1995), Brynjolfsson (1996), Mahmood (1993), Mitra (1996) and Rai (1997) that are focusing on relationships between investment in IT and organizational performance and productivity have reported positive and significant effects of such investment [1, 2, 3, 4, 5, 6].

As organizations become more and more dependent on their computer-based information systems, which play a vital role and important part in their business operations, there needs to be a greater awareness and concern about the security of these systems. Information security appears on the list of critical success factors of most major organizations today [7]. The successful management of information security within an organization is vital to its survival and success. Any company's balance sheet has a finite tolerance for risk; security advisors contribute to business success by purifying the overall risk the company holds, allocating more of the available risk tolerance to risks that actually could bear fruit by removing the risks that, at best, lead nowhere. But to do this, the ROI of security interventions must be measured [8]. However, the organization needs an effective methodology to manage the ROI of information security investments. According to Checkland (1981), "a methodology can be described as a set of tools, perhaps even research methods, or a bridge, translating management theory to management practice" [9]. But, previous researches on an economic justification of information systems lack in providing practical methods

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1239 – 1248, 2005.

for security investments. In this paper, we defined key characteristics of an economic justification methodology of information security systems as following: 1) Security concern: concentration on information security controls; 2) Methodology concern: provision of steps and tools for an economic justification; 3) Practicality: practical usability of a methodology.

In this paper, we provide a process model and analysis criteria of cost factors and benefit factors. This paper also provides a case study to prove practical values of this methodology.

2 Previous Researches

We categorized existing researches that are related with the economic justification of security investments for information systems into three groups: evaluation of information systems, investment evaluation of information systems, and investment evaluation of information security systems. These previous researches are introduced in chapter 2.1, 2.2 and 2.3. Limits of previous researches are summarized in table 1.

Table 1. Limits of previous researches

Research \ Limits	Security concern	Methodology concern	Practicality
Evaluation of information systems	Do not consider security related factors and controls	Do not consider economic justification issues	Do not provide benefit criteria of system operations
Investment evaluation of information systems	Do not consider security related factors and controls	Applicable	Lack in provision of practical usage
Investment evaluation of information security systems	Applicable	Only provides high level category of cost factors and benefit factors	Do not provide case studies or practical usage

2.1 Researches on the Evaluation of Information Systems

Researches related with an evaluation of information systems consist of success studies which explain operational influences such as productivity, user satisfaction, and profitability, etc. and models which measure the strategic influences such as buyer and consumers, competitive rivalry, suppliers, etc. The original research of a success of information systems is Delone, et al. (1992) [10]. Afterwards, many researches were verified and extended. These are summarized in table 2.

Mahmood, et al. (1991) made a general model to measure potential influences of information systems on an enterprise and validated it by interviewing strategic managers and reviewing researches [11]. Palvia (1997) made GLIT model by extending that of Mahmood's assuming conditions of an enterprise [12]. But, Palvia's model focused on establishing a logical framework, so the practical application of this model is very difficult.

Table 2. Comparison of IS success studies [10, 13, 14, 15]

	Delone, et al. 1992	Saarinen, 1994	Grover et. al., 1996	Torkzadeh et. al., 1999
Research method	Literature review	Literature review & statistical verification	Literature review	Literature review & statistical verification
Evaluation factors	System quality, Information quality, Information use, User satisfaction, Individual impact, Organizational impact	Development process, Use process, Quality of the IS product, Impact of the IS on organization	Infusion measure, Market measure, Economic measure, Usage measure, Perceptual measure, Productivity measure	Task productivity, Task innovation, Customer satisfaction, Management control

2.2 Researches on the Investment Evaluation of Information Systems

Renkema, et al. (1997) discerned four basic approaches such as the financial approach, multi-criteria approach, ratio approach, and portfolio approach [16]. Dan, et al. (2000) group evaluation approaches into four categories: economic appraisal techniques, strategic approaches, analytical appraisal techniques, and integrated approaches [17].

But, complex methods are not accepted usually. As a result, professionals of information systems often adopt the analytic hierarchy process (AHP). AHP includes three major steps: identifying and selecting criteria, weighting the criteria and building consensus on their relative importance, and evaluating the IS using weighted criteria. So, the selection of criteria is very important [18]. Bacon (1992) found that criteria of the support of explicit business objectives and response to competitive systems are important in selecting IS investments [19].

2.3 Researches on the Investment Evaluation of Information Security Systems

Security investments are justified based on the cost of doing business and the protection of corporate assets [20]. Often, security investments are likened to an insurance policy - they are used to mitigate risk but usually do not provide a quantifiable return on investment (ROI). Another justification for security investment is a situation where regulatory compliance is required. Many incidents show that security investments could be justified as a competitive advantage [21, 22]. Bob (2001) defined security ROI: Returns on Investment = ((Change in Revenue)+(Cost Savings))/(Investment) [23]. Investment includes an initial purchase cost, renewal cost, administrative expense, and so on. Change in revenue means how an investment in security might increase revenue. Adding security functions will allow to do things that would have been too risky to do otherwise. Cost savings is really about in a security context is loss avoidance. Witty (2001) suggested "The Information Security TCO chart" that has accounts categorized into five major sections: 1) Hardware, 2) People, 3) Soft-

ware, 4) External services, 5) Physical security [24]. These five sections are then divided into these security activities: authentication, authorization, code protection and promotion, cyber-Incident response, content monitoring, digital rights management, employment due diligence, encryption , firewalls, information protection insurance, information security program, internet content filtering and monitoring, intrusion detection, license compliance, logging, reporting and auditing, malicious code management, message integrity, privacy management, public-key infrastructure, record classification and retention, remote access, risk assessment, security administration, security architecture, single sign-on standards certification, transaction incident management, and vulnerability assessment and management.

3 Approach of This Methodology

Malik (2001) provided three categories of decision makers for information security systems: the informed, risk takers and the security ignorant [25]. The informed typically understand technology, information security risk and the need for controls. Informed decision makers drive the process and seek input from internal and external sources before funding security initiatives. Risk takers are somewhat aware of security issues, such as threats and vulnerabilities, but are much more willing to take a chance. The security-ignorant minimally understand information security risk, but prefer to avoid funding initiatives rather than get involved or take responsibility for the security issue.

There is little value in identifying and quantifying risk without implementation of cost-effective security measures to counter the risk. An important aspect of a security analysis and management methodology is the provision of some means by which countermeasures can be selected and implemented on a cost-effective basis to support decision makers.

This methodology suggests cost and benefit factors to provide an analytical view of the economic justification of an investment on information security systems. Benefits have various types of expression. And, decision makers concern not only monetary terms which are forced with irrationality but also strategic terms which take relative and descriptive reporting forms of professional and analytical view point [25]. So, case specific methods which cooperate with process model and cost-benefit factors of this methodology should be defined considering on their own status, characteristics, and purpose. This methodology can be used in two cases: an economic justification of an investment planning (before implementation of information security systems) and a cost-benefit assessment of implemented information security systems(after implementation of information security systems).

4 Process Model

The process model of an economic justification methodology is shown in fig. 1 which is derived from Kim (2000)'s model [26]. Kim's model focuses on an economic justification of information systems in performance and productivity issues, and lacks in providing detailed processes and considering two aspects of time situation. We pro-

vide two justification paths of "before implementation" and "after implementation". Upper half side of fig. 1 shows a process of "before implementation" justification which focuses on a justification of the investment planning. Lower half side of fig. 1 shows a process of "after implementation" justification which focuses on a cost-benefit assessment of invested resources. Characteristics of each step are described in table 3.

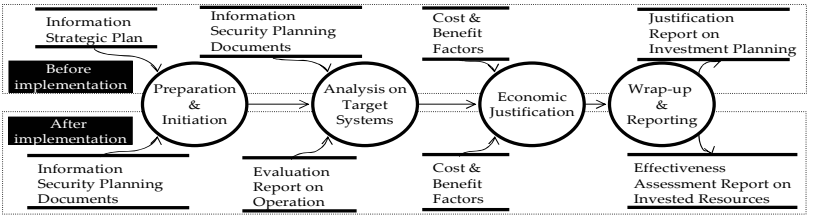


Fig. 1. Process model

Table 3. Characteristics of process model

Case / Step	Justification on planning	Assessment on operational benefits
Preparation & initiation	analyze security-related investment factors that can be found in information strategic planning	analyze an investment objective, background and categories (including detailed sub-items) on information security planning
Analysis on target systems	analysis on information security planning (planned investment, estimated effects of information security systems)	analysis on invested resources and operational effects of current systems
Economic justification	overhead cost and strategic benefits must be considered carefully	assessment on finished investment can find ignored or missed cost factors which had not been included on past investment planning but must be cared for next planning periods.
Wrap-up & reporting	reporting can be used for decision making on investment	reporting may be used for performance appraisal on a past investment and reformation of planned investment (on-going investment)

5 Analysis Criteria

This methodology provides two kinds of criteria: cost factors, benefit factors. Cost factors are classified in nine groups. Benefit factors consists of operational benefits and strategic benefits.

5.1 Cost Factors

Cost factors mean allocated resources such as equipment and labor. Almost of these factors could be quantified. That is to say, the measuring of cost factors is easier than the measuring of benefit factors.

Kim (2001) suggested cost factors (classified into 12 categories) of information systems [26]. Witty (2001) suggested "The Information Security TCO chart of accounts" which is categorized into five major sections: hardware, people, software, external services, and physical security [24]. Harris (2001) suggested cost factors need to be considered and evaluated when deriving the full cost of a countermeasure: product cost, design/planning costs, implementation costs, environment modifications, compatibility with other countermeasures, maintenance requirements, testing requirements, repair/replace/ upgrade costs, operating/support costs, and effects on productivity [27]. Roper (1999) suggested cost factors: purchase price; life cycle maintenance costs (installation, preventive maintenance, repair, warranty, and replacement); life expectancy; salaries for staff/contractors to implement, maintain, monitor, or train others to use countermeasure [28]. Cost factors are classified in nine groups with two perspectives of the lifecycle and control categories. Control categories are derived from researches [29, 30, 31, 32, 33]. Cost factors are provided in table 4.

Table 4. Cost factors of a security investment

	Administrative	Logical	Physical
Planning	- Loss of working - Staffing(planning TFT) - Consulting - Awareness/training/education	- Computing - Communication equipment - Analysis tool - System downtime	- Space - Supporting utility - Tempest/shielding - Monitoring/alarming
Implementation	- Loss of working - Public charge - Staffing(implementation TFT) - Outsourcing (app. development, system build up) - Awareness/training/education	- OS - S/W - H/W - DB - Contents - Comm. equipment - System downtime	- Space - Supporting utility - Tempest/shielding - Monitoring/ alarming
Operation	- Insurance - Public charge - Staffing(CISO, administrator, auditor, helpdesk, guard, housekeeping) - Certification - Awareness/training/education - Loss of productivity	- Insurance - Maintenance - Repair/replace /upgrade	- Insurance - Maintenance - Repair/replace/ upgrade

5.2 Benefit Factors

Scott (2001) suggested factors of potential loss for emergency status which can be taken place if security controls were not workable: productivity, revenue, damaged reputation, financial performance, and other expenses [34].

We classified benefits into operational benefits and strategic benefits. Operational or strategic benefits may be categorized into one of three types of expression factor: economic factor which is measured and evaluated with monetary terms, numerical factor which are measured and evaluated with number or volume, and qualitative factor. Operational benefits mean the enhanced efficiency of organization's opera-

tions. They consist of cost saving, added profitability, enhanced decision-making and enhanced business function. It is described in table 5.

Table 5. Operational benefit factors

	Measurement index	Measurement factor	Characteristic
Cost saving (Prevention of potential losses)	Revenue	Direct losses	Economic
		Compensation payment	Economic
		Billing losses	Economic
		Investment losses	Economic
	Service cost saving	A/S cost saving	Economic
	Firm infrastructure cost saving	Equipment rental saving	Economic
		Space rental cost saving	Economic
		Travel/shipping cost saving	Economic
	Human resource cost saving	Prevention of potential work losses	Economic
		Temporary employment saving	Economic
		Overtime cost saving	Economic
	Financial performance	Lost discounts	Economic
Added profitability	Increase of sales	New customer occurrence	Numerical
		Existing customer preservation	Numerical
	Increase of profitability	Added premium value	Economic
		Enhanced productivity	Economic
Enhanced decision making	Time reduction	Reduced decision making time	Numerical
		Reduced decision making step	Numerical
	Enhanced quality	Enhanced problem cognition	Qualitative
		Enhanced correctness	Qualitative
Enhanced business function	Enhanced flexibility		Qualitative
	Enhanced credibility		Qualitative

Table 6. Strategic benefit factors

	Measurement Index	Measurement factor	Characteristic
Reduced threat of rivalry	Differentiation	Intensifying product functionality	Qualitative
		Intensifying product awareness	Qualitative
		Intensifying switching cost	Qualitative
	Cost advantage	Credit rating	Qualitative
		Stock price	Qualitative
Enhanced supplier relationship	Increased supplier	Supplier extension	Qualitative
		Availability of company searching	Qualitative
	Enhanced supplier manipulation	Enhanced negotiation	Qualitative
		Enhanced quality management	Qualitative
Enhanced customer relationship	Increased customer	Customer extension	Qualitative
		Availability of customer searching	Qualitative
	Enhanced service	Availability of product/service information	Qualitative
		Administrative support	Qualitative
		Effective A/S	Qualitative

Strategic benefits mean enhanced competitive advantages. According to Porter (1979)'s five competitive forces model, there are five threats such as the threat of new

entrants, the power of suppliers, the threat of substitute products, and the rivalry among existing competitors [35]. We classified strategic benefit factors into the reduced threat of rivalry, enhanced supplier relationship, and enhanced customer relationship. It is described in table 6.

6 A Case Study

We performed a case study to show the practical values of this methodology. We implemented the contents security system in ABC Co., Ltd. The features of ABC Co. Ltd. follows: 1000 employees, located in Seoul, public service provider. We monitored traffics with measuring systems during five days before implementing the contents security system. The Internet traffic of one business day is about 23.3Gbyte. The HTTP traffics of one business day is about 15.8Gbyte. The average byte of HTTP session is 14.4Kbyte. The non-business HTTP traffics of one business day is about 6.6Gbyte. The rate of non-business HTTP traffics over total Internet traffics is 27.9%. We monitored traffics with measuring systems during five days after implementing the contents security system. The key characteristics of Internet traffics are as following: number of daily HTTP sessions is 649,966 sessions, number of daily blocked sessions is 4,574 sessions. 447,538 sessions are reduced after implementing the contents security system. 6.4Gbyte of traffics is reduced (= 447,538sessions * 14.4K/sessions). The rate of reduced Internet traffics "after implementation" over "before implementation" is 27.5%.

Table 7. Cost-benefit factors

Factor			Value
Cost	Initial Invest	Logical	Hardware/vendor price = 50 ,000$ (2 set) Software/vendor price = 40 ,000$
		Admin.	Staffing(IT dept. staffs of ABC Co. Ltd.): 25,000$
	On-going Invest.	Logical	Upgrade (S/W & URL DB) = 800$/month Replacement (H/W) = 50 ,000$ (after 5 years)
		Admin.	Maintenance price = 800 $/month (after 1yr) Staffing: 5,000$/month,man*0.5man = 2,500$/month
Benefit	Human Resource	Prevention of work loss	Employees surfing time = 1.5 hours/ day Employees salary per minute = 0.25 $/min per user Reduction of wasted surfing: 6.4Gbyte/15.8Gbyte*90min/day=36.5 min/day Saved cost: 36.5min/day*0.25$/min,user*1000users=9,125 $/day
	Revenue	Saving of direct loss	Withdrawal of investment plan on E1 line: according to ISP price = 2,000$/month

With factors of table 7, we calculated pay-back periods. ABC Co., Ltd. invested initial cost of 11 5,000$ (= 50,000$ + 40,000$ + 25,000$) and needs to invest monthly on-going cost of 3,300$ in first year and 4,100$ after first year. ABC Co.,

Ltd. can save monthly cost of 184,500$ (= 9,125$/day * 20day/month + 2,000$/month). So, its pay-back period is shorter than one month. Regarding the term of validity of implemented systems, we calculated total investment and cost savings during five years. Total investment during five years is 351, 400$ (= 115,000 $ + 3,300$/month * 12month + 4, 100$/month * 48month). Total cost savings during five years is 11,070,000$(= 184,500$/month * 60month).

7 Conclusion

Information security as a discipline was first studied in the early 1970s, although the issues had influenced the development of many earlier systems such as the Atlas system and MULTICS (NIST). Previous researches focus on technical issues like cryptography, networking, protocol, programming, and evaluation of these technology and products.

This paper answers on difficult problems that organizations face in business environments when they invest in information security systems by suggesting the economic justification methodology. The methodology provided in this paper includes a process model and analysis criteria of cost factors and benefit factors. Limitation and further research issues are summarized as following: 1) Monetary quantification methods of numerical and qualitative benefit factors should be developed. 2) Case studies should be executed, analyzed, and combined into a repository to provide various practical use cases.

With this methodology, organizations can manage their security investments to take effects of a cost saving, added profitability, enhanced process, enhanced competitiveness, and enhanced outer relationship.

References

1. Alpar, P. and Kim, M.A.: A Microeconomic Approach to the Measurement of Information Technology Value. Journal of Management Information Systems **7** (1990)
2. Barua, A. et al.: Information Technologies and Business Value: An Analytic and Empirical Investigation. Information Systems Research **6** (1995)
3. Brynjolfsson, E. and Hitt, L.: Paradox Lost? Firm-level Evidence on the Returns to Information Systems. Management Science **42** (1996)
4. Mahmood, M.A. and Mann, G.J.: Measuring the Organizational Impact of Information Technology Investment: An Exploratory Study. Journal of Management Information Systems **10** (1993)
5. Mitra, S. and Chaya, A.K.: Analyzing Cost-effectiveness of Organizations: the Impact of Information Technology Spending, Journal of Management Information Systems **13** (1996)
6. Rai, A. et al.: Technology Investment and Business Performance. Communications of the ACM **40** (1997)
7. Eloff, J.H.P. et al.: A Comparative Framework for Risk Analysis Methods. Computers & Security **12** (1993)
8. Geer, D.E.: Making Choices to Show ROI. Secure Business Quarterly **1** (2001)
9. Checkland, P.: Systems Thinking, Systems Practice. John Wiley & Sons (1981)

10. Delone, W.H. and McLean, E.R.: Information Systems Success: The Quest for the Dependent Variable. Information Systems Research (1992)
11. Mahmood, M.A.: A Comprehensive Model for Measuring the Potential Impact of Information Technology on Organizational Strategic Variables. Decision Sciences **22** (1991)
12. Palvia, P.C.: Developing a Model of the Global and Strategic Impact of Information Technology. Information & Management **32** (1997)
13. Grover, V. et al.: Information Systems Effectiveness: The Construct Space and Patterns of Application. Information & Management **31** (1996)
14. Saarinen, T. and Scheer, A.W.: Business Process Engineering. Springer-Verlag (1994)
15. Torkzadeh, G. and Doll, W.J.: The Development of a Tool for Measuring the Perceived Impact of Information Technology on Work. Omega (1999)
16. Renkema, T.J.W. and Berghout, E.W.: Methodologies for Information Systems Investment Evaluation at the Proposal Stage: A Comparative Review. Information and Software Technology **39** (1997)
17. Remenyi, D. et al.: Effective Measurement and Management of IT Costs and Benefits. Butterworth-Heinemann (2000)
18. Jiang, J.J. and Klein, G.: Information System Project-selection Criteria Variations within Strategic Classes. IEEE Transactions on Engineering Management **46** (1999)
19. Bacon, C.J.: The Use of Decision Criteria in Selecting Information Systems / Technology Investments. MIS Quarterly, September (1992)
20. Scott, D.: Security Investment Justification and Success Factors. Gartner (1998)
21. Bates, R.J.: Disaster Recovery Planning. McGraw-Hill (1991)
22. Power, R.: CSI/FBI Computer Crime and Security Survey. Computer Security Issues & Trends **18** (2002)
23. Blakley, B.: Returns on Security Investment: An Imprecise but Necessary Calculation, Secure Business Quarterly **1** (2001)
24. Witty, R. et al.: The Price of Information Security, Strategic Analysis Report. Gartner (2001)
25. Malik, W.: A Security Funding Strategy. Gartner (2001)
26. Kim, S.: A Study on Enterprise Information System Investment Evaluation, Master thesis. Yonsei University (2000)
27. Harris, S.: CISSP All-in-One Exam Guide. McGraw-Hill (2001)
28. Roper, C.A.: Risk Management for Security Professionals. Butterworth Heinemann (1999)
29. Fites, P.E. et al.: Controls and Security of Computer Information Systems. Computer Science Press (1989)
30. Hutt, A.E., Management's Roles in Computer Security, in Computer Security Handbook. Macmillan Publishing Company (1988)
31. Vallabhaneni, R.: CISSP Examination Textbooks. SRV Professional Publications (2000)
32. Krutz, R.L. and Vines, R.D.: The CISSP Prep Guide: Mastering the Ten Domains of Computer Security. John Wiley & Sons (2001)
33. Schweitzer, J.A.: Protecting Information in the Electronic Workplace: A Guide for Managers. Reston Publishing Company (1983)
34. Scott, D. and Malik, W.: Best Practices in Business Continuity Planning. Symposium/ITxpo 2001 (2001)
35. Porter, M.E.: How Competitive Forces Shape Strategy, Harvard Business Review **57** (1979)

A Modeling Framework of Business Transactions for Enterprise Integration

Minsoo Kim[1], Dongsoo Kim[2,†], Yong Gu Ji[3], and Hoontae Kim[4]

[1] Department of System Management and Engineering, Pukyong National University, San 100, Yongdang-dong, Nam-gu, Busan, Korea
minsky@pknu.ac.kr
[2] Graduate School of Healthcare Management and Policy, The Catholic University of Korea, 505 Banpo-dong, Seocho-gu, Seoul, Korea
dskim@catholic.ac.kr
[3] Department of Industrial and Information System Engineering, Soongsil University, 1-1 Sangdo 5-Dong, Dongjak-Gu, Seoul, Korea
gilbreth@ssu.ac.kr
[4] Department of Industrial and Systems Engineering, Daejin University, San 11-1, Sundan-dong, Poch'on-si, Kyonggi-do, Korea
hoontae@daejin.ac.kr

Abstract. A modeling framework and an integration platform for unifying B2Bi (Business to Business integration) and EAI (Enterprise Application Integration) systems are required to effectively manage collaborative B2B business processes. In this research, we suggest a modeling framework for enterprise integration. The suggested framework includes various activity controllers derived from several business process modeling standards. An extension mechanism is included in the suggested framework for incorporating future changes in those standards or system's operational requirements. The modeling framework has been implemented as a real business integration platform and runs successfully in production systems to execute enterprise B2B collaborations.

1 Introduction

There are various classification schemes for enterprise collaboration. According to the level of information technology involved, enterprise collaboration can be classified into 3 key types: supply chain, extended enterprise and virtual enterprise. And infinite variations in the range and scope of collaborations can exist in those three types. Tight collaboration of supply chain is characterized as long-term relationship and information sharing between their business systems. To the contrary, collaboration in virtual enterprises is so loose and the relationships are temporal and dynamic [4].

There have been lots of efforts in the area of EAI (Enterprise Application Integration) and B2Bi (Business to Business integration) to support B2B collaborations. EAI systems have been developed to meet individual business needs in a proprietary man-

† Corresponding author. Tel.:+82-2-590-2484; fax.+82-2-590-1393

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1249 – 1258, 2005.

ner by solution providers; however, B2Bi should reflect various emerging e-Business standards. Therefore, separate systems for B2Bi and EAI are used for executing public and private processes respectively. As there is an increasing need for the integration of the two systems, companies are requiring an integration method that satisfies the requirements of both systems. In this research, we have designed a unified modeling framework for integrating B2Bi and EAI processes and implemented a business integration platform to support various types of business collaborations.

This paper is composed of the following chapters. Chapter 2 introduces and analyzes business process modeling standards. Chapter 3 defines activity controllers and service elements used for executing business processes. Chapter 4 describes the procedure for defining business transactions in our modeling framework. Chapter 5 presents a business integration platform that executes those business transactions. Finally, the conclusion of this research is given in chapter 6.

2 Standards for B2B Collaboration Modeling

Numerous modeling frameworks and languages have been proposed to describe business processes in various researches. Recently, there have been several efforts to define B2B collaborations using Web Services. Among them, we have focused three standards and they are BPEL4WS (Business Process Execution Language for Web Services), BPML (Business Process Markup Language) and WSCI (Web Services Choreography Interface). We reviewed all components of those standards and extracted common modeling constructs to define a unified modeling framework.

WSCI has been released for dynamic Web Services configuration. Web Services describes input/output arguments and invocation method via WSDL (Web Services Description Language) and thus its services are restricted only to static applications. WSCI, however, supports specification of dynamic services that are described by combing WSDLs of various Web Services to support complicated interactions [1,5,6].

BPML and BPEL4WS are process definition languages for describing business logics of distributed Web Services. Both provide abstract model and syntax for defining general business activities and can be used for modeling of business processes, B2B collaborations and choreography among complex Web Services [2,3,7].

2.1 Simple Activity

Two types of activities, what are called simple and complex, are used for the configuration of business processes. Simple activity is atomic and thus indecomposable, while complex activity is composed of other activities. Business choreography languages can be related to WSDL 'operations' via simple activities. In those standards, business processes among partners are defined using logical constructs such as 'message', 'portType' and 'operation'. Descriptions of physical protocol binding and specific service types are also delegated to WSDL. <Table 1> shows simple activities defined in those B2Bi standards.

Table 1. Simple activities of B2Bi standards

standard	list of simple activity
BPEL4WS	receive, reply, invoke, assign, throw, terminate, wait, empty, pick, compensate
BPML	action, assign, call, compensate, delay, empty, fault, raise, spawn, synch
WSCI	action, delay, empty, fault, call, spawn, join

2.2 Complex Activity

BPML and WSCI define a business process in a block structure. A block that can recursively contain other blocks is an important structure for declaration, definition and execution of business process. BPML and WSCI present seven complex activities for controlling flows between those blocks. In BPEL4WS, parallel blocks are designed using 'flow' activity. Description of interdependencies between inner activities is also supported in a directed graph format. There are five complex activities in BPEL4WS including 'scope' activity, which is used for sharing context information between blocks. <Table 2> summarizes complex activities of B2Bi standards.

Table 2. Complex activities of B2Bi standards

Classification	BPML/WSCI	BPEL4WS
parallel structure	all	Flow
serial structure	sequence	Sequence
split structure	choice, switch	switch
repeatable structure	until, while, foreach	while
others		scope

2.3 Business Process Initiation

In BPML and WSCI, a new business process can be started by using activities such as 'call', 'compensate' and 'spawn' or by receiving messages from external entities. Process initiation by activities is frequently used to start a nested sub-process from the parent process. And process initiation by receiving input message is usually performed in one-way or request-response communication with external company. In BPEL4WS, a new business process can be initiated only by message exchange. The 'receive' activity can initiate a business process after receiving a message and the 'invoke' activity can call external Web Services by sending message.

2.4 Exception Handling and Transaction Control

A business process often runs long and may exchange asynchronous messages. Therefore, ACID transaction properties should be relaxed considering the difficulties of reliability assurance and long-lasting locking or isolation. Consequently, compensation could be used for exception handling of business process. Compensation is a kind of activity for undoing the effects of already finished activities in abandoned process.

BPEL4WS and WSCI also support LRTs(Long-Running Transactions) and nested transactions. They also have exception handling and compensation mechanism. The 'throw' activity casts out a failure when an exception occurs and the failure is compensated by 'compensate' activity. When a transaction is cancelled, all compensation activities in the same execution context are executed backward. BPEL4WS does not specify transaction property directly in its specification, and it has a complementary standard called WS-Transaction. Technical details of transaction properties are defined in WS-Transaction [3].

BPML also provides 'exception' and 'compensation' processes, which can be executed when exception handling and compensation are needed with their own execution context, event, and component activities.

3 Unified Modeling Framework

As is described above, each standard has similar activities with little differences. New standard may emerge with new activities afterward. Therefore it is important to support existing B2Bi standards and also has the flexibility and extensibility for additional activities. This can be helpful to block the impact of B2B changes on the backend systems. To satisfy these requirements, our unified modeling framework and business integration platform provide various activity controllers and system services.

3.1 System Activity Controller

Our unified modeling framework has 17 system activity controllers, which can be categorized into following 6 types.

***(a) Flow Controller*:** 'Sequence', 'Escape', 'If' and 'Switch' controllers are included in this type. 'Sequence' controls serial execution of process blocks and corresponds to complex activity of 'sequence' in the above B2Bi standards. 'Escape' is used to finish the whole process or escape from loop controllers such as 'While' or 'Repeat'. When 'Escape' is used to finish the whole process, the status of a finished process can be specified as a success or a failure. 'Escape' is similar to 'throw' of BPEL4WS, 'raise' of BPML and 'fault' of WSCI. 'If' is used for selecting conditional routes from true and false blocks. 'Switch' provides many conditional branches and the evaluated value is used in selecting correct case among those branches. 'If' and 'Switch' are similar to the complex activities such as 'choice' and 'switch' of the B2Bi standards.

***(b) Loop Controller*:** 'While', 'Repeat' and 'Foreach' controllers are included in this type. 'While' controller is used to repeat enclosed services or controllers for a given condition is met. 'Repeat' controller is used to iterate fixed number of times that are specified by other variable or constant. These controllers correspond to the complex activities of 'until' and 'while' in the B2Bi standards. 'Foreach' is a controller for handling data structure of the same topology that repeatedly occurs in the input data. It corresponds to 'foreach' complex activity of other B2Bi standards. This controller can be very useful for handling XML elements of the same XPath in the XML business

documents. Since a concept similar to 'scope' complex activity of BPEL4WS is supported within the 'Foreach' block, enclosed services or controllers can share accumulated results of previous iterations and access current iteration element.

***(c) Transaction Controller*:** 'StartTX' and 'EndTX' are controllers for specifying transaction boundary. All the transaction services that are demarcated by these controllers are combined as a global transaction and controlled by the same transaction manager. 'CheckPoint' controller stores intermediate result of a transaction thus can be used only within 'StartTX' and 'EndTX'. When an unexpected system failure occurs during the transaction execution, it is possible to resume that transaction to the point where the last 'CheckPoint' is executed, instead of restarting the whole transaction.

***(d) Concurrency Controller*:** 'Split' and 'Merge' controllers execute services in parallel. They correspond to the complex activities such as 'all' and 'flow' of the B2Bi standards. In our business integration platform, 'Split' and 'Merge' should be used in pairs with the same number of parallel branches, and they can be nested in any depth.

***(e) Nested Process Controller*:** 'Sub-Process' controller synchronously spawns a nested process from a parent process and waits until the child process finishes and returns the result. 'Chained-Process' controller is used for invoking a nested process asynchronously. Parent process continues its subsequent activities without waiting the result of the child process. These two controllers correspond to the simple activities such as 'spawn', 'synch', and 'join' controllers of the B2Bi standards.

***(f) Timing and Exception Handling Controller*:** All services defined in our business integration platform can have 'Timeout' and 'Exception' controllers. 'Timeout' controller specifies the time limit of a successful transaction completion. 'Exception' controller executes an exception handling process that does the compensation or cleanup when a failure occurs. 'Wait' controller blocks the progress of process for a specified time period. These controllers are similar to the activities such as 'compensate', 'fault', and 'wait' of B2Bi standards.

3.2 User-Defined Activity Controller

In our business integration platform, you can define a new type of controller called user-defined activity controller by combining basic services and other activity controllers. You can easily support new activities of emerging B2B standards by this extension mechanism. This enables B2B collaboration to operate without adopting a new application system or solution.

For instance, 'N-out-of-M' and 'XOR' controllers can be defined as user-defined activity controllers. Figure 1 illustrates the composition of '2-out-of-3' controller by combining the system activity controllers and controller state matrix.

A common storage space is required in defining a user-defined activity controller to manage the status information of involved controllers. For this purpose, the integration platform proposed in this work provides system services such as status management service and runtime object storage.

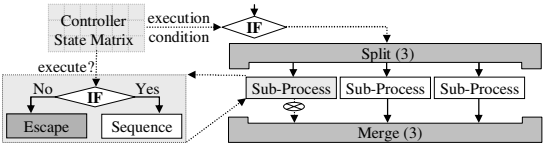


Fig. 1. Composition of '2-out-of-3' Controller

3.3 System Service

Status management service provides state information of controllers and business process instances. State information is stored in a state matrix, where you can conditionally select next state among the candidate states that are movable from current state. The execution results of other processes can be used as transition condition of the state matrix. Controller state matrix and conversation state matrix are used for storing the status of activity controllers and business process instances respectively.

Conversation state matrix enables state transition of business process by exchanging business messages or signals between business entities. This matrix is used for keeping the overall execution status of business conversations performed in the choreography pattern of one-action or two-action. While state matrix provides status information of activity controllers and business process instances, runtime object storage preserves real data or document objects exchanged. Once an XML document is delivered and parsed as a memory object, subsequent processing to that object can be performed on the runtime object storage via object reference.

4 Business Transaction Modeling

In this section, the procedure for defining business processes using our modeling framework is described. The whole business processes designed to perform B2B collaboration should be redefined to reflect the role of participating company. Individual company's process model redefined to meet partner's role and information system capability is called business transaction. The procedure for business transaction modeling consists of the following 6 steps.

4.1 Decomposition of Business Process

A business process can be composed of complex activities occurring across companies. Therefore, participating companies should decompose the process and redefine their own processes according to their roles in the B2B collaboration. At this time, the business transaction to be executed in each company can be identified. The entire input and output of the business process are usually in the format of XML business documents defined by the B2B standards. The business transaction defined by decomposing the business process thus creates some partial sections of the XML business documents. This means that we can also define input and output of business transaction by decomposing the input and output of the business process.

4.2 Design of Business Transaction

Once input and output are defined for business transaction in XML format, it is required to define how to create or transform those XML documents. The 17 system activity controllers are used in this step. The procedure for designing the business transaction is very similar to that of programming using conventional programming languages. If designing a business transaction with given 17 activity controllers is somewhat complex, then you can divide it with several sub level business transactions or introduce a new user-defined activity controller or even code the whole transaction logic as a user-defined service with programming language. User-defined service is a service module that is developed using conventional programming language and executable on our integration platform.

4.3 Design of Sub Business Transaction

Sub business transaction can be designed similarly like the parent business transaction. When we invoke a sub business transaction, it should be specified whether the invocation is asynchronous or synchronous at the parent business transaction.

4.4 Design of User-Defined Controller

If it is difficult to design business transaction only with system activity controllers and services, you can define and use new controllers by combining existing controllers and services. When designing a new controller, input variables to capture the state information of that controller are introduced and mapped to the XML input of business transaction. If the controller needs conditional decision logic, you can define that condition with XPath statement that returns boolean value.

4.5 User-Defined Service

If there is a difficulty in designing business transaction by a user-defined activity controller or dividing it into several sub business transactions due to the complicated functionality or complex logic, then you can model it as a user-defined service that is coded with conventional programming language. User-defined service may not be recommendable because it hides business logic in source codes. However, it can be used as a way of avoiding performance degradation that can be caused by the interpretation overhead of complex business transaction.

4.6 Debugging and Deployment

The final implementation logic of business transactions should be verified at the level of the whole business process. First of all, debugging task for verifying the input and output values of individual business transaction should be performed by participating enterprises. Once all the business transactions are verified, they are deployed as Web Services for testing. After all business transactions are deployed, the company who takes the role of initiator calls the starting Web Service and tests the entire process. In this step, the whole process including choreography and messaging can be verified. If

a certain business transaction fails, then the company in charge of that failed transaction is requested to modify the business transaction.

5 Implementation of Business Integration Platform

5.1 System Architecture of Business Integration Platform

The whole architecture of business integration platform implemented in this research is presented in Figure 2. Communications via well-known protocols such as HTTP(S), FTP and SMTP are supported by corresponding communication protocol adapters. Specific application adapters are also provided to support legacy protocols and applications such as JDBC connection of DBMS, RFC protocol of SAP/R3. These adapters take the role of communication with external business entities by exchanging business documents. When a new business document arrives to an adapter, an instance of corresponding business transaction is created and dispatched to the execution engine.

Execution engine refers to the state matrix of corresponding business transaction model. It creates a new state log and updates that state log to control the execution of runtime business transaction instance. Execution engine is equipped with the functionalities of condition evaluation and data transformation. Debug engine is provided to verify business transaction before it is deployed.

Execution engine basically operates on a stack data structure to handle a business transaction and the stack can be extended to a tree of stacks to support 'Split' and 'Merge' controllers. That is to say, when execution engine encounters a 'Split' controller in the business transaction, that transaction's current stack becomes the parent of child stacks newly created as many as the number of parallel branches of that 'Split'. Each child stack is concurrently executed by separate thread and synchronized at the 'Merge' controller. Execution engine advances the processing of business transaction by selecting and executing popped node at the leaf stacks of that tree.

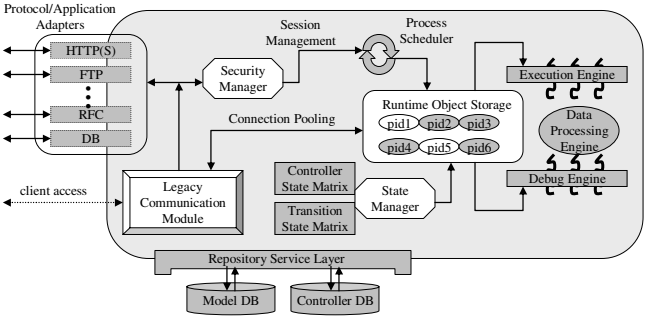


Fig. 2. Architecture of Business Integration Platform

Repository service layer coordinates and controls access to underlying databases that store business transaction model, definition of activity controller and status information of runtime instance.

5.2 Activity Controller Handling

Controllers and business transaction models of our business integration platform are stored at the process instance table, state matrix and runtime object storage, all of which tightly interact to advance transaction processing and data transformation of input business document. Business message issued by external business entity is detected by communication protocol adapters of the business integration platform and thus creates a new transaction instance with its input business document.

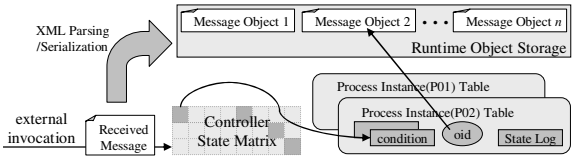


Fig. 3. Execution of business transaction

Business documents are stored in the database through repository service layer and at the same time they are also delivered to the runtime object storage. Storing in database is needed for non-repudiation of original document. In addition, archiving prevents data loss from unexpected system failures.

Created business transaction instance is stored at the process instance table together with its unique process ID and runtime object reference. And the execution of the transaction instance advances by updating state transition matrix and controller state matrix. All of above operations are recorded in the status log.

Evaluating conditional statements, which are defined in the transaction model, is needed for the transition of the status of the state matrix. Spawning of new transaction instance is also possible by evaluating those conditional statements. All of those conditions are maintained at the process instance table. Figure 3 illustrates the processing scheme of business transaction with various system services.

6 Conclusions

To implement enterprise collaboration system effectively, a systematic reconstruction between pre-existing B2Bbi and EAI systems is required. To support the systematic reconstruction, it is needed to provide unified modeling framework that integrates functional requirements of B2Bi and EAI systems from the beginning. By doing this, the impact of B2Bi processes that frequently fluctuate in the enterprise information system can be minimized or blocked against the backend EAI processes.

The unified modeling framework that is presented in this research can keep the backend legacy system operational from the repeated change of B2Bi standards and stay away from reintegrating backend system iteratively. Our business integration platform supports web services, JMS messaging, global transaction and RosettaNet transaction as well as various system services and integration methods [8]. Further enhancement to this platform and its modeling concept is planned as subsequent research issues and its production deployment is expected in the nearby future.

References

[1] A. Arkin et.al., Web Services Choreography Interface 1.0, <http://www.w3.org/TR/wsci/> (8 Aug. 2002).
[2] Assaf Arkin, Business Process Modeling Language, <http://www.bpmi.org> (13 Nov. 2002).
[3] F. Cabrera et al., Web Services Transaction, <http://www.ibm.com/developerworks/library/ws-transpec/> (9 Aug. 2002).
[4] Jagdev, H. S., and Thoben, K.-D., Anatomy of enterprise collaborations, Production Planning and Control, 12 (5), 437-451, 2001.
[5] W3C Note: WSDL 1.1: Web Services Description Language (WSDL) 1.1. World Wide Web Consortium. http://www.w3.org/TR/wsdl.html (2001)
[6] R. Lay, J2EE Platform Web Services, Prentice Hall, 2003.
[7] T. Andrews et al., BPEL for Web Services 1.1, <http://www.ibm.com/developerworks/library/ws-bpel/> (5 May 2003).
[8] RosettaNet Implementation Framework: Core Spec. V02.00, <http://www.rosettanet.org> (13 Jul. 2001).

Process-Oriented Development of Job Manual System

Seung-Hyun Rhee[1], Hoseong Song[1], Hyung Jun Won[1], Jaeyoung Ju[1], Minsoo Kim[2], and Hyerim Bae[3,†]

[1] Handysoft Corp., 1708-2, Seocho Dong, Seocho Gu, Seoul, Korea
{toy7sam, hodorii, hjwon, jyju}@handysoft.co.kr
http://www.handysoft.co.kr
[2] National Computerization Agency, NCA Bldg, 77, Mugyo_Dong, Jung-Gu, Seoul, Korea
kms@nca.or.kr
http://www.nca.or.kr
[3] Pusan National University, San 30, Jangjeon Dong, Guemjoeng Gu, Pusan, Korea
hrbae@pusan.ac.kr
http://www.ie.pusan.ac.kr

Abstract. Today's business environment involves increasingly complex and uncertain processes. To maintain or increase their competitiveness, businesses look for new methodologies to permit them to manage their processes more effectively. One of these methodologies, Business Process Management (BPM), which includes the workflow concept, has gained some attention. The process-centric integration of BPM has demonstrated the capability to more effectively manage business processes; however it is difficult to integrate systematically using existing IT development paradigms. Therefore, a new IT development paradigm is required, where business processes are managed independently and automated with calling relevant business applications. This new IT paradigm is called the '*Process Orientation*' (PO) concept, which focuses on independence of the business processes and the BPM system. In this paper, we explain a new kind of job manual system as a business application under the PO concept. In the new job manual system, the BPM system becomes a development platform to implement various applications. Contrary to existing job manual systems, it is possible for this new system to directly transfer a manual prepared in it to a BPM process definition. The new system also enables more efficient management of business processes and helps users of both systems perform relevant tasks more easily.

1 Introduction

Over the past 20 years, many businesses have invested heavily in their IT systems. However, IT has been viewed only as an ancillary element to support various business activities rather than as a strategic element to create business value [9]. To be more flexible and agile in an ever more competitive marketplace [3], top global companies, including General Electric, have examined changing their IT strategy from skeptical data-centric IT to process-centric IT. Process-centric IT focuses on utilizing IT to

† Corresponding author.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1259 – 1268, 2005.

maximize strategic business value by integrating existing systems and managing them effectively rather than developing new systems [9, 10]. An effective methodology to execute and implement this new strategy is gaining importance in the business community.

Process-centric IT regards business processes as independent elements in the BPM system, like database in a Data Base Management System (DBMS). This enables the development of business applications based on a new architecture that manages the processes separately. The separation of the processes is essential for a system to achieve process-centric IT. Business Process Management (BPM) is a software system to define, execute, monitor, analyze, and improve dispersed business processes that has been recognized as a solution for industries.[6] In order to introduce and implement BPM, a company has to conceive the processes as independent objects that can be defined, executed, and managed apart from applications [8]. A new information system architecture is considered that utilizes BPM as a platform in developing applications, and we introduce a process-oriented 'Job Manual System' (JMS) developed under that architecture. In this new system, users write a job manual by modeling a business process corresponding to the manual. Thus, each manual generated in our system includes one business process in it. The features of the system support simplified clarification of business structures and processes and transfer of manuals to the BPM process definition in an executable form.

2 Background

2.1 Process Orientation

Since the Relational Database Management System (RDBMS) appeared in the 1970s, most business application systems (solutions) have been developed utilizing the DBMS. The DBMS can separate data from applications, and the applications can deliver relevant data to the DBMS to store, retrieve, and update. As a result, the DBMS enables business applications to become more flexible and less dependent on changes of the data [9]. After the appearance of the Workflow Management System (WFMS) [5, 13] in the 1990s, business processes could also be separated from applications like database, as shown in Fig. 1. Recently, as the WFMS concept has expanded into the BPM concept, the BPM system has become the technological basis for the development of new kinds of applications through the independence of business processes.

As known well, BPM originally automates, executes, and improves business processes [7, 11]. The business processes are directly related to business activities, and act as an agent to invoke various applications such as SCM and CRM. With regard to these properties of the business processes, the BPM system plays an essential role as a 'Business Operating System'. In other words, the BPM system not only can manage business processes systematically while maintaining original functionality, but also becomes a solution platform in order to develop new applications, enabling the solution to be flexible to changes of processes. In this paper, the development framework of the solution to develop new applications is called '*Process Orientation*' (PO).

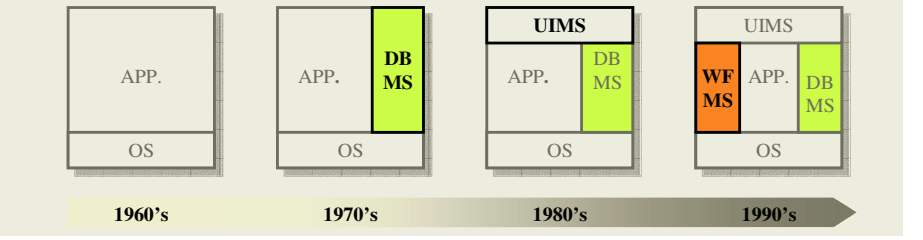


Fig. 1. Process Independence

Business solutions under the PO concept can be classified into two types. First, there are packaged solutions, modeled on BPM, to execute and manage specific processes efficiently. For example, the 'SOXA Accelerator' on which the Sarbanes-Oxley Act [6], the law enacted in the USA for accounting reform, is modeled as business processes on the BPM. Because it is very important to comply with the Act, it is required that it is led to observe the Act by the characteristics of BPM such as execution and control of the processes. Second, other solutions are developed on the BPM platform and expected to interact with the platform. These solutions include 'KMS based on BPM', 'the performance management system based on BPM', and JMS introduced in this paper.

2.2 Role of BPM

BPM is an integrated methodology to design, manage and improve business processes, in order to enhance productivity, or software systems to support the methodology [6, 7, 11]. BPM comprehends automation of business processes (workflows) and

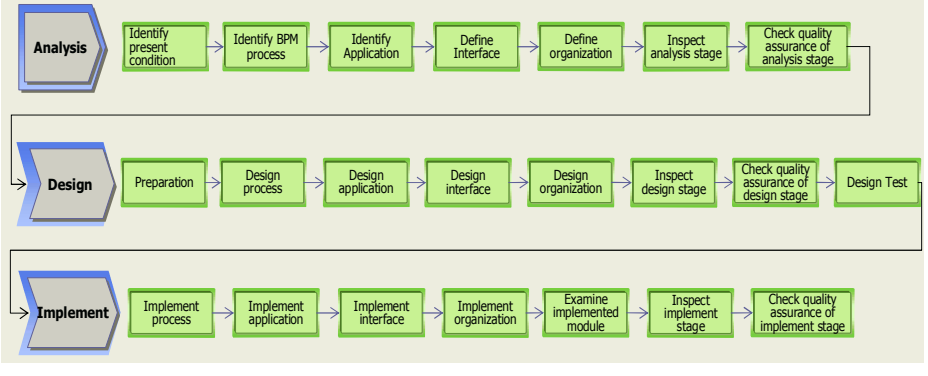


Fig. 2. The Procedure of BPM Implementation

integration of business applications (EAI, Enterprise Application Integration) [2, 5, 8]. While it designs, executes, monitors and analyzes the processes [2, 9], it calls or integrates external systems [6, 7]. BPM products of worldwide vendors including

Handysoft Global, FileNet and Staffware have already been utilized in various industries and are being introduced to an increasing number of companies [4].

When a company tries to introduce the BPM system, it generally follows the series of procedures presented in Fig. 2. First, it identifies business processes to be managed by BPM [8]. It then defines applications and organizations relevant to those business processes. Once relevant information has been identified, it is designed concretely and implemented to be executed in the BPM system [8]. Significant time and manpower can be expended in the 'Analysis' stage of the procedure below if the processes have not previously been managed or defined systematically.

2.3 Job Manual System (JMS)

JMS is a software system to frame manuals including procedures and related details of various tasks and manage them [2]. The system can provide information relevant to specific activities for task performers who are not proficient in those activities. In addition, when a task performer is changed, it can simplify a process to tutor the new performer. Finally, because users utilize know-how accumulated in the system, their abilities and performance can become significantly enhanced. This helps the users to perform their jobs more effectively, and with increased productivity.

Generally in a traditional job manual system, a series of business procedures is implicit in each job manual to which users refer. The manual and the procedures are ordinarily expressed in a simple text format, so that it is not only inconvenient for the users to read and utilize the manuals but also difficult to understand the overall work flow [2]. In addition, the systems do not support interfaces with other applications required to perform the jobs. If the business procedures in a job manual can be represented and managed in a BPM-understandable form, those business procedures can operate efficiently and can effectively integrate other systems relevant to the procedures. Therefore, the 'job manual system based on BPM' that is developed on the BPM platform can correct the inadequacies of existing systems. That system provides unique UI to enable users to draw up a job manual related to processes and contributes to promoting business processes through interaction with the BPM system.

3 Development of JMS Based on PO Concept

This chapter explains how the new JMS is designed and implemented on the BPM platform and how it is utilized and interacts with BPM. Our system can be a typical one developed in accordance with the PO concept.

3.1 Job Manual and Process Definition of BPM

In our system, each manual is considered to be related to a single process and is managed by the BPM system. In other words, preparation of a manual includes modeling a business process corresponding to the manual. Unlike traditional systems, a manual is expressed in a graphical form, so that the users can enhance the readability of the manual and the understandability of the tasks detailed in it [2]. To model the process in the system, we provide a design tool, 'Simple Designer', which is a simplified version of an original BPM's process designer. Simple Designer provides user interfaces in order to

manage information relevant to a process (name, description, deadline and attached file) and a task (name, participants, applications, deadline, and attached file). The information is modeled according to the structured DB schema in Fig. 3.

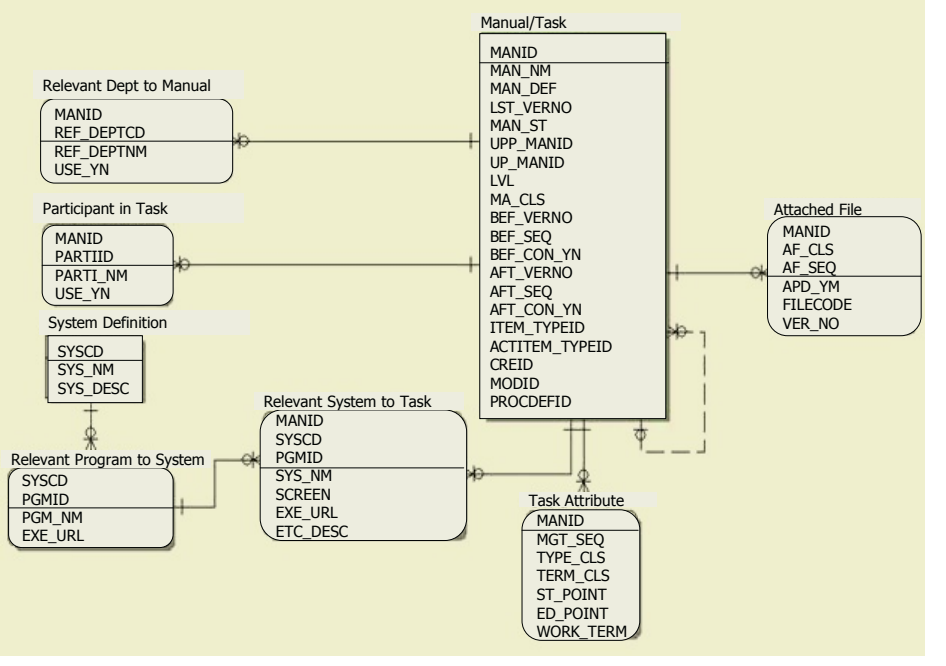


Fig. 3. DB Schema of main entities in the JMS

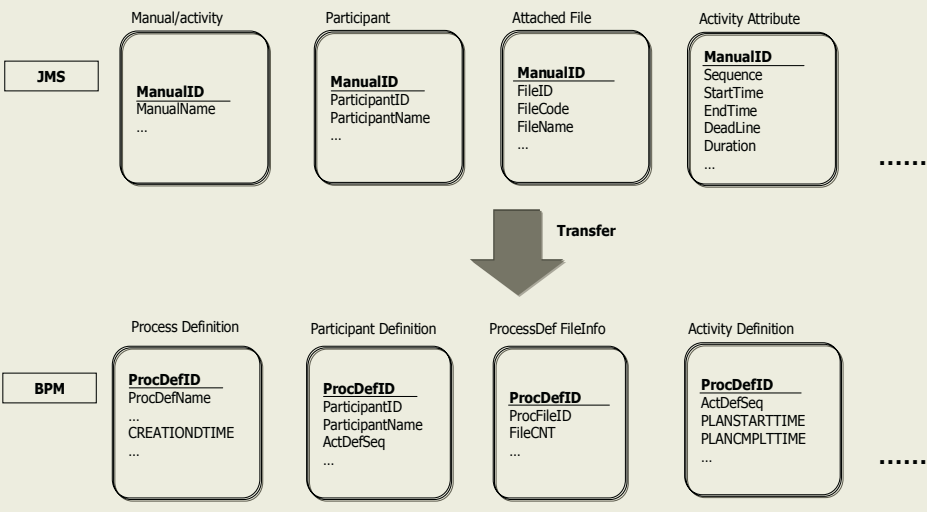


Fig. 4. Transferring data from the JMS to the BPM system

A manual with a customized process model can be utilized as a process definition in BPM through the structural property. In this sense, our system is an application developed on the basis of the BPM platform. In this system, the manual can be transferred directly to the BPM process definition and executed directly by the BPM engine. The transfer is enabled by the data structure between the JMS and the BPM system, which is designed to deliver directly relevant DB tables of the JMS such as process definition, application definition, participant, activity definition and attached file, to those of the BPM. The data structure is presented in Fig. 4. The figure illustrates which DB tables are utilized in the migration of common information between the manual and the process definition.

3.2 Interaction Between the JMS and the BPM System

As mentioned in Section 3.1, our system interoperates with BPM. The interoperability supports the system to interact with the BPM system in several ways. First, because the system transfers a manual to an executable process definition in the BPM platform, it contributes to simplifying identification, design, and implementation of the process definition using only simple steps. The system comprehends the analysis stages (process identification, application identification, organization identification), the design stage (process design), and the implementation stage (process implementation) in the procedure of BPM implementation. Therefore, if a company already utilizes this system, it can construct a BPM system rapidly.

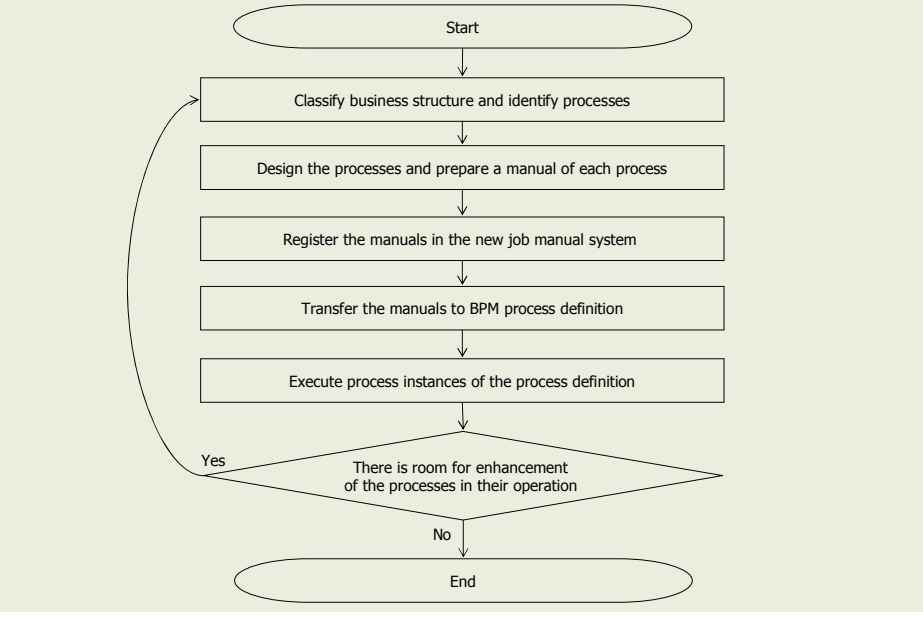


Fig. 5. The flow of continuous process improvement

Interoperability also enables continuous improvement of business processes, as shown in Fig. 5 First, business structure and business processes in a company are defined systematically. Then they are registered in the new JMS as manuals. The manuals, including the business processes, are transferred to the BPM system and executed in the platform. Enhancements of the operations of business processes are modified and updated into the manual system. As the following flow chart shows, continuous improvement of the processes can be carried out.

4 System Implementation

4.1 System Architecture

The JMS described in this paper is a typical software product designed and implemented under the PO concept. With development based on the BPM platform, the system has a unique architecture to utilize original properties of BPM. In addition to this system, other applications based on BPM are expected to be constructed on similar architecture. Therefore, general system architectures of the applications will most likely be designed as presented in Fig. 6.

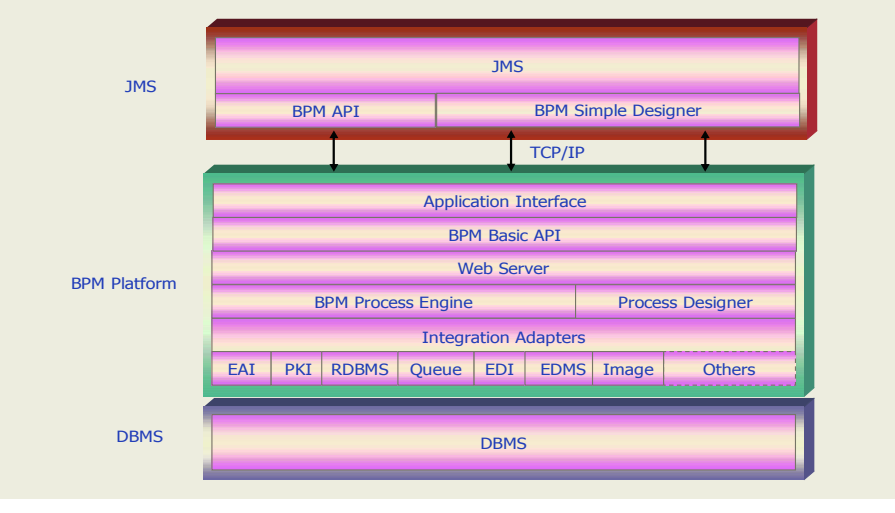


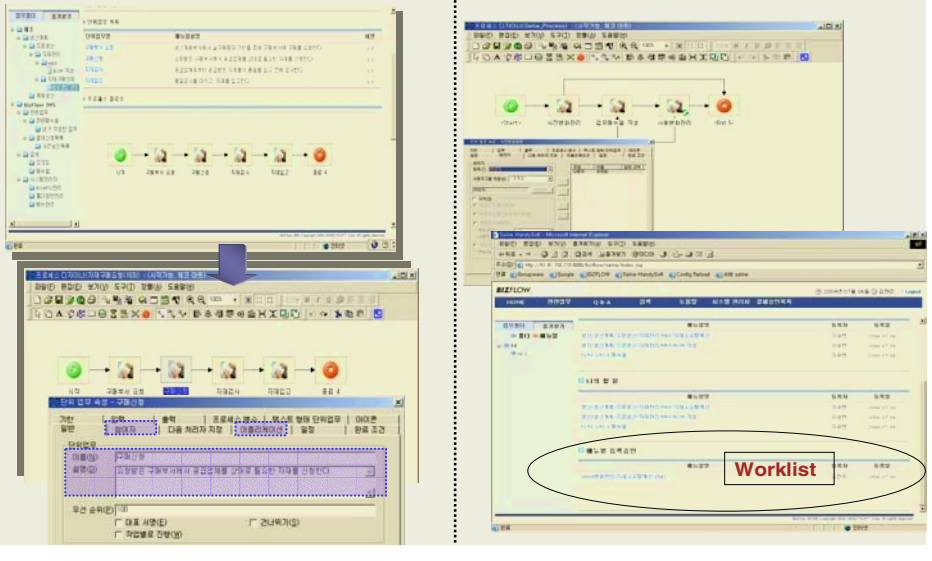
Fig. 6. System architecture of the JMS based on BPM

As shown in Fig. 6, the BPM system, which plays the role of a platform, manages process-relevant data in the DBMS and includes original functions, a process engine and a designer. It is equipped with a web server to handle Java Servlet and the BPM API, which starts the execution of a process or completes a task. The BPM system also provides several adapters to interoperate with back-end systems such as ERP (Enterprise Resource Planning) and integration systems such as EAI (Enterprise Application Integration) and EDMS (Electronic Document Management System) [8, 9].

This implies that the BPM platform plays a role as the Enterprise Nervous System proposed by Gartner [12]. Process-oriented applications including the JMS can interface with BPM. This interface can transport process related properties between the BPM system and the JMS through TCP/IP communication. The interface in this system delivers parts of the constituent information of Simple Designer to the DB tables of the specific properties in Process Designer to support the interoperability. Thereby, the system architecture enables the process-oriented development of applications based on BPM.

4.2 System Implementation

The JMS is implemented as a process-oriented application. In this section, the two main functions of the system are discussed. Figure 7 shows screen captures of the implemented system. Originated from properties of the BPM system, these functions demonstrate that the JMS is a solution based on BPM. The first function is the transfer of a manual to the BPM process definition, which is represented in Fig. 7 (a). A user designs a manual with Simple Designer and a processes included in the manual can be directly automated by the BPM engine. The second function is the registration of a manual by the BPM engine, which is shown in Fig. 7 (b).



(a) The transfer of a manual to BPM process definition

(b) Design and execution of a process in a manual

Fig. 7. Main functions of the system

Figure 7 (a) demonstrates, as previously mentioned, the transfer of a manual to the BPM process definition. The upper and lower sides of the figure show a manual in the system and the BPM process designer, respectively. The squares of dotted lines indi-

cate the common properties (task name, description, participants, and applications) of the manual and the process designer. Once the manual is transferred, a process definition is designed using the properties delivered from the manual, and then it is in an executable form.

The upper part of Fig. 7 (b) shows a manual registration process modeled with the BPM process designer. Applications and participants are suitably allocated to constituent tasks of the process. Once the process is executed, each user of the system can see a new manual to register or approve in his/her worklist, as shown in the lower part of the figure. If the process has to be changed, it is only necessary to modify the process model and develop relevant applications. As a result, the function discussed in Fig. 7 (b) increases the flexibility of the alteration of business processes. This is an important strength of general applications based on BPM.

5 Conclusions

In this paper, we proposed a new application development paradigm and developed the JMS as a sample application. We also discussed the technical issues of the system architecture and the utilization of the BPM platform. The 'Process Orientation' concept is a new framework to develop an application by separating business processes from existing applications. A 'process-oriented JMS based on BPM' is one of the new applications under the PO concept. The introduced system has a unique system architecture to interoperate with the BPM system. Once a manual, including a process, is prepared in the new JMS, the process can be transferred to the BPM process definition and executed by the BPM engine.

The JMS based on BPM has several advantages over existing systems. These mainly arise from the fact that the system observes the PO concept. First, a manual managed in our system can be viewed as a process to users. Therefore, it is possible to establish a business structure and clearly define dispersed business processes through the manuals. Second, a manual is not only utilized for references, but also has a direct connection to corresponding processes. As a result, the manual is transferred to the BPM process definition and is operated by the BPM engine with little additional work. Third, the system, by interacting with the BPM system, enables continuous improvement of business processes. Finally, companies that have not managed business processes systematically can arrange their business processes and clarify them using the JMS, and then can introduce the BPM easily.

Acknowledgements

This research was carried out at Handysoft Corporation which is a worldwide BPM vendor with 250 or more customers including GE, J&J and Samsung Electronics. It was supported by "Seine Project" for the product development in Handysoft Corp.

This work was also supported by "Research Center for Logistics Information Technology (LIT)" hosted by the Ministry of Education & Human Resources Development in Korea.

References

1. M. Gellevij, H. van der Meij, T. de Jong and J. Pieters, "The effects of screen captures in manuals: a textual and two visual manuals compared," IEEE Transactions on Professional Communication, Vol. 42, No. 2, 1999
2. D. Grigori, F. Casati, M. Castellanos, U. Dayal, M. Sayal and M.-C. Shan, "Business Process Intelligence," Computers in Industry, Vol. 53, No. 3, 2004.
3. M. Hammer, "Agenda: What Every Business Must Do to Dominate the Decade," Random House, 2001.
4. N. Latimer, "Business Process Management Preliminary Market Size and Forecast," SWSI-WW-MT-0123, Gartner, 2004.
5. S. –H. Rhee, H. Bae, Y. Kim, "A dispatching rule for efficient workflow," Concurrent Engineering – Research and Applications, Vol. 12, No. 4, 2004
6. E. Eugene Schultz, "Sarbanes-oxley – a huge boon to information security in the US," Computers & Security, Vol. 23, No. 5, 2004.
7. H. Smith, "Business process management—the third wave: business process modeling language (bpml) and its pi-calculus foundations," Information and Software Technology, Vol. 45, No 15, 2003
8. H. Smith and P. Fingar, "Business Process Management – The Third Wave," Meghan-Kiffer Press, 2003
9. H. Smith and P. Fingar, "IT Doesn't Matter-Business Processes Do: A Critical Analysis of Nicholas Carr's I.T. Article in the Harvard Business Review", Meghan-Kiffer Press, 2003
10. K. Steenstrup, "Business Process Fusion: Powering the RTE," CNET-Gartner Summit Korea – Real Time Enterprise Ahead, 2004, Seoul, Korea
11. M. Weske, W. M. P. van der Aalst and H. M. W. Verbeek, "Advances in business process management," Data & Knowledge Engineering, Vol. 50, No. 1, 2004
12. B. Wood, "CPM Is Key to Business Process Fusion," COM-21-8797, Gartner, 2004
13. WFMC-TC-1011, Terminology & Glossary, Workflow Management Coalition, Lighthouse Point, Fla, 1999

An Information System Approach and Methodology for Enterprise Credit Rating

Hakjoo Lee[1], Choonseong Leem[2], and Kyungyup Cha[1]

[1] IT Research and Consulting, 35-6 Yeouido-dong, Yeongdeungpo-gu, Seoul, Korea
{hjlee, gycha}@itr.re.kr
[2] Department of Information and Industrial Engineering College of Engineering
Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul , Korea
leem@yonsei.ac.kr

Abstract. This research is going to construct an enterprise credit evaluation model which consists of non-financial elements of IT index considered influencing to the financial output of the business entities. To examine the influence closely, this paper analyzes the relationship between the IT investment of enterprises and the elements of IT level and their productivity level or financial output first, and then is going to propose the new evaluation model of which IT elements as a non-financial one activates in a method of evaluating credit of each enterprises by discovering the linkage between the IT level and their financial output finally.

1 Research Background and Its Needs

Currently contribution level to market from the credit information gets more and more important in the variable ways such as economic side of which the credit value, as a parameter, promotes the credit exchanging economic society, business one of which it improves the process of decision-making of enterprises, banking side of which it increases the credit loan of business entities, and managerial one of which it secures the market transactions through the credit management. So when we approach the subject, the credit information, we should consider the mixed reciprocal effects of lots of elements of which enterprises have faced. Nevertheless, general system of credit evaluation about enterprises has analyzed only financial statements and social elementary data, because it is easy to be measured, and these phenomena have conceived the fundamental problem of its credibility.

So, this research proposes that we should accept this essential limitation of easily measured data like financial statements, for we are obliged to acknowledge the utility of current accounting system, and we should assist the system with the other important approach that we suggest in this paper.

The construction of information system through IT investment of enterprises promotes them to be able to achieve the financial output of enterprise more efficiently by diverse IT elements. A recent article of a research shows that one-dollar IT expenditure stimulus about the same amount of raise in productivity of the enterprise within a year and from double to eight times as the time goes in some cases. In addition to the increase in productivity, another research shows that as enterprises who have raised one more dollar IT investment the value of them have increased from

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1269 – 1278, 2005.

five-dollar at least to twenty-dollars as maximum. [1]. We could consider the linkage between the IT investment of enterprises which raises their IT level and the credit level of enterprises from the financial output which is stimulated by their IT level.

2 Direction of Research

This research is based on the hypothesis that "The construction of information system through IT investment causes the increase of expense savings and utility level of whole production processes of an enterprise, the improvement of production output revenue through the increase of efficiency level of input resource, and finally contribution to building the stable business environment of it." To examine this hypothesis, this research classifies and analyzes the subject of study by sizes (large enterprises & small and medium enterprises) and by industries (manufacturing enterprises, constructing enterprises, and distributing/service enterprises), and then examines each financial indices by details following the each steps of financial output and IT investment rate as an actual IT index and the relationship between each indices of the IT level of an enterprise.

Likewise, the same as generally known, we can say that the IT investment has a trend that occurs and concentrates on at the specific term and then the amount gradually decreases, and the IT level after the IT investment grows rather slowly and the link to the financial output reflects its effects more gradually and long-term based. Therefore for the analysis of "time" dimension between the IT investment and its financial output, this research extracts the result of three consecutive years from 2001 to 2003 of "Evaluation of IT Level of Enterprises" [2] by IT Research & Consulting (we call it "ITR" from now on) which has researched and analyzed more than 500 enterprises in Korea and produces the common indices through utilizing the result. To raise the credibility, this research also has analyzed the relation between IT investment and its financial output for the three consecutive years from the year of IT investment using the financial Information [3] of Korea Information Service (we call it "KIS" from now on). Furthermore, to observe the financial output by IT level, this research has inquired the alteration of relationship between the level of detailed indices of each IT elements which consists of IT strategy, IT organization, and IT application level and its financial output.

Finally this research proposes the enterprise credit evaluation model suitable to the size of enterprises and the characteristics of industries of which influence the IT indices for the non-financial elements.

3 Contents of Research

3.1 Analysis of the Relationship Between the IT Investment and IT Level and the Financial Output of Enterprises

Research Design

To analyze the relationship between the IT investment of an enterprise and the financial output, this research has organized a research model that utilizes pooling data to measure the IT investment and its time factional financial output. Because by

the results according to the internal or external environment like enterprise or economic environment when a research measures only the output of IT investment by years there can be some differences, it is sure that there will be a limitation measuring the actual IT investment output. And since the IT investment of an enterprise creates the gradual and consecutive output, it cannot produce accurate result when a research compares the IT investment and its output of a certain year. So to measure the IT investment output accurately, this research suggests that an analysis of IT investment and its output should exclude the environment of a certain year and organize a model not using standard time but using generation time.[1]

Table 1. Classification of Model Enterprises by industry

Unit: Number of enterprise (%)

Section	Manufacturing	Constructing	Distributing/service	Total
Large enterprise	121(28.74)	24(5.70)	70(16.63)	215(51.07)
Small and medium enterprise	126(29.93)	20(4.75)	60(14.25)	206(48.93)
Total	247(58.67)	44(10.45)	130(30.88)	421(100.0)

Primarily this research organized the population by the enterprises of which ITR researched and analyzed in its "Evaluation of IT Level of Enterprises" project using EIII(Evaluation Indices of Industrial IT)[4] developed by ITR and the project includes about 1,500 companies from 2001 to 2003. This paper has organized the final model enterprises that have their financial Information and utilized the data [3] aided by KIS.

The statistical variables used in this research are defined as IT investment variable, IT index one, and financial index one. At first, related to the IT investment variable, existing researches by PIMS(1984) and Turner(1985) made efforts to examine the relationship between the IT investment correlated with sales amount and financial output using IT investment variable. Secondly Hannu and Timo(1995) implemented a research not only about the former one but also about the IT investment correlated with the number of employees and users of IT system and financial output using IT investment variable. As a new index of IT investment, this research has used the index of IT investment rate correlated with the total asset, the rate is {IT investment/ total asset} *100, and it proposes that it could closely examine the relationship between the IT investment and financial output since this index considers the relative importance of IT investment related to the real and concrete value of enterprises.

Related to the IT index variable, this paper has intended to explain the relationship by IT environment using the user satisfaction variable of Mahmood and Becker(1985) and Hannu and Timo(1995) who implemented their research with the variable of the user satisfaction and ability of data management, the inclusive character of IT works, and the maturity level of IT function. This research has considered IT strategy, organization, and application as IT indices, and we believe that this approach of

[1] Generally to consider the characteristics of specific year, there is a method of adjusting IT investment by depreciation, but it is not be commonly utilized because the basis of the depreciation rate is not sufficient. So this research has analyzed the financial output in the year, we call this t year, of IT investment year by year.

various IT environment would be able to analyze the detailed relationship between the IT investment and financial output and to overcome the limitation of situational element of influence between the relationship caused by diverse changeable environment which is essentially included in it.

Table 2. IT index variables

IT Area		IT index
IT Strategy		Whether using Analysis of investment decision-making or not, Method of analyzing investment decision-making, Whether using analysis of afterward effectiveness or not, Whether establishing ISP or not
IT Organization		Object area of analyzing user satisfaction, Term of researching user satisfaction, Whether implementing IT education or not, Authority of IT organization, Formation of IT executives
IT Application	System within an enterprise	ERP, Human resource management, Accounting management, Groupware, KMS
	System Among enterprises	B2B Sales, B2B Service, B2B partner management
	System to customers	B2C Sales, B2C Service, B2C Marketing, Integration of customer database

Related to the financial index variable, Hannu and Timo(1995) did it using the financial index of the increase rate of sales amount, pure profit from sales amount, and liability ratio. This research has intended to use diverse approaches using five points of view and 27 indices as those of analyzing financial output.

Analysis and the Result

Above all, to analyze the relationship between the IT investment and its financial output, this research has analyzed 6 enterprise groups considering their industries (manufacturing industry enterprises, constructing ones, and distributing/service ones) and their sizes(large ones and small and medium ones). The banking industry enterprises are omitted in this research because their quiet different characteristics of financial indices do not help this comparative analysis meaningful. To closely examine the similarity of the relationship between the IT investment and its financial output, this paper has implemented Pearson correlation analysis. And to analyze the financial output caused by IT investment overcoming time gaps and to examine the similarity, it has analyzed the data from the year of IT investment 1 year and 2 year after the IT investment by turn. This treatise proposes the concrete result of the group of large manufacturing enterprises only.

From the analysis of the relationship between the IT investment and its financial output of large manufacturing enterprises, this research gets the following result; IT investment contributes to enterprises by the effect of utility increase and expense saving. But the effect of actual financial output of an enterprise such as the improvement of revenue structure or enlargement of circulation is not enough to be positive. But we have discovered that the correlation between the IT investment and revenue type indices has a trend of showing positive effect in long-term base.

Table 3. Correlation between the IT investment and its financial output of the large manufacturing enterprises

Section	Financial rate	t year	t+1 year	t+2 year
Revenue Type index	Rate of increase in sales amount	0.03	0.12	0.10
	Rate of current profit from Total capital	0.10	0.23(*)	0.12
	Rate of pure profit from total capital	0.08	0.10	0.12
	Current profit rate of enterprise	0.09	0.22(*)	0.23(*)
	Pure profit rate of enterprise	0.05	0.07	0.15
	Current profit rate of sales amount	0.08	0.17(*)	0.12
	Operational profit rate of sales amount	-0.01	0.13	0.03
	Sales amount versus Appropriation Expense for bad debt	-0.05	-0.06	-0.05
	Sales bond versus appropriation expense for bad debt	-0.07	-0.10	-0.09
Expense index	Cost rate of sales	-0.12	-0.16	-0.14
	Salary expense versus sales amount	-0.09	0.002	-0.01
	Distribution expense versus sales amount	-0.04	-0.11	-0.20
	Sales administration expense versus sales amount	0.17	0.07	0.13
Utility index	Total capital circulation rate	0.31(*)	0.36(*)	0.49(*)
	Capital circulation rate	0.29(*)	0.04	0.02
	Managerial capital circulation rate	0.40(*)	0.44(*)	0.60(*)
	Fixed asset circulation rate	0.17(*)	0.20(*)	0.29(*)
	Visible fixed asset circulation rate	0.24(*)	0.30(*)	0.47(*)
	circulation rate	0.80(*)	-0.01	0.39(*)
	Product circulation rate	0.19(*)	0.03	0.06
	Sales bond circulation rate	0.17(*)	0.33(*)	0.27(*)
Productivity index	Total capital investment efficiency	0.23(*)	0.26(*)	0.50(*)
	Equipment investment efficiency	0.23(*)	0.26(*)	0.45(*)
	Machinery investment efficiency	-0.01	0.01	-0.01
Circulation index	Circulation rate	-0.02	0.08	0.15
	Current deposit rate	0.01	0.17	0.24
	Fixed rate	-0.05	-0.12	-0.23

※ Star (*) is the mark of useful index in the 5% usefulness level.

Then this research has intended to measure the size of effect of which IT level in addition to the investment reaches to the financial output of enterprise through implementing analysis of relationship between IT investment and its financial output of enterprises. To measure the size of financial output according to different IT levels, this research has classified and analyzed three IT application and utilization level. Because it has some difficulties in translating the result which utilized and analyzed simple IT indices such as user satisfaction or job of CIO as well as the limitation of data, existing researches had trouble in examining the relationship between IT indices and its financial output. In this paper we have utilized the model of "Evaluating IT level of enterprises" using EIII developed by ITR, intended to measure the financial output according to IT environment by quantitatively analyzing detailed IT indices of each actual areas, and then closely examined the relationship between the detailed indices and its output.

This research has measured the similarity level between IT level and the financial output dividing into revenue type, expense savings, activation, productivity, and circulation areas, and, to overcome the differences in translation between each index and to raise the explanation capacity of a financial output as a single index, we have implemented making elements by areas not of examining similarity between each index but of Factor Analysis in the financial output area. Moreover, to measure the differences between financial output levels caused by the IT level, this paper has implemented comparison of average figures of groups through the analysis of variance (ANOVA) using IT level by areas and their scores of elements.

As a result, even though there is a difference by their industries we have discovered that the enterprises having high level of IT strategies get more financial output in expense savings and revenue type than the ones having low level of IT strategies in the domestic large manufacturing industry groups. That is, this paper proposes an important discovery that even though IT investment could add the managerial executives high financial pressure by increase of operational and current expenses to an enterprise it could save IT investment expense and maximize its operational efficiency through IT application and utilization based on IT strategy.

3.2 Credit Evaluation Model by Enterprise's IT Level

Research Design

Model enterprises in this research are the same model that has been proposed in the research "Analysis of relationship between the IT investment and IT level and its financial output of enterprises." This research has analyzed the large and the small and medium manufacturing enterprises and the large and the small and medium non-manufacturing ones, constructing ones and distributing/service ones. The reason why we have excluded the classification of constructing ones and distributing/service ones is that the number of the superior and inferior groups is too small for producing a meaningful value and in the actual expectation of credit level of enterprises people use the classification system of manufacturing and non-manufacturing. We have organized the standard of classifying the superior or normal enterprises and the inferior ones following the criteria shown in the table 7 and through this procedure the table shows the classification system of the large manufacturing enterprise and the small and medium ones and the large non-manufacturing enterprise and the small and medium ones.

Table 4. Classification system of superior and inferior enterprise

Section	Classification System
Inferior enterprise	[Total capital ≤ 0] or [An enterprise whose cash flow after 2 consecutive years of operational activity is below zero], or [An enterprise whose current pure profit is below zero for 2 consecutive years]
Superior enterprise	An enterprise which is not an inferior one

Table 5. Superior and inferior enterprise model organization

Section		Manufacturing	Non-manufacturing
Large enterprise	Superior	98(81.0%)	73(77.7%)
	Inferior	23(19.0%)	21(22.3%)
Small and medium enterprise	Superior	80(63.5%)	32(40.0%)
	Inferior	46(36.5%)	48(60.0%)

In the research of forecasting the inferiority of an enterprise or credit level, there is no theoretical basis to select the financial rate variable as a independent variable, explanatory one. Therefore this paper has classified and analyzed 17 financial rate by revenue type, stability, growth, activation, and circulation.

Using "Evaluation of IT level" data of ITR as an IT index, this research has utilized the non-financial information of IT level of enterprises which is very trustful, and our use of the data let us expect that the analysis level of this treatise would be superior to the other existing papers studying the relationship between the IT investment and its financial output. We have analyzed 12 detailed IT indices by classifying IT strategy, IT Organization, and IT application.

Analysis and the Result

Especially in this paper, we would closely examine what IT level of enterprises by positive analysis, as a non-financial element, roles in evaluating enterprise's credit and how much the level contributes to the credit level. As we have explained at the former model selection, the dependent variable dealt in credit level model of this research is the classification of the enterprises into superior group or inferior one. So, the IT level variable as a financial and non-financial variable which can be used as an independent variable of this model selects the variable having meaningful statistical difference as a result of the t-examination between superior and inferior groups usable as a dependent variable, Logit, and X^2 examination. The reason of choosing t-examination, Logit, and X^2 examination as a choosing criteria of an independent variable is because we could distinguish the difference between superior and inferior groups when the difference of average between levels or explaining capability of model is considered to be meaningful variable. This analysis has implemented t-examination and Logit in analyzing numerical variable, and X^2 examination in analyzing category type variable, considering whether there is consistency of dispersion in financial rate of non-manufacturing enterprise groups or not it has used dummy variable in industries of constructing and distributing/service as an independent variable of model. Therefore, to examine the contribution of financial variable to this model, this research defines financial variable as independent variable and brings out the organized model by dummy variable by industry, in case of non-manufacturing industry

Considering the result of the analysis of the t-examination and Logit about the large manufacturing enterprise, we could concluded that there is a difference between detailed financial indices of superior groups and inferior ones, since until the 8th ranked indices in the financial variable showing superior discrimination rate of the analysis result of Logit shows meaningful result of t-examination within 10% meaningfulness level.

Table 6. Analysis result of X^2 examination about each IT level candidate variables of large manufacturing enterprise

Variable		Examination		Rank
		X^2 Value	Rate of meaningfulness	
IT Strategy	Whether using analysis of investment decision-making	4.09	0.043	1
	Method of analyzing investment Decision-making	1.37	0.242	5
	Whether using analysis of afterward effectiveness	2.31	0.128	4
	Whether establishing ISP	2.63	0.268	6
IT Organization	Whether implementing IT education	0.00	0.972	12
IT Application	Whether applying ERP	3.25	0.072	3
	Whether applying groupware	3.36	0.067	2
	Whether applying B2B sales system	0.95	0.331	7
	Whether applying B2B service system	0.27	0.603	10
	Whether applying B2B partner management system	0.39	0.533	9
	Whether applying B2C service system	0.02	0.905	11
	Whether applying B2C marketing system	0.39	0.531	8

Following the above result, in case of the large manufacturing enterprise, the stable formula of the discrimination shows below one.

$Z = \beta_0 + \beta_{11}$(total capital current revenue rate) + β_{12}(Sales amount current revenue rate) + β_{13}(Total capital pure revenue rate) + β_{14}(Banking expense versus sales) + β_{15} (Own capital rate) + β_{16} (Liability versus sales) + β_{17} (the increase rate of sales amount) + β_{13} (Sales bond circulation rate) + β_{21} (Whether using the investment decision-making analysis or not) + β_{22} (Whether applying groupware or not) + β_{23} (Whether applying ERP) + β_{24} (IT investment in comparison with total asset)

Through the above whole examinations on the candidate financial variable by classes and the IT variable, this research selected the first financial and IT variable. And then through the entry standard stepwise method, it makes the discrimination formula of financial element suitable, and finally through combining the selected IT element as the non-financial indices with the suitable discrimination formula for only

Table 7. Discriminating model by selecting financial variable of the large manufacturing enterprise

Name of variable	figure(β)	Standard Deviation	wald X^2	Rate of Meaningfulness
Sales amount revenue rate(□)	0.0394	0.0231	2.9118	0.0879
Own capital rate(□)	0.0355	0.0126	7.8986	0.0049
The increase rate of sales amount(□)	0.0399	0.0223	3.2154	0.0729
Constant term	0.4081	0.5251	0.6040	0.0371
Forecasting ability of model	83.3%			
Forecasting model	$Z = 0.4081 + 0.0394$□ $+ 0.0355$□ $+ 0.0399$□			

the financial elements, it makes the final discrimination formula stable. In case of financial variable about the large manufacturing enterprise, we have discovered that the financial indices of revenue type which includes the sales amount current revenue rate variable, own capital rate, the increase rate of sales amount, and so on, stability, and growth are the discriminating variables in classifying the superior enterprises and inferior ones, and the forecasting ability of this model shows 83.3%.

The discriminating model of financial variables including IT ones shows that, in financial variables, variables of sales amount current revenue rate, own capital rate, and the increase rate of sales amount and, in the IT level variables of enterprises, variables of IT investment in comparison with total asset, whether using investment decision-making analysis or not, whether applying ERP or not, and whether applying groupware or not could discriminate superior enterprise and inferior ones, and the forecasting ability of this model shows 88.4%.

Table 8. Comparison of discrimination rate by models

Enterprise classification		Discrimination rate		Increasing level of Discrimination rate
		Financial variable	Financial variable + IT variable	
Manufacturing Industry	Large enterprise	83.3%	88.4%	**+5.1%p**
	S/M enterprise	-	-	**-**
Non-manufacturing Industry	Large enterprise	88.0%	91.2%	**+3.2%p**
	S/M enterprise	77.3%	85.5%	**+8.2%p**

We could bring out meaningful financial variables and IT level ones in enterprise groups except for the small and medium manufacturing enterprise group. In case of large manufacturing enterprises, the increasing level of discrimination rate including IT level variables shows 5.1% higher discrimination hit ratio than that including financial variables only. In case of large non-manufacturing enterprises, the increasing level of discrimination rate including IT level variables shows 3.2%p higher discrimination hit ratio than that including financial variables only, and In case of small and medium manufacturing enterprises, the rate including IT level variables shows 8.2%p higher discrimination hit ratio than that including financial variables only. This result means that the classification of superior groups and inferior ones by IT level is meaningful and the IT level of enterprises can contribute to forecast the credit level of enterprises.

4 Conclusion

This research, by bringing out the result whether IT investment of enterprises directly influences to the sectional financial outputs, which includes revenue increase, expense savings, utilization increase, productivity increase, and circulation increase, or not by industries and sizes, has verified the relationship between the IT investment and its financial outputs. Moreover, this research has verified whether the improvement of enterprise's IT level by IT investment influences the financial outputs of enterprises

or not by bringing out the indirect relationship by industries and sizes. Based on the above analysis result, we have implemented analysis of the contribution level of which IT influences stimulus enterprise's credit level, and in case of large manufacturing enterprises, the increasing level of discrimination rate including IT level variables shows 5.1% higher discrimination hit ratio than that including financial variables only. Through this result, we could verified that forecasting credit level of enterprises by IT level is meaningful, and based on the result, could expect that further research can develop more clear system and model of translating relationship between IT level and credit level of enterprises afterward by implementing supplementing research based on the more reliable data collected for longer period.

References

1. Jeon Jong Gyu.: Why the effect of the IT investment of every enterprise is different. information Technology Management(2002)
2. IT Research and Consulting.: Result of IT Level Evaluation of Enterprise Report. enterprise ITR (2001, 2002, 2003)
3. Korea Information Service.: Enterpise financial Information. KIS (2001, 2002, 2003)
4. Kim In Ju.: Development of Integrated Evaluation System based on IT Maturity Model Basis. Yonsei University Doctoral Degree treatise (1999)
5. The PIMS Program.: Management Productivity and Information Technology. the Strategic Planning Institute. Cambridge, MA (1984)
6. Turner, J.: Organizational Performance, Size and the Use of Data Processing Resource. Working Paper, No. 58. Center for Research in Information Systems, New York University, Now York (1985)
7. Hannu Kivijarvi and Timo Saarinen.: Investment in information systems and the financial performance of the firm. Information & Management 28(1995) 143-163

Privacy Engineering in ubiComp

Tae Joong Kim[1], Sang Won Lee[2], and Eung Young Lee[3]

[1] Privacy Protection Team, Korea Information Security Agency
78, Garak-dong, SongpaGu, Seoul 138-803, Korea
taej@kisa.or.kr
[2] Dept. of Management Engineering, Korea Advanced Institute of Science and Technology
207-43, Chongyangni-Dong, Dongdaemoon-Gu, Seoul 130-012, Korea
sangwonlee@kgsm.kaist.ac.kr
[3] Information Security Policy Research Team, Korea Information Security Agency
78, Garak-dong, SongpaGu, Seoul 138-803, Korea
eylee@kisa.or.kr

Abstract. In the ubiquitous age, privacy will be the matter of trade-offs about pros and cons of revealing personal information for personalized services. Ubiquitous computing demands a fundamental shift in the control of personal information and requires disclosure of personal information. As we enjoy comfortable life, the invasion of personal information can be occurred at the same time. The privacy requires the effective security, but the effective security does not guarantee the effective privacy. We present the privacy engineering in order to prevent the privacy invasion and measure the economic value of privacy. We hope this privacy engineering in ubiComp will be used as one of the tool for protecting the users in the ubiquitous age. The approach includes the followings; the architecture of the privacy engineering, the database modeling, the privacy impact assessment, the economic value assessment of the privacy.

1 Introduction

Privacy can have many aspects, but for purposes of this principle and the corresponding criteria, privacy is defined as the rights and obligations of individuals and entities with respect to the collection, use, disclosure, and retention of personal information. Personal information is defined as any information relating to an identified or identifiable individual. [1]

The threat of privacy break-in may not seem real until it happens. Specific risks of having inadequate privacy policies and procedures include; Damage to the organization's reputation, brand, or business relationships, Legal liability and industry or regulatory sanctions, Customer or employee distrust, and Disruption of international business operations. The future is uncertain. Firms can use scenario planning so as to develop privacy strategies in the technological, regulatory, and competitive landscape.

Firms must establish internal control and monitoring measures so that they may be enforced. For firms to prevent these risk potentials, it is desirable that they should put into three specific actions which would be difficult as the follows. The wireless world,

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1279 – 1288, 2005.

in surfing the Internet or viewing e-mail, has the same privacy concerns as the wired world. The first consideration in developing a sustainable approach to privacy is the creation of an underlying enterprise data strategy. The second is the issue of technology infrastructures. The third element of a privacy strategy is business operations- the side of the triangle that shapes processes and procedures that guide consumer interactions. [2]

We present the privacy engineering in order to prevent the privacy invasion and measure the economic value of privacy. Now let us study on the followings; Tech Architecture, Legal Architecture, and Privacy Engineering.

2 Technical Architecture for Privacy

2.1 Mapping

The technology infrastructure depicts the component parts (e.g., applications, host platforms, databases, networks, etc.) of the information technology environment that enable the flow of personal information. Application systems and business processes that collect, use, disclose or store personal information must be identified and understood. To assess compliance with privacy goals and objectives at the most granular level, the organization must develop an inventory. If notice and choice are to be offered at every collection point, then the inventory can be used to identify all collection points. The inventory also serves to identify where security must be in place to protect personal information. [3]

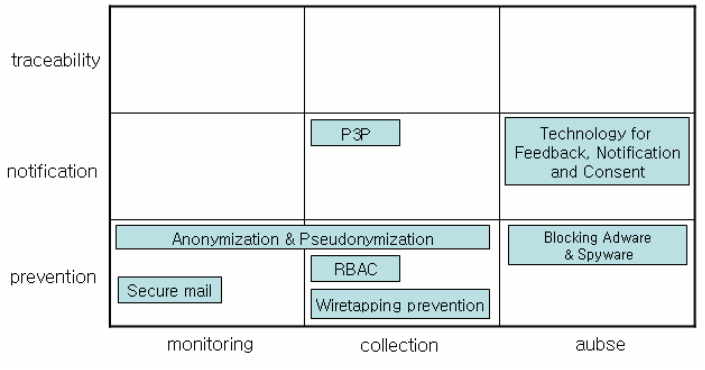


Fig. 1. In order to strengthen the protection of personal information, a foundation of laws and polices must have been established. Also, the government should prepare the related technology which can handle the privacy issue. In this process, we have to make an effort not to omit the whole technology to handle the increasing the future privacy issues. We suggest the following technological framework to prepare the safe ubiquitous computing environment. [4]

With the above matrix, it leads to a new framework for categorizing privacy protection mechanism. In this framework, “Anonymization & Pseudonymization” technologies are preventative methods that can be used on monitoring and collection

of privacy. “Secure mail” also can be used as prevention for monitoring and collection process. In contrast, “P3P (Platform for Privacy Preferences)” tells web users how their information will be used, so this notification mechanism might be a solution for data collection phase. And “Technology for Feedback, Notification and Consent” could be efficient method for abuse phase. [4]

2.2 Control Point

In many organizations, to mitigate the risks associated with information systems, the IT audit function is assigned the responsibility of implementing a system of internal controls. Because of additional risks associated with e-commerce systems and the resulting need for strong control procedures, it is important that management appreciates the significance of having IT auditors participate in the systems development process. These internal controls are activities performed to eliminate risks or minimize them to an acceptable level. In most cases, it is cost-prohibitive to implement every type of control in an effort to eliminate all elements of risk. [5].

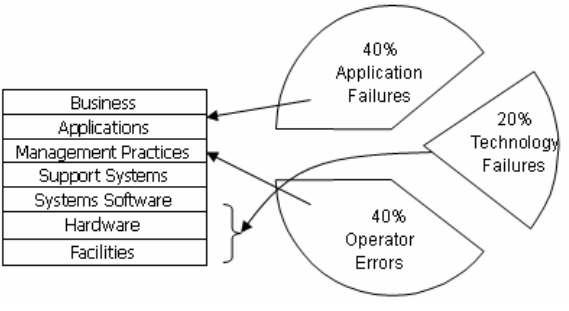


Fig. 2. By Disaster Recovery Institute, even though the cause of service suspension was the facility part in days past, it moves to the higher stack nowadays. [6] It is so because users come in contact with various programs more easily and then many side-effects appear

Of the internal attacks cited in Ponemon's report, about almost 40 percent occurred because well-intentioned employees inadvertently caused security problems by how they handled sensitive information. Only 30 percent were attributed to malicious employees. "Most internal security issues are due to organizational sloppiness," he said. "These aren't bad people. They are just trying to get a job done, but they aren't considering all the consequences to their actions." [7]

The operational errors must be reduced by use of management practices and the application and technical failures must be reduced through standardized design and systematic procedures. The internal controls must be added in compliance with organizational regulations. Moreover, every view must be offered to the authorized user according to data constraints. Traceability should be guaranteed by using information systems.

3 Legal Architecture for Privacy

3.1 Considerations

Information technologies have increased the interests about the right of privacy since the 1960s and 1970s. The Council of Europe's 1981 Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data and the Organization for Economic Cooperation and Development's (OECD) Guidelines Governing the Protection of Privacy and Transborder Data Flows of Personal Data set out specific rules covering the handling of electronic data. All require that personal information must be; Obtained fairly and lawfully, Used only for the original specified purpose, Adequate, relevant and not excessive to purpose, Accurate and up to date, Accessible to the subject, Kept secure, and Destroyed after its purpose is completed. [8]

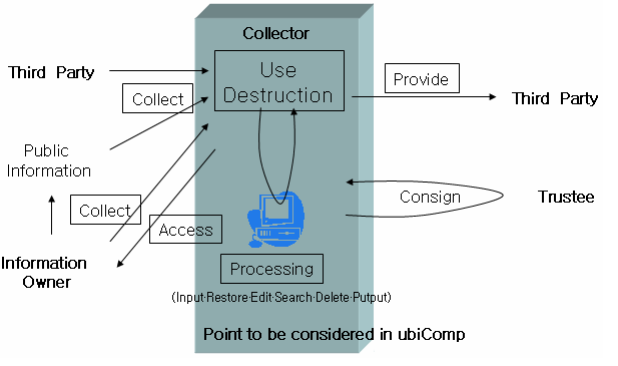


Fig. 3. In the ubiquitous environment, a company should collect the agreement with the personal information owner. The ubiquitous network and data mining techniques will increase the tracking and profiling of personal information. The shift to the pervasive computing paradigm brings forth new challenges to security and privacy, which cannot be addressed by mere adaptation of existing security and privacy mechanisms. Unless security concerns are accommodated early in the design phase, pervasive computing environments will be plenty of vulnerabilities and exposures

It should add the use and processing step of the life cycle of personal information on the current law. The current law just focused on the collect, access, provide and cosign step of the life cycle of personal information. In ubiComp, the analysis of a company will increase because of the ease of collection for personal information.

A customer could be monitor in a shop by a company for marketing. A company will collect and analyze using accumulated personal information. A company can sell and distribute the personal information intentionally or accidentally. The law should stipulate how to control the personal information on processing step. [4]

3.2 Implementations

A full assessment identifies both gaps in the design of the privacy program in considering business risks and opportunities, as well as gaps in the operating effectiveness of current policies and practices. It tells you what will never work because of flawed design and what isn't working because of flawed execution. The resulting strategy and plan will document recommended solutions as well as how to ensure those solutions are implemented. The assessment should take advantage of graphical representations to depict the results since many of us respond more to pictures than words. [3]

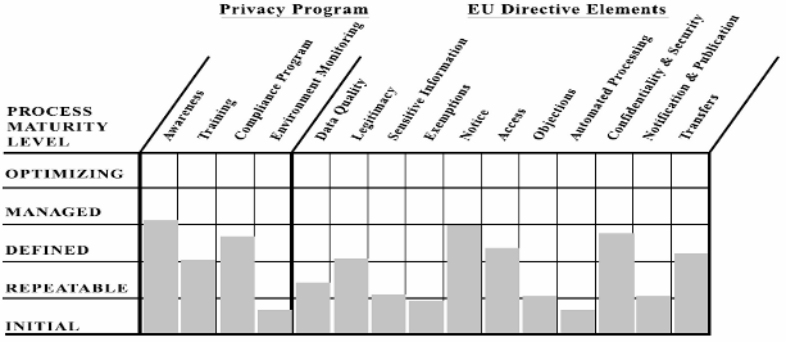


Fig. 4. This depicts the results of an assessment where the Capability Maturity Model was adapted to indicate the process maturity of key elements of the privacy program and major ideas and principles in the EU Directive

In the final analysis, the assessment should set forth a thoughtful strategy and plan for closing the gaps and achieving the goals and objectives. Findings and recommendations should be presented in consideration of the organization's culture and values as well as risk management priorities. Finally, consideration must be given to overcoming anticipated barriers to change, as well as the trade-off between cost and benefit. Now, with plan in place, it is time to implement. [3]

The implementation method of related laws should be linked to Privacy Engineering. In designing database, the constraints for securing privacy should be defined at the early stage.

4 Privacy Engineering

Building the world's datasphere is a three-step process—one that we've been blindly following without considering its ramifications for the future of privacy. First, industrialized society creates new opportunities for data collection. Next, we dramatically increase the ease of automatically capturing information into a computer. The final step is to arrange this information into a large-scale database so it can be easily retrieved at a moment's notice. Once the day-to-day events of our lives are

systematically captured in a machine-readable format, this information takes on a life of its own. It finds new uses. It becomes indispensable in business operations. And it often flows from computer to computer, from business to business, and between industry and government. If we don't step back and stop the collection and release of this data, we'll soon have a world in which every moment and every action is permanently "on the record." Now we show solutions for privacy with Modeling, Security Implementation, Impact Assessment, and Economic Value Assessment.

4.1 Database Modeling

Proper authentication is a critical component of DB modeling. It is because if once a party has been accepted into the system, a legally binding transaction process has begun.

Database Security Control consists of the Flow Control, the Inference Control, and the Access Control. DB Security Control measures restrict who gets in, manage their identities and access rights, Column level controls, and Audit to hold users accountable for their actions.

Classification of Security Threats according to the way they occur: (1) Accidental; Human errors (incorrect input, incorrect use of applications), Errors in software (incorrect application of security policies, denial of access to authorized users), Natural or accidental disasters (damage of hard-/software) (2) Intentional; Authorized users who abuse their privileges and authority, Hostile agents (improper users, insiders / outsiders) executing improper reading or writing of data legal use of applications can mask fraudulent purpose. (Viruses, Trojan Horses, Trapdoors) [9]

To ensure successful identity management, a digital identity solution should support at least the following basic requirements. (1) Reliability and dependability: Identity theft is one the fastest growing electronic crime and it is expected to accelerate. Digital identity must offer protection against forgery and related attacks. (2) Controlled information disclosure: Users must be given control on what identity to use in specific circumstances. Control must also be given with respect to possible replication and misuses of the identity information a party reveals in a transaction. (3) Mobility support: The mobile computing infrastructure can keep track of an individual's physical location. In addition, mobile computing bears some peculiarity such as limited bandwidth and limited display size. [10]

There are several ways to look at approaches to identity management. One may look at this question as a matter of policy and law; or as business cases and practices; or as technical architectures and technologies; or even at guiding philosophies and principles. [11] We must pay attention to understand how business process controls and identity management work together, which could ensure that the identity management and its supporting infrastructure are delivered within the context of business objectives.

4.2 Security Implementation

The Ponemon Institute's data security study asked respondents what type of leaks they'd suffered. Because respondents could cite more than one category per incident, the percentages don't total 100. (1) 22 percent of leaks involved customers' personal

data. (2) 10 percent involved workers' personal data. (3) 39 percent disclosed confidential business data. (4) 14 percent leaked intellectual property, including software code. (5) 16 percent: "Other." [7]

The emerging threats described above are becoming increasingly complex and difficult to resolve. Like the newest viruses and worms, the threats can be blended, where one threat facilitates another. Once inside a company or one of its partners, a trusted employee can do enormous damage. Often such leaks disclose the most sensitive of data. In addition to products that control who gets access to what information, a slew of new start-ups focus on securing digital content and watching where it goes. Products in this category vary in their approach. Some focus solely on protecting intellectual property from being leaked, while others also perform forensics analysis, digital rights management and security policy management. [7]

Effective application security controls spring from such standards as least privilege and separation of duties. These controls must be precise and effective, but no more precise or granular than considerations of cost and value dictate. At the same time, they must place minimal burdens on administrators, auditors, and legitimate users of the system. Controls must be built on a firm foundation of organizational policies. Although all organizations probably need the type of policy that predominates in the commercial environment, some require the more stringent type of policy that the U.S. government uses, which places additional controls on use of systems. [12]

Users typically cannot access these applications without being registered to use that software, i.e., having a user account set up for them. As a result, user accounts need to be added, deleted and modified constantly in response to people joining or leaving the organization, changing roles, or moving locations. Subsequently, a typical user needs accounts on multiple platforms and access to a number of applications. As such, organizations conducting business on the Internet must have or develop the ability to authenticate, authorize and provide user access rights in a unified, consistent and effective manner. [13]

4.3 Impact Assessment

When we use the word “sensitive data,” it implies that if these kinds of data are misused or breached, the individuals would suffer from severer damage than they would suffer if their name, address or other information is misused. The clearer conceptualization and practical thinking concerning the trust and the relationship between them are required.

In this point, PIA has been welcomed as a tool for identification of privacy risks in proposed government and private sector initiatives. PIA is an analysis of how information is handled; (1) to ensure handling conform to applicable legal, regulatory and policy requirements regarding privacy (2) to determine the risks and effects of collecting maintaining and disseminating information in identifiable form in an electronic information system (3) to explain and evaluate protections and alternative processes for handling information to mitigate potential privacy risks [14].

This definition makes the linkage between PIA and risk assessment, placing the concept of risk at the center of process. To some extent, we may infer risks to privacy that comes from security breaches or breakdowns of information equipment. But, in the ubiquitous computing, this term needs to be understood from a social scientific

perspective in terms of the different characteristics of individuals, or of categories of individuals, who experience different kinds of and degrees of privacy breaches and who enjoy different level and forms of privacy protection in the ubiquitous computing.

It would be impossible to know which data of someone are more at risk for privacy than others in the ubiquitous computing. The construction of 'data subject' is enriched not by understanding of various kinds of data subjects, but by explanations of what differentiates them and by recognition of the different consequences; they experience through involvement in the variety of information processes that put them at risk or that protect them.

The PIA process ensure that privacy issues are identified and addressed by policy makers at the initial stages of a new project or policy- at the conceptual stage, the design approval stage, and the funding stage [15]. Accordingly, it is important to implement the PIA early in the project life cycle, at the conceptual stage of project.

And it is important as well for the PIA to be performed by an independent officer or entity not linked to the project. And Privacy Commissioner should takes charge of this role. Many officers in various countries have authority over private information and governmental databases. An important thing to success of Privacy Commissioners is to provide with adequate resources to conduct oversight and enforcement.

4.4 Economic Value Assessment

Most items of personal property (mobile phones, houses, etc.) have a value that can be determined in the marketplace. In the case of privacy, the value and trust equation is even less clear, making corporate responses to consumers' expectations difficult to reflect in policy. Perceptions about what is or isn't personal information varies by company, industry, transaction type and the expected consequences of violation. The value of privacy is not only elusive- determining it can be fraught with emotion and intangible considerations.

It is difficult to distinguish clearly between online personal information and offline one. In most real cases, the legacy systems and existing infrastructure prohibit the distinguishing between online and offline personal information. Online personal information includes users' propensity, product evaluation, search behavior, purchasing history, and so forth. Online personal information is a kind of customers' behavior in the view of economic model. In electronic commerce, online personal information is intimately associated with cookies and IP addresses which track customers' behaviors. Offline personal information presents real personal identification-information and is presented by identification information such as credit card number and social security number.

The economic value can be distinguished in the viewpoint of enterprises and users. The economic value of enterprises consists of compliance cost and non-compliance cost of law. The non-compliance cost of law includes fine, direct loss, indirect loss, potential loss, and so on. The non-compliance cost of law includes the cost of policy development and maintenance, educational cost, implementation cost of technology, audit or control cost, and so on. The economic value should include the benefits caused by law compliance. The economic cost of users consists of the cost of loss and the cost of indemnity. The cost of loss includes the cost of users' loss, purchasing

cost, spam mail, ID theft, and so on. The cost of indemnity includes value and the amount of indemnity. The economic value of users should include the benefits caused by information sharing. And then the economic analysis can be used to define the structure for privacy protection.

Costs may be classified as tangible or intangible. The tangible costs of Data Protection comprise two main elements: (1) The cost of running the supervisory authority and the payment of any fees for notification (2) Compliance with the Data Protection principles, in particular the costs of the provision of data subject access. Intangibles include: (1) The perception that compliance might reduce competitiveness, by limiting what can be done with customers' information; and (2) Inefficiencies introduced by restrictions on the sharing of data and the additional bureaucracy associated with compliance activities. [16]

The assessment of the shortfall between what we need and what we have is a tedious process. It involves analyzing the process maps from a "what can go wrong" perspective vis-à-vis privacy and data protection. It involves assessing every requirement of the entity's privacy policy and charter (compliance with asserted actions, requirements of specific laws, etc.) against each and every collection, use and disclosure of personal information. It involves assessing the security of personal information whether electronic or physical. [3]

The economic value aims at grasping the disaster differences according to types of countermeasures, by use of the model which calculates the loss. And by suggesting the security level of personal information, it shows the method to compute the criteria in order to elevate the standards such as technological counter-plans, administrative counter-plans, authentication, and so on.

5 Conclusions

In the electronic commerce, it is important for companies to build trust for creating end-to-end privacy practices, which promise value to consumers. By combining these elements into an integrated privacy strategy, a company can greatly increase its success in building sustainable relationships, and in doing so, achieve competitive advantage over businesses that are judged less trustworthy by potential customers, vendors and partners.

What causes firms the most difficulty is ensuring that online privacy policies are consistent with the approach taken in the offline world. Up to now, we presented the privacy engineering in order to prevent the privacy invasion and measure the economic value of privacy. The approach is the followings; the architecture of the privacy engineering, the database modeling, the privacy impact assessment, the economic value assessment of the privacy.

To avoid the gap between perceived and actual practice, monitoring is key. Maintaining compliance requires your organization to monitor internal processes and the regulatory environment on an ongoing basis, while responding to any changes effectively and on a timely basis. [16]

Addressing privacy and security concerns encourages customers to transact using Web-based channels that are presumably the lowest cost among sales channels. If the fears and concerns about transacting online are addressed, then it follows that more

business will flow through the Web site. For the business, this means more profitable customers. And for the customers, this means lower cost. [3]

As the privacy program matures, the issue of compliance will emerge. The privacy task force will have confidence that the privacy program is well-designed but be unsure whether it is complied with throughout the organization.

References

1. US: US Safe Harbor Privacy Principles (2000)
2. Deloitte Research: Creating a privacy value strategy (2002)
3. Glasser Legalworks: Privacy- Taking Action to Safeguard Customer Loyalty (2001)
4. TJ Kim, EY Lee, IH Kim, KI Yoon, YJ Kim: Architecture of the Privacy Governance in ubiComp (2004)
5. Jagdish Pathak: Information Technology Auditing and cybersommerce (2004)
6. Disaster Recovery Institute International: Evaluating & selecting the most appropriate continuity strategy for your organization (2002)
7. CNET: Securing data from the threat within (2005)
8. Linda Ackerman, James Kempf: Wireless Location Privacy- A Report on Law and Policy in the US, the EU, and Japan (2003)
9. Addison-Wesley: Database Security (1994)
10. 10 Information Society Technologies: Identity management PIM Roadmap- Multiple and Dependable Identity Management- R & D Issues (2002)
11. The National Electronic Commerce Coordinating Council: Identity Management- A White Paper (2002)
12. Stanley Kurzban: Implementation of Access Controls (1998)
13. IT governance Institute: Enterprisewide Identity Management (2004)
14. Office of Management and Budget: OMB Guidance for implementing the privacy provision of the E-Government Act of 2002 (2003)
15. Privacy and E-Government: privacy Impact Assessments and Privacy Commissioners- Two Mechanisms for Protecting Privacy to Promote Citizen Trust Online (2003)
16. Peter R. Harris: The European perspective is Data Protection value for money? (2004)

Development of a BSC-Based Evaluation Framework for e-Manufacturing Project

Yongju Cho[1], Wooju Kim[2], Choon Seong Leem[2], and Honzong Choi[1]

[1] Manufacturing Process Research Center, Korea Institute of Industrial Technology, 994-32, Dongchun-dong, Yeonsu-gu, Incheon Metropolitan City 406-130, Korea
{yjcho, choihz}@kitech.re.kr
[2] Department of Industrial System Engineering, Yonsei University, 134 Seodaemun-gu, Seoul 120-749, Korea
{wkim, leem}@yonsei.ac.kr

Abstract. Recently, in order to swiftly satisfy the various product demands by customers reinforced collaboration among industries or business types as well as increased outsourcing have emerged as indispensable constitutive elements of modern manufacturing. An e-Manufacturing pilot project for supporting design collaboration of a domestic mould company is now underway. To drive efficiently the project, different evaluation frameworks per stakeholder in the e-Manufacturing project are required. In this study, we proposed a unified evaluation template which is composed of two viewpoints (user and operator) and three steps (pre-evaluation, operation evaluation and post evaluation). Particularly, we proposed the evaluation framework for contractor mould companies from the user's viewpoint using BSC concept.

1 Introduction

In the current global business environment, manufacturers experience intensified competition between one another due to oversupply; additionally corporate survival itself is threatened with the lack of speedy performance in meeting diverse customer demands. For these reasons, product manufacturing in the manufacturing industry has become transformed from the mass production structure in the past into a small or large quantity of multi-typed products to meet customer demands, which themselves have recently become more diversified. In addition, labor division and business cooperation have grown stronger both by area and by scope in order to secure low-cost, high-efficiency production. As such, in order to swiftly satisfy the various product demands by customers reinforced collaboration among industries or business types as well as increased outsourcing have emerged as indispensable constitutive elements of modern manufacturing.

Technological cooperation in the product and mould design process is extremely important among partners in the mould industry, which is a basis for the manufacturing industry. The goal of technological collaboration in design is to preview material, product, and part requirements, equipment for shaping/assembly production, or production process elements in the early design stage prior to actual product manufacturing. In this milieu, it can be seen that effective operations through collaboration in the

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1289 – 1296, 2005.

design process can greatly reduce opportunity and time costs that normally are incurred in ensuing stages. Such technological collaboration is achieved when scattered specialized businesses participate at real time over the Internet and synchronously discuss operation sharing and cooperation. Consequently, to allow scattered specialized businesses and actual related companies to collaborate easily, a Hub for connecting them must be built.

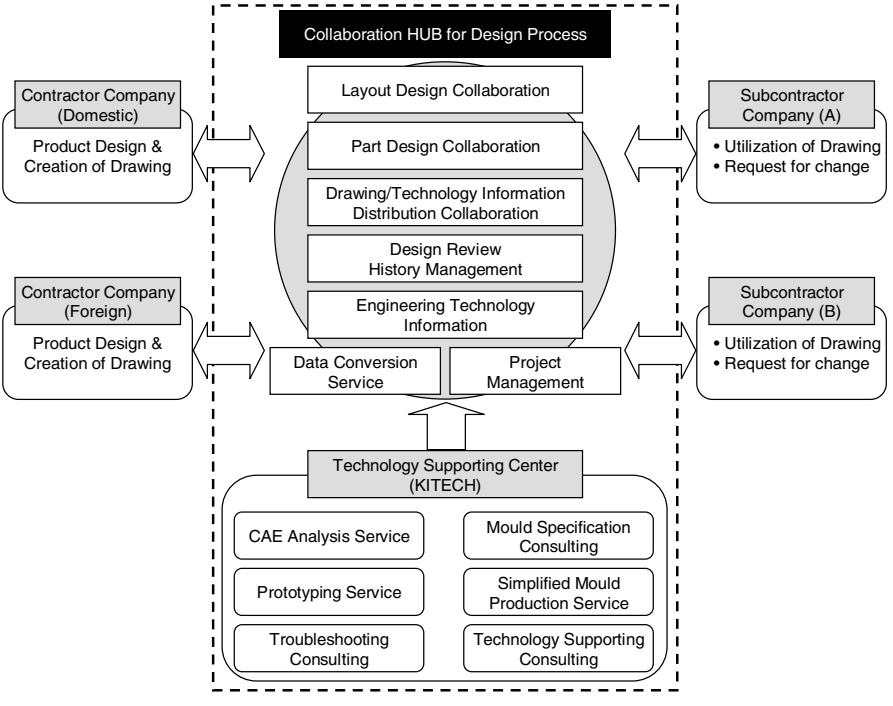


Fig. 1. TO-BE Model of e-Manufacturing project

An e-Manufacturing pilot project for supporting a domestic mould company is now underway for 14 months (2004. 1 – 2005. 2) under the title of "Collaborative Hub Establishment for Design Process." This project will be performed over 10 years from this year. The final goal of e-Manufacturing is the reduction of problems that occur during the design process using collaborative Internet technology. The TO-BE model of the e-Manufacturing project is shown in figure 1. It is a construction of a collaborative hub system that can support collaboration between the principal contractor and subcontractor companies.

2 Necessity and Contents of the Study

In progressing with the e-Manufacturing project, a systematic evaluation framework designed to evaluate the project from the viewpoints of the contractor and subcontrac-

tor companies as users of the hub, the government (MOCIE) and KITECH as operators, is absolutely required. For example, from the viewpoint of the contractor and subcontractor companies, which are users of the collaboration hub, financial success with efficient collaboration will be the most important evaluation index. On the other hand, from the government's point of view, propagation of the result of the e-Manufacturing project is more important than financial success. After all, assurance of national industrial competitiveness will be the most important evaluation index. Next, from the viewpoint of KITECH, which is an enforcement and operation organization, customer satisfaction, security for the hub system, and its maintenance will be the most important evaluation index. Thus, different evaluation frameworks per stakeholder in the e-Manufacturing project are required. Using this evaluation framework, long-term plans will be objectively verified.

Also, to ensure the success of the e-Manufacturing project, step-by-step evaluation must be effectively performed. These steps are divided into the following steps: pre-evaluation, operation evaluation, and post evaluation of e-Manufacturing (Table 1). First, in the pre-evaluation step, a derivation from the evaluation index is required from the user's viewpoint such that the user utilizes outputs of the e-Manufacturing project. Secondly, in the operation evaluation step, real-time monitoring should be performed whether project outputs correspond to the final goal or not. And confirmation of the progress status is needed whether or not evaluation indexes derived are correctly developed. Thirdly, an evaluation about the outcomes achieved by users introducing the collaborative hub must be performed. Finally, feedback of user requests must be performed effectively if the collaborative hub system is to be applied to mould companies.

After all, to evaluate the e-Manufacturing project, a unified evaluation must be executed from different viewpoints (those of the user and operator) and steps (pre-evaluation, operation evaluation and post evaluation) as in Table 1. That is, 12 evaluation systems (A – L) should be integrated and managed.

Table 1. Template about evaluation framework of e-Manufacturing project

Viewpoint / STEP	User		Operator	
	Subcontractor Mold Company	Contractor Mold Company	KITECH	Government
Pre-evaluation	A	D	G	J
Operation evaluation	B	E	H	K
Post evaluation	C	F	I	L

In this study, from the user's viewpoint, research is to be performed on the evaluation framework for contractor mould companies. To develop the evaluation framework, the BSC (Balanced Scorecard) concept is then introduced. First of all, we must establish a vision for the e-Manufacturing project. Then, to achieve this, we must devise a strategy. Ultimately, four viewpoints will be drawn for each BSC with three total scopes (D, E, F), and a measurable CSF (Critical Success Factor) is then to be developed.

3 Related Works on BSC

Over the last 10 years, performance management or measurement systems have been an important focus of interest for both the academic and the societies engaged in the practical application of academic discoveries and innovations. These discussions, in particular, are progressing dynamically based mostly on the concept of BSC (Balanced Scorecard) as suggested by Kaplan and Norton [1, 2, 3]. Companies have found that performance systems dependent solely on financial measurements like ROI and payback period are unable to satisfy all the core fields of interest for a company [4]. As a result, they suggest that financial measures be supplemented with additional ones that reflect customer satisfaction, internal business processes, and the ability to learn and grow. Their BSC is designed to complement "financial measures of past performance with measures of the drivers of future performance" [5]. The name of their concept reflects an intention to keep score of a set of items that maintain a balance "between short- and long-term objectives, between financial and non-financial measures, between lagging and leading indicators, and between internal and external performance perspectives [5]. Table 2 shows the four different viewpoints of BSC, while figure 2 schematically depicts the relationship between the four viewpoints.

Table 2. The four perspectives in a balanced scorecard

Customer perspective (value-adding view)	Financial perspectives (shareholders' view)
Mission: to achieve out vision, by delivering value to our customers	Mission: to succeed financially, by delivering value to out shareholders
Internal process perspectives (process-based view)	Learning and growth perspectives (future view)
Mission: to satisfy our shareholders and customers by promoting efficiency and effectiveness in our business processes	Mission: to achieve our vision, by sustaining our innovation and change capabilities, through continuous improvement and preparation for future challenges

For studies using BSC, the most important task is to convert corporate strategic goals into a series of major performance indicators. For this, the stages for developing a performance system using BSC are as follows: establishment of a corporate vision and strategy, the defining of the company's most important performance area for achieving corporate vision and strategy (financial, customer, internal process, learning and growth), the planning of major activities to be performed in order to achieve the strategic goal of each performance area (CSF decision), and finally, the defining of the KPI (Key Performance Indicator) for measuring CSF. At this point, decisions on CSF and KPI must take into consideration both cause and effect relations.

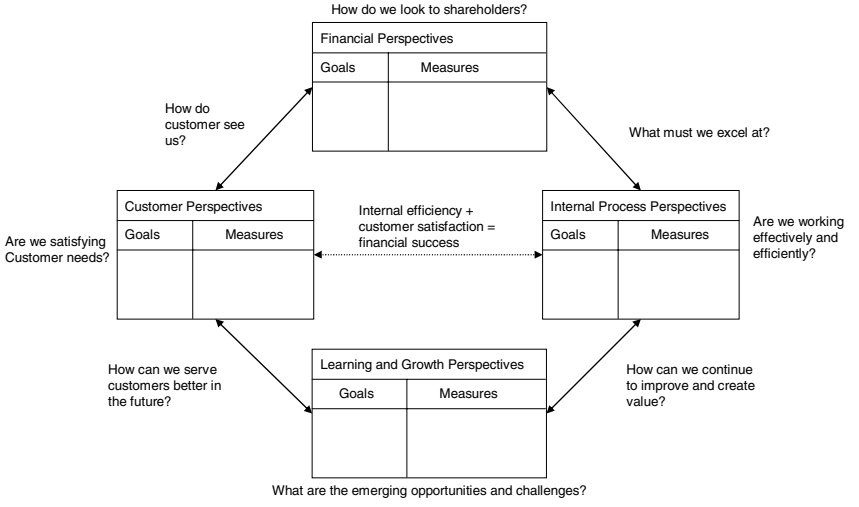


Fig. 2. Relationships between the four perspectives in balanced scorecard ([1])

As a result of analyzing previous studies with regard to BSC, it was found that manufacturers were the largest group to introduce BSC [6, 7]. Some public sector cases of BSC introduction were also found [8]. Particularly, with the rapid advancement in IT (Information Technology), the introduction of IS (Information system) has become a requirement for manufacturers. However, IT and IS introduction is not simply a matter of corporate cost and benefit over a certain period of time but a vital means for corporate survival and growth. Therefore, an accurate evaluation of how IT and IS affect corporate management performance has become an extremely important issue. Nevertheless, IT and IS evaluation has also focused heavily on financial perspectives until now, making it difficult to measure their impact on corporate vision or long-term plans for securing a superior position in market competition. In addition, the introduction and usage of IT and IS contribute less directly to corporate goals, but more directly to the long-term creation of corporate profit, and therefore, it is difficult to evaluate IT and IS.

Unfortunately, evaluation methods that rely on financial measures are not as well suited for newer generations of IT applications. These computer-based IS typically seek to provide a wide range of benefits, including many that are intangible in nature [9]. The introduction of a collaborative Hub system, discussed in this study, may also be considered as an introduction of IS using the latest Internet and IT. In other words, when a domestic mould company introduces the collaborative Hub system, the manager's analysis of possibly introducing a collaborative Hub system, analysis of investment benefits from the Hub system, and the consequent implementation of standards for monitoring the Hub system progress status from a variety of perspectives have become necessary. Also, future-oriented, process-based metrics are seen as a key element in a strategic management system that drives performance improvement and enables the top management team to make well-informed decisions that prepare their organization for the future [10, 11].

As a result, we propose an evaluation model for measurement and evaluation from the four viewpoints of BSC.

4 BSC-Based Evaluation Framework

Figure 3 shows a BSC-based evaluation framework. As mentioned above, this study's evaluation framework skipped the KPI defining stage, providing up to four stages for development. First, the outcome that mould companies want to achieve through the collaborative Hub is benefit maximization. Of course, mould companies may have differing visions, but in reality, the current domestic mould industry experiences a very adverse environment for manufacturing. Due to the surrounding environment of Japan, with an advanced mould industry, and China, with rapidly growing mould industry, domestic mould companies are competing intensively against one another. Thus, their most important goal and vision for its outcome would be securing maximum profit.

Secondly, five strategies were developed for realizing the ultimate vision of the mould industry. Traditionally, the mould industry has considered short delivery and troubleshooting reduction as the most important factors impacting sales. Today, every company is inevitably integrated within a complex supply chain network, playing the roles of both customer and vendor to other companies. Hence, a manufacturing company cannot become agile unless its relationships with the supply chain are also agile [12]. Due to such changes in the manufacturing environment, maximum satisfaction for customers within the industry's supply chain has emerged as the most significant factor. Maximum customer satisfaction typically would lead to increased sales. And such effective collaboration with customers is also considered an important strategy. As the final strategy, the order-securing factor was pointed out, having direct connection to sales.

Thirdly, four perspectives were proposed based on BSC. Among the four perspectives of BSC presented by Kaplan and Norton, financial, customer, and learning and growth perspectives were introduced in the same way. However, the internal process perspective was modified as a unified internal and external perspective. In light of the necessity of collaboration as mentioned above, internal processes as well as external processes of the mould industry should be considered as an important perspective. Furthermore, the IT/IS viewpoint should be included in an internal and external perspective. The perspective bearing on a collaborative Hub system which is an information system, as discussed in this study, should also be considered at the same time.

Fourth, the CSF of each of the four perspectives described above was drawn (Fig. 3). In figure 3, the gray box means addition, change and deletion are possible. In other words, it shows that everything related to BSC can be changed. First of all, the financial perspective used traditional evaluation factors. Next, the customer perspective considered mould delivery and mould quality to be the most important factors. Also, individual customer satisfaction was considered as a CSF. In the internal and external process perspective, the IS utilization mentioned above was included, and the meeting reduction factor for evaluating effective performance through introduction of a collaborative Hub was also considered. And by introducing the collaborative Hub

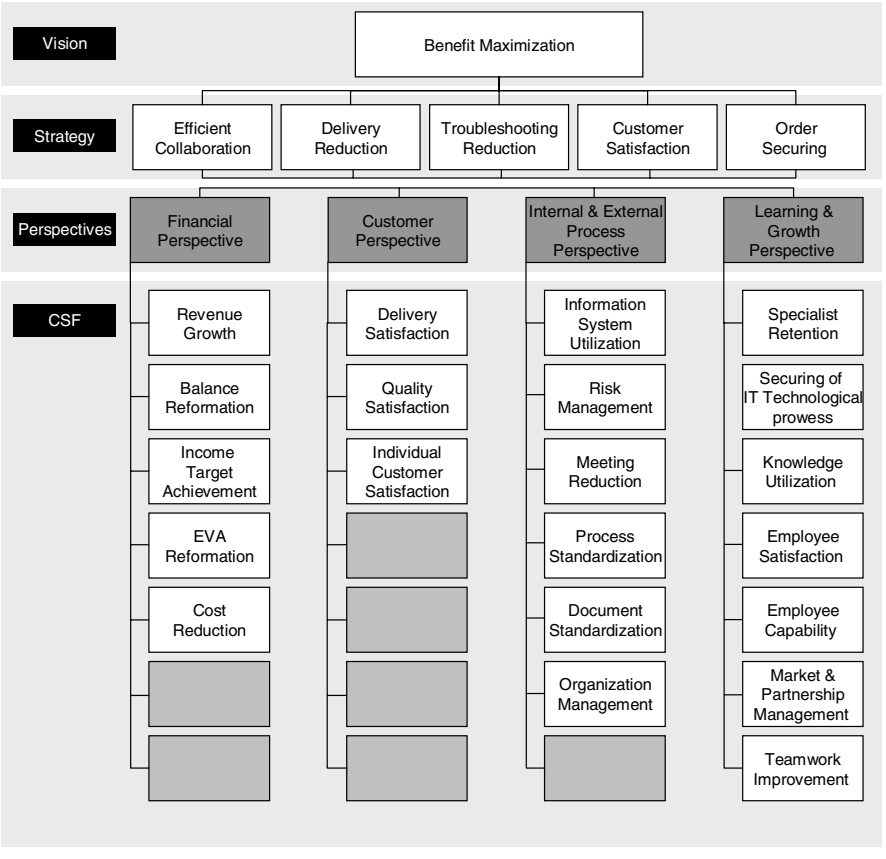


Fig. 3. Evaluation Framework

system, three factors for evaluating process, document management and organization, which are parts that can be changed for mould companies, were added. Also in the final learning and growth perspective, specialist retention, securing of IT technological prowess, and knowledge utilization factors were added. Here, knowledge utilization factor is related to KM (Knowledge Management). Other factors were considered previously in manufacturing industry.

5 Conclusions and Future Works

In this study, a BSC-based evaluation framework was developed for an effective introduction of a collaborative Hub system by the mould industry. First of all, the vision and strategy to be introduced to the mould industry were developed. In this stage, a vision and strategy reflecting the characteristics of the mould industry were built, with particular emphasis on the collaboration factor. Second, in the four existing perspectives of BSC, the internal process perspective was added with external process. Third,

the CSF was developed for each perspective. What is notable is that IT/IS related factors were added through the introduction of a collaborative Hub system, and that factors for evaluating process, document management and organization were added. Furthermore, factors for evaluating KM, which has recently emerged as an important factor in the organization, have been integrated.

For the evaluation framework developed in this study, it is necessary to create the KPI for measuring CSF. The interactive relationship within CSF and KPI should also be analyzed. The evaluation framework developed in this study is targeted at the Contractor Mould Company from the user perspective in the final model mentioned in Table 1. In the future, an evaluation framework for the Subcontractor Mould Company, another user of the collaborative Hub system, as well as a framework from the perspective of KITECH, the operator of the e-Manufacturing project, and the government will need to be developed. The four developed frameworks should be integrated. And finally, pre-, operation, and post evaluation should be performed through an integrated evaluation framework.

References

1. Kaplan, R.S., Norton, D.P.: The balanced scorecard: measures that drive performance. Harvard Business Review 70(1) (1992) 71-79
2. Kaplan, R.S., Norton, D.P.: Putting the balanced scorecard to work. Harvard Business Review 71(5) (1993) 134-142
3. Kaplan, R.S., Norton, D.P.: Using the balanced scorecard as a strategic management system. Harvard Business Review 74(1) (1996) 75-85
4. Merchant, K.: Control in Business Organization. MA, Harvard Graduate School of Business (1985)
5. Kaplan, R.S., Norton, D.P.: The Balanced Scorecard: Translating Strategy into Action. Harvard Business School Press, Boston (1996)
6. Van Veen-Dirks, P., Wijn, Martin,: Strategic Control: Meshing Critical Success Factors with the Balanced Scorecard. Long Range Planning 35 (2002) 407-427
7. Ahn, H.: Applying the Balanced Scorecard Concept: An Experience Report. Long Range Planning 34 (2001) 441-461
8. Irwin, D.: Strategy mapping in the public sector. Long Range Planning 35 (2002) 637-647
9. Martinsons, M., Davison, R., Tse, D.: The balanced scorecard: a foundation for the strategic management of information systems. Decision Support Systems 25 (1999) 71-88
10. Brynjolfsson, E., Renshaw, A.A., Alstyne, M.V.: The matrix of change. Sloan Management Review 38(2) (1997) 37-54
11. Rainer, R.K., Watson, H.J.: What does it take for successful executive information systems?. Decision Support Systems 14(2) (1995) 147-156
12. Lee, R.S., Chen, Y.M., Cheng, H.Y., Kuo, M.D.: A Framework of a concurrent process planning system for mold manufacturing. Computer Integration Manufacturing System 11(3) (1998) 171-190

Design of a BPR-Based Information Strategy Planning (ISP) Framework

Chiwoon Cho[1] and Nam Wook Cho[2]

[1] School of Industrial Engineering, University of Ulsan,
Ulsan, 680-749, South Korea
chiwoon6@mail.ulsan.ac.kr
[2] Department of Industrial and Information Systems Engineering,
Seoul National University of Technology, Seoul, 139-743, South Korea
nwcho@snut.ac.kr

Abstract. In this paper we propose a BPR-based Information Strategy Planning (ISP) Framework, which aims to fully incorporate the BPR concept into ISP. Since business processes and information systems (IS) are tightly linked together, business processes need to be reengineered during information systems planning. Five sub-processes are defined in the framework – Business Strategy Analysis, Process Analysis and Redesign, IS Analysis and Modeling, Organization Analysis, and ROI (Return On Investment) Analysis and Integrated Execution Planning of IS. With this framework, it is expected that enterprises will be able to perform more effective, efficient, and strategic IS planning.

1 Introduction

The importance of Information Systems (IS) has rapidly increased in recent years, and the strategic use of IS has become one of the critical success factors (CSF) of many companies [4, 16]. Therefore, most enterprises pursue the renovation of business processes and strategies through IS [10, 11]. During the last decade, a number of leading companies have implemented solutions such as ERP, SCM, CRM, PLM and so on, all of which inevitably require a large capital investment. To make the investment more effective, Information Strategy Planning (ISP) has been widely adopted prior to the implementation of IS. The aim of ISP is to align business strategy with information system planning [17]. It identifies prioritized information systems that are efficient, effective and strategic [1]. It is true that ISP facilitates systematic IS master planning; however, some companies still consider their IT investment to be unsatisfactory, which often results in overall poor IS performance. The overall poor performance principally results from ISP being executed without a fundamental redesign of business processes [13], which shows that the goal of ISP cannot be achieved by simply automating existing business processes [5].

Of course, existing ISP methodologies include the BPR concept to some extent. However, since they mainly focus on the identification of prioritized IS planning rather than business reengineering, business processes are analyzed in order to identify IS requirements. Thus, redesign or reengineering of business process has been often neglected.

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1297 – 1305, 2005.

Much research has been conducted on ISP, but the majority of it has focused on improvement and evaluation from the IS perspective [10, 11, 14, 15]. Despite the importance of business processes in IS planning, little ISP research has been conducted from the BPR perspective. Taylor [16] attempted to link BPR and SAP, an ERP Solution, by using a Sociotechnical System (STS). NG *et al.* [13] proposed a framework that integrates BPR and ERP.

To overcome the limitations of current ISP methodologies, this research aims to provide an ISP framework that incorporates the BPR concept. The new approach proposes that business processes and information systems should be linked together during the planning stage of enterprise information systems, and that these processes need to be redesigned or reengineered during ISP. This study also provides an ROI analysis flow, which is a critical step for the economic justification of an IT investment.

2 BPR-Based ISP Framework

Figure 1 depicts the four key components of an enterprise ISP: Business Strategy, Business Process, Information System, and Organization. To maximize the effectiveness of ISP while minimizing the loss resulting from inadequate output, these four key components should be fully considered during the entire planning process.

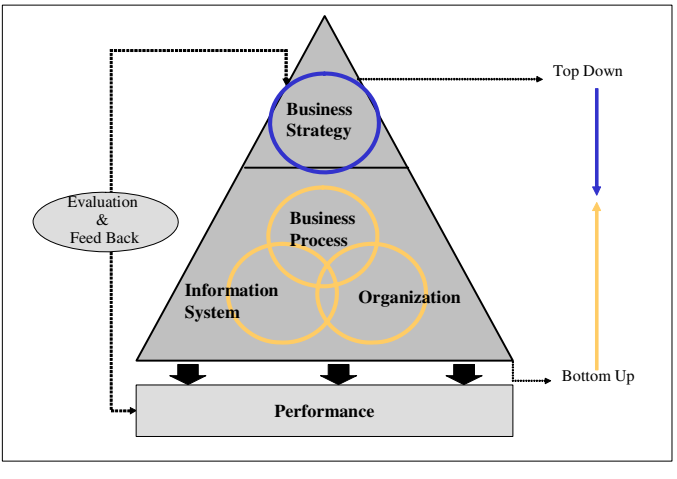


Fig. 1. Components of information strategy planning

The business strategy component aligns enterprise strategy with IS planning from a top down perspective. The business process, information system, and organization components constitute the bottom layer in which all of the components are analyzed and designed concurrently. The proposed ISP framework considers the four key components simultaneously to derive successful ISP results. The links and interaction

between the business process and information system components are focused upon in the framework, which is why it is called a BPR-based ISP framework. It aids in the design and planning of IS under a BPR context, and supports the transformation of process redesign results into IS implementation.

2.1 Key Features

A framework is proposed to reinforce the BPR concept while retaining the advantages of existing ISP methodologies. The proposed framework, as compared to ISP methodologies (IE-Expert, METHOD/1, INNOVATOR, IPR) developed by IS solution and consulting companies, consists of supplementary sub-processes with a different structure. Three key features are embodied in the proposed framework, and these features differentiate our framework from other ISP methodologies.

First, the framework requires that ISP should be tightly coupled with BPR. Business process reengineering (BPR), the fundamental rethinking and radical redesign of business processes, is closely related to the process of IS planning. IS planning can hardly be successful without the comprehensive redesign of business processes. Second, a sub-process is added to analyze the IS department. Despite the importance of the IS department, its analysis is under-emphasized in existing methodologies. The framework proposes that the role of the IS department should be evaluated and the department should be reorganized during ISP. Third, as most enterprises require ROI analysis prior to IT investment, an ROI analysis flow for the quantification of investment effectiveness is proposed in the framework. Table 1 shows a comparison of the proposed framework and other ISP methodologies.

Table 1. A comparison of the proposed framework and other ISP methodologies. IE-Expert is originated from James Martin's Information Engineering Methodology. INNOVATOR is an ISP framework developed by Samsung SDS, Korea

	Proposed Framework	IE-Expert™	INNOVATOR™
Enterprise model	Focus on data, technology, process, and organization	Focus on data, technology, and process	Focus on data, technology, and process
Applicability of BPR	Yes	No	Limited
Applicability of IS department analysis	Yes	No	No
Availability on economic valuation	Yes	No	No
Time horizon (months)	5-7	8-10	3-4

2.2 Framework

The proposed framework largely consists of five sub-processes: Business Strategy Analysis, Process Analysis and Redesign, IS Analysis and Modeling, Organization Analysis, ROI analysis and Integrated Execution Planning. A structure of the framework is illustrated in Figure 2. Of the five sub-processes, two remain similar to previous ISP methodologies: Business Strategy Analysis and IS Analysis and Modeling. The other three sub-processes are redesigned.

CSFs are critical inputs for Process Analysis and Redesign, IS Analysis and Modeling, and Organization Analysis. Therefore, these sub-processes are followed by the Business Strategy Analysis. The sub-processes, and the performance of each activity pertaining to them, generate the backbone of the final integrated execution of IS planning.

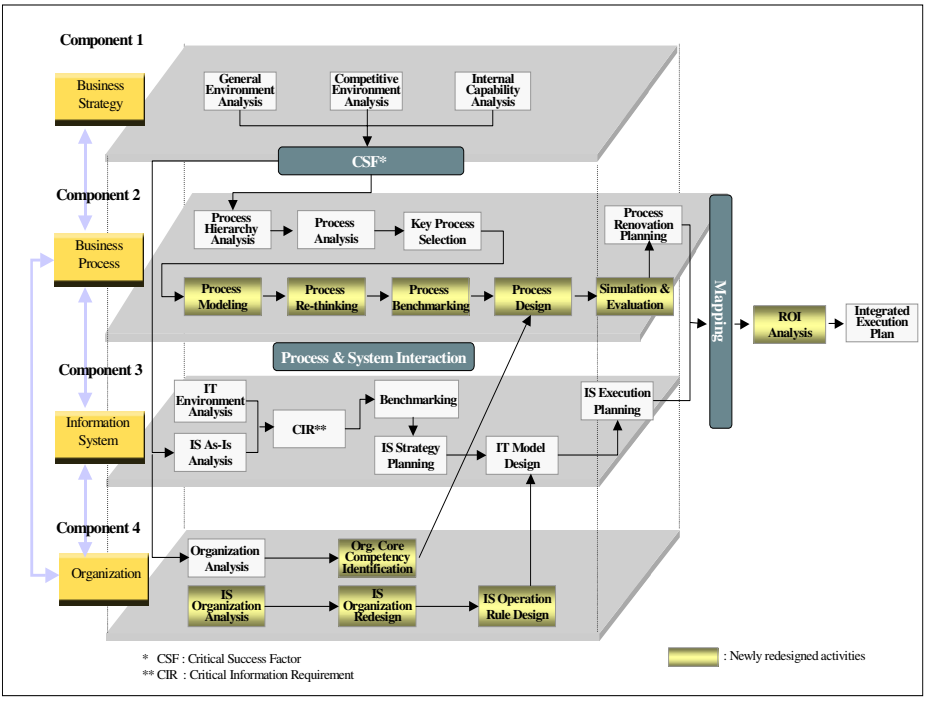


Fig. 2. Framework of BPR-based ISP

2.2.1 Business Strategy Analysis

Business Strategy Analysis is the sub-process for identifying the business strategies of the corporation. Distinct business strategies are identified as based on the analysis of the business environment, internal capacity, and CSF (Critical Success Factors). The Business Strategy Analysis sub-process is divided into four activities: general environment analysis, competitive environment analysis, internal capability analysis, and identification of CSF. The following is a description of the four activities.

General Environment Analysis. The business environment needs to be analyzed to identify the current situation of the company. The business environment can be categorized as a remote environment, which includes political, social and economic components, as an industry environment, and as a market environment [15].

Competitive Environment Analysis. The factors that can be considered in the competitive environment analysis are threat of mobility, segment rivalry, buyer's bargaining power, supplier's bargaining power, and threat of substitutes.

Internal Capacity Analysis. Value chain analysis and BSC (Balance Scorecard) analysis can be applied to analyze internal capabilities.

Identification of CSF. Based on the analysis of the external and internal business environments of the corporation, the CSF can be identified. The CSF identified at this phase are used for identifying the business strategy [1].

2.2.2 Process Analysis and Redesign

Accomplishment of process analysis and redesign is based on the CSF identified in the previous sub-process. The results of process analysis are used not only for business process reengineering, but also for IT model design in the IS analysis and modeling sub-process. The process analysis and redesign sub-process can be divided into nine activities, as shown in Figure 2.

Process Hierarchy Analysis & Process Analysis. The process analysis can be completed by auditing the flexibility, accuracy, connectivity and efficiency of business processes. The process designers must completely familiarize themselves with the current processes of the corporation [3].

Key Process Selection. Due to practical limitations, not all processes that have problems can be improved. Therefore, the processes most in need of improvement are selected as key processes.

Process Modeling & Process Re-thinking. The process designers should query the assumptions that underlie current processes in order to design new and improved processes. In this way, a set of processes can be identified for process benchmarking.

Process Benchmarking & Process Design. This provides the company with examples of best practices and offers distinct guidance in terms of new processes and their manner of implementation. Modeling techniques can be employed to map out the logical flow of processes [6].

Simulation & Evaluation. Simulations are used to visualize and evaluate the redesigned processes. After evaluation and modification, new processes are finally established and are ready to be used in the IS design stage for process and system interaction.

Process Renovation Planning. After generating new processes, the project team that will carry out process renovation is organized. The team makes a plan that includes a time horizon, detailed activities, roles, a performance index, and a brief of the techniques that will be followed in conducting the project.

2.2.3 IS Analysis and Modeling

IS Analysis and Modeling is the sub-process where IT opportunities and strategies for IS implementation and investment are identified. As illustrated in Figure 2, the sub-process can be classified into seven activities, which are described as follows.

IT Environment Analysis. The IT environment is analyzed to identify IT trends in similar business environments and new capabilities of IT [11].

IS As-Is Analysis & Identification of CIR. By these activities, the current IS is evaluated to see how it can support the CSF identified in the business strategy analysis sub-process. Additionally, Critical Information Requirements (CIR) are generated. The process of specifying CIR can be quite tedious, but its importance should not be underestimated [14].

Benchmarking & IS Strategy Planning. Benchmarking is an excellent tool for continuous improvement. It is used to gain information regarding IS and strategies of top-tier companies. IS strategies are established at this stage, and are based on benchmarking and the analysis of business strategies and current IS.

IT Model Design & IS Execution Planning. After obtaining IS strategies and CIR, the specifics for IT model design and implementation of strategic IS are defined. Information architecture and IT solutions are decided through IT model design activities. The actual sequence for IS development is established during the IS execution planning activities. Also, investment is estimated after consideration of key criteria such as economic and competitive advantages [1, 13]. The IS execution plan and the generated process renovation plan are combined before ROI analysis is conducted.

2.2.4 Organization Analysis

This is the sub-process where all organizations including the IS department are analyzed and evaluated. The core competency and roles of the organizations are also identified. The sub-process can be divided into five activities as shown in Figure 2. The following is a description of the five activities.

Organization Analysis & Core Competency Identification. Through these activities, organizations and their business process structures are analyzed. The core competency of each organization should be identified based on relevant circumstances concerning the CSF that were generated in the business strategy analysis sub-process. The results are referred to in the process design activity of the process analysis and redesign sub-process.

IS Organization Analysis & Redesign. In conventional ISP methodology, there is not any activity or stage for analysis of the functional structure of the IT department. The functions and roles of the IT department should be evaluated and newly identified in accordance with the IS strategies generated during the execution of the IS analysis and modeling sub-process.

IS Operation Rule Design. For the effectiveness of IT department, IS operation policies should be set. For example, the functions and tasks between local help desk and global help desk should be differentiated. The IS operation rules of the IT department are referred to in the IT model design activity of the IS analysis and modeling sub-process.

2.2.5 ROI Analysis and Integrated Execution Planning of IS

After combining the process renovation plan with the IS execution plan, IT ROI analysis is performed. Enterprises invest in IT to raise their level of competitiveness in the market, but they don't make investments unless distinct objective results are provided [12]. In the proposed framework, an ROI analysis flow is designed to yield

qualitative economic effects through informatization according to various scenarios. The details are described in the next section.

ROI Analysis. This activity consists of conducting an analysis of investment effectiveness of the enterprise informatization.

Integrated Execution Planning. The actual sequence and detailed plan for the IS development that accompanies process innovation are decided through this activity.

2.3 ROI Analysis

As most companies require economic justification prior to IT investment, ROI analysis has become a critical component of ISP. In the proposed framework, an ROI analysis module is composed of four steps: IS investment assessment, performance indicator identification, data collection, and ROI calculation.

Step 1. *Assess IS Investment:* As the first step of ROI analysis, the total investment amount is assessed based on IS investment planning. Cost factors including hardware cost, man-hours, and the investment horizon are fully investigated. As a result, a year-by-year investment schedule is generated.

Step 2. *Identify IS Performance Indicators:* Based on critical success factors, performance indicators of IS are identified in Step 2. In this step, performance indicators of IS are categorized into three groups – the cost saving, the productivity benefit, and the business benefit. Tangible cost saving such as reduced cost of labor, material, and maintenance can be classified as the cost saving indicators. The productivity benefit indicators can be identified from the increased productivity of a business value chain. The business benefit reflects intangible and strategic aspect of IS benefit, which can be selected to raise the competitiveness in the market and to improve the customer relationship. Increased market share and enhanced customer satisfaction are typical examples of the business benefit indicators.

Step 3. *Collect Data:* This is the most difficult part of an ROI analysis, as the future effectiveness of information systems has to be predicted during the planning stage. Step 3 constitutes the basis of the ROI analysis. Data such as sales, profit margin, employee salary, and number of employees using a specific information system can be obtained from internal sources. Also, performance guideline data for the calculation of ROI must be collected through literature review or case studies. Since a new information system has yet to be implemented, one needs to estimate the effectiveness of the information system based upon case studies. For example, let us suppose productivity savings of a Knowledge Management System (KMS) per employee could be as much as 2 hours per day for one company or less than a few minutes for another. A company considering the implementation of a KMS can use such data as performance guidelines.

Step 4. *Calculate ROI:* For each performance indicator, this step requires the conversion of IS performance indicators to monetary values. This requires a conversion formula for each performance indicator, which is defined and the data for which are collected at step 3. What-if analysis is conducted under assumed optimistic, moderate, and pessimistic scenarios. In addition to traditional ROI, Net Present Value (NPV),

Internal Rate of Return (IRR), and Payback Period (PB) are also evaluated to determine the economic value of an IS investment.

3 Conclusion

The enterprise information system has become a business process innovation tool, and it has been shown that ISP can hardly be successful without comprehensive business process reengineering. To reinforce the business process perspective in ISP, this paper introduces a BPR-based ISP framework. The new framework aims to reflect the effects of process innovation in the planning of an IS road map. An approach to ROI for IT investment justification is also provided. Even though the proposed framework is intended to fully incorporate the BPR concept into the ISP framework, a full scale BPR may not be possible during ISP due to time or budget constraints. In this case, the framework recommends that the level of process depth can be reduced during the Process Analysis and Redesign sub-process while maintaining the BPR-based ISP concept. With this framework, it is expected that enterprises will be able to perform more effective, efficient, and strategic IS planning.

Acknowledgement. This research was supported by a research grant from the University of Ulsan in Korea.

References

1. Baker, B.: The role of feedback in assessing information systems planning effectiveness. Journal of Strategic Information Systems, 1995. 4(1): p.61-80
2. Davenport, T. H.: Process innovation: Reengineering work through information technology. Harvard Business Press, 1993. Boston, USA
3. David, K. C., Henry, J. J.: Best practices in reengineering: What works and what doesn't in the reengineering process. McGraw-Hill, 1995. New York, USA
4. Guimaraes, T., Bond, W.: Empirically assessing the impact of BPR on manufacturing firms. Int. J. Operations & Production Management, 1996. 16(8): p.5-28
5. 5. Hammer, M.: Reengineering work: Don't automate, obliterate. Harvard Business Review, 1990. July-August: p.104-112
6. Hammer, M., Champy, J.: Reengineering the corporation: A manifesto for business evolution. Nicholas Brealcy, 1993. London, England
7. Hong, S. W., Park, G. H., Seo. H. J.: IT ROI. Daechung Media, 2004. Seoul, Korea
8. Kim, S., Leem, C. S.: An information engineering methodology for the security strategy planning. ICCSA, 2004.: p.597-607
9. Lee, S. C., Hong, J. W.: Analysis of the indices for economic effects through informatization according to industry types. IE Interface, 2001. 14(4): p.421-428
10. Leem, C. S., Kim, S.: Introduction to an integrated methodology for development and implementation of enterprise information systems. J. Systems and Software, 2002. 60: p.249-261
11. Leem, C. S., Kim, I.: An Integrated Evaluation System based on the Continuous Improvement Model of IS Performance. Industrial Management and Data Systems, 2004. 104(2): p.115-128

12. Leem, C.S., Yoon, C.Y., Park, S.K.: A Process-Centered IT ROI Analysis with a Case Study, Information Systems Frontiers, 2004. 6(4): p.369-383
13. NG, J. K. C., IP, W. H., Lee, T. C.: A paradigm for ERP and BPR integration. Int. J. Production Research, 1999. 37(9): p.2093-2108
14. Premkumar, G., King, W. R.: The evaluation of strategic information system planning. J. Information and Management, 1994. 26: p.327-340
15. Sergas, A. H., Grover, V.: Strategic information systems planning success: an investigation of the construct and its measurements. MIS Quarterly, 1998. 22(2): p.139-163
16. Taylor, J. C.: Participative design: linking BPR and SAP with an STS approach. J. Organizational Change Management, 1998. 11(3): p. 233-245
17. Zani, W.M.: Blueprint for MIS. Harvard Business Review, 1970. 48(6): p.95-100

An Integrated Evaluation System for Personal Informatization Levels and Their Maturity Measuement: Korean Motors Company Case

Eun Jung Yu, Choon Seong Leem, Seung Kyu Park, and Byung Wan Kim

Information & Industrial Engineering, Yonsei University, Korea
IT Research and consulting, Korea
alex@itr.re.kr

Abstract. It is an incontestable fact that an enterprise's competitiveness these days is reflected by whether an individual has the application capability and knowledge related to information technology. This study improves upon the previous research of evaluation systems for personal informatization level and its maturity measurement. To measure the personal informatization level, three types of questions are developed: Answer-Driven Question(ADQ), Personality - Fit Question(PFQ) and Understanding-Oriented Question(UOQ). These three types are fitted to the characteristic of evaluation indices. Evaluation indices consist of three evaluation domains: Mind of IT, Knowledge of IT and Application of IT. Each domain is composed of three evaluation factors and their specific items. They have cause-and-effect relation among them. We applied this evaluation system to individuals working in an informatization environment at a representative Korean motor company, and verified the application and practicality of evaluation systems through presenting the evaluation results of 18,898 individuals.

1 Introduction

As human resources have come to be viewed as more critical to organizational success, many organizations have realized that it is people in an organization that can provide a competitive advantage [1]. These days, the enterprise environment, which is built on information technology, is also increasing the range and investment of information technology's application in the enterprise process, and considers information technology to be one of the most important elements for improving competitiveness of organizations. Many viewpoints have been presented as being a primary factor for informatization success, such as alignment to strategy, IT investments that take into consideration work process quality, superiority of information system to innovate processes, and quality of system and organization's culture that can maintain a foundation for informatization. But most importantly, in today's end user computer (EUC) environment, the bottom line is not how good information systems (IS) are, but rather how well they are used; how well these systems are used is a function of EUC skills, which empower individuals to utilize IT and perform a variety of functions in organizational context [2]. The members of enterprises are the subjects who produce prod-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1306 – 1315, 2005.

ucts (or services) of real value, using the processes and information technology being invested for the creation of value, and in the end, it is the well-trained members of an organization who can produce great products and services that can effect the outcome of the enterprise [3].

Accordingly, enterprises feel the need for practical tools that can measure and manage an individual's capability for informatization, and related researches have been carried out. However, the researches done so far are only at the level of systematically arranging these thoughts, and they cannot yet present a model or management tool that can be applied practically in enterprises. In this paper, we will present a case where we developed and applied a D-EIPI (Dynamic Evaluation Indices of Personal Informatization) model, which overcomes the limitations of existing researches. The case study involving 18,000 staff members of 'H Co.' in Korea, a Fortune 500 global motor company, verified our research model and gave us various discussion points. The Integrated Evaluation System for Personal Informatization Levels developed in this research can become a guide for improving the personal informatization levels in future enterprises, and be used as a quantitative index in making deliberations and developing roadmaps for training investments.

2 Previous Research

2.1 Personal Informatization Level Evaluation Domain

Personal informatization level is a person's ability to act on information related to his or her capability. According to Rodriguez., et al, a person's capability refers to knowledge, skill, ability, and action – in forms that can be measured – which are needed for the person to successfully perform a job function or role [4]. Furthermore, according to S.C. Parry, a person's capability can even include attitudes which can influence the results [5]. Accordingly, Personal Information Level can be classified as measurable forms of knowledge level, application capacity, and attitude needed by the person to successfully perform his or her essential role using information technology. There are existing researches on evaluating a person's capability such as O'Driscoll & Eubanks(1993), Jenkins & McHarg(1998) and Rifkin(1999), but their works are either limited to specialized types of workers or deficient as a comprehensive evaluation system that satisfies the definition of Personal Informatization Level as mentioned above.

When we search among the existing researches in the industry and academia for systems that evaluate the informatization capacities of organization members, we can mention the evaluation system for computer competency by McCoy(2001), and the system for evaluating informatization levels of staff and new employees at such places as the Korean MIC (Ministry of Information and Communication) and Samsung SDS. These evaluation systems are one-time evaluations that do not consider job positions or functions, but in contrast, Yoon(2003) developed a model with 13 detailed evaluations that considered the characteristics of staff's job functions and positions, and carried out a research that connected the resulting levels of each individual to a Maturity model.

Table 1. Existing research regarding organization staff informatization level evaluation

Groups	Brief	Details
McCoy (2001)	Computer skills	Evaluate general knowledge level for hardware, software, programming & integration
The Ministry of Information and Communications, Korea (2003)	Staff informatization levels	Evaluate IT policy mind, IT technology comprehension, IT usage capability & IT Growth efforts
Samsung SDS (2003)	Informatization level of new employees	Evaluate practical sense for information technology, using information I (HTML, programming languages etc.) & information II (OA)
Yoon (2003)	Informatization level classified by job position / function	Evaluate mind of IT, knowledge of information technology, using information technology & information growth ability

2.2 Personal Informatization Level Evaluation Methods

In order to organize the methods used when evaluating personal informatization level, we can look at the trends according to evaluation methods and types of indicators. Evaluation methods are divided into Questionnaire test, which usually consists of several types of questions, and Application test, which directly evaluates one's ability to use information technology. When the Questionnaire test is examined according to the types of indicators, it generally divides into a dichotomous question test, selective question test, and test with questions based on level of measurement. Fig. 1 shows a diagram that connects the methods mainly used in each evaluation area classification and types of indicators.

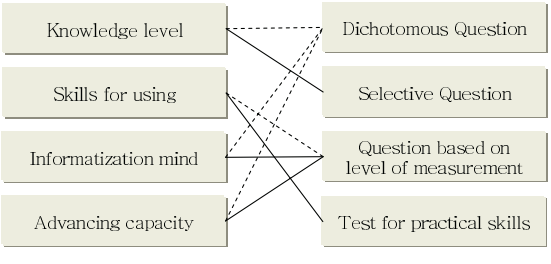


Fig. 1. Measurement type trends classified by evaluation areas

Selective question test, in which one selects several questions, is used for general knowledge level diagnosis, and to diagnose each individual's capability for using skills, test for practical skills is mainly used. However, test with questions based on level of measurement, in which one selects established measurements, is mainly used for a very sincere individual's mind or an advance ability area, due to the special characteristics of such evaluation area [6], [7].

2.3 Limitations of Existing Research

Whereas the existing personal informatization level researches mainly focused on defining a relationship between a person's competency and the person's informatization competency, more recent researches are suggesting various evaluation areas and measurement methods to structuralize and evaluate personal informatization levels, and are trying to apply the measured results to managing systems. However, recent researches cannot sufficiently explain the connections among the factors of personal informatization level.

There has to be an evaluation system where the indices of each evaluation area are linked with cause-and-effect relations in order for an HR manager to take the results of the evaluation, find out the present levels and their causes, and assign priorities for improvements using limited resources to enhance the levels. In other words, after an Informatization Level evaluation, the information about the cause-and-effect relations among the evaluation areas are crucial in suggesting a guideline for level enhancement, but most current researches are insufficient in their solutions. In addition, in terms of measurements, a measurement method that conforms to the evaluation contents must be introduced, but at present, an indiscriminate measurement method that ignores evaluation areas is mainly being applied. In this paper, we will present the D-EIPI model, which overcomes these problems.

3 D-EIPI (Dynamic-Evaluation Indices of Personal Informatization)

D-EIPI (Dynamic-Evaluation Indices of Personal Informatization) is a comprehensive model for evaluating and measuring Personal Informatization Levels, and it contains a system of measurement types specialized for evaluation areas and evaluation indices that are made up of cause-and-effect relations and classified by evaluation areas. The word 'Dynamic' refers to the fact that the model's Informatization Level is measured through the cause-and-effect relations causing one index variable to affect another variable index, and it signifies that these indices are systemized as measurement types that reflect the characteristics of areas.

3.1 Cause-and-Effect Relation Model for Areas

Dynamic model means a model where the cause-and-effect relations among the three mutually exclusive and inclusive areas, the mind of IT, knowledge of information technology, and application of information technology, are set up. The aptitude and attitude to receive and diffuse personal informatization, the interest in informatization plans eventually generate the motive to investigate knowledge regarding information foundation technologies such as OA and OS, application technologies and trends. A person's application ability that corresponds to the person's mind of IT and knowledge level gets cultivated, and eventually, it leads to enhancement of personal informatization level and job accomplishments. Furthermore, when the enhanced personal informatization level creates job accomplishments, it ultimately strengthens the inclination to participate in informatization activities, and it in turn, results in the elevation of knowledge and use of information technology.

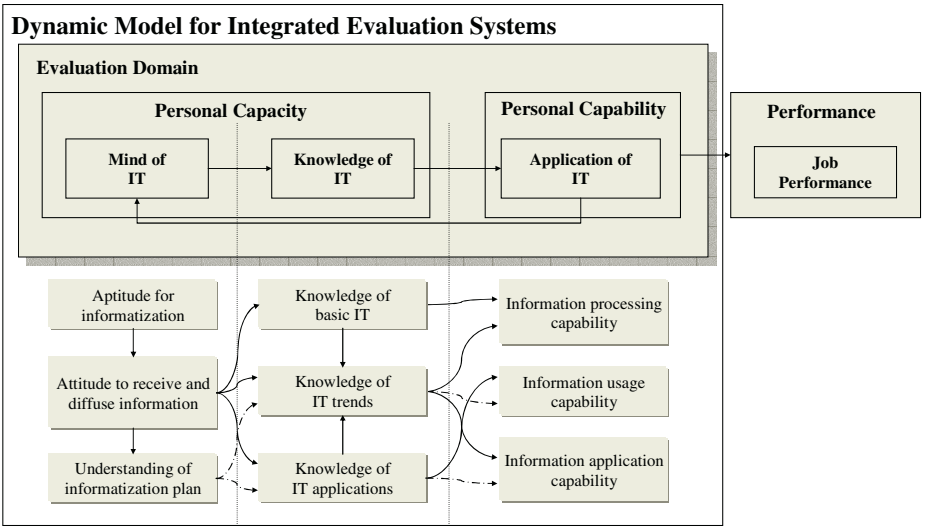


Fig. 2. Dynamic model for diagnosing personal informatization level

This cause-and-effect model can reduce the logical jump found in the existing static models that focus on the past-oriented evaluation of resulting status, and can strengthen the explanations for the relations among the factors of personal informatization level. With this model, an enterprise can forecast the progress of personal informatization and find a quantitative basis to construct directions for improving personal informatization levels in the future. Table 2 shows three evaluation areas and their sub-categories, developed from the basis of logic stated above.

Table 2. Three evaluation areas and sub-categories

Evaluation area	Sub-categories	Evaluation subject indices
Mind of Information Technology	Understanding of informatization plan	Understandng informatization goal, strategy, and effects
	Attitude to receive and diffuse information	Receiving, diffusing, and managing information
	Aptitude for informatization	Informatization leadership, participation, and ethics
Knowledge of Information Technology	Knowledge of basic IT	H/W, S/W, N/W and DB
	Knowledge of IT applications	Solutions, IT methodology, e-Business
	Knowledge of IT trends	H/W, new S/W technology, N/W, new DB technology, new e-Business technology
Application of Information Technology	Information processing capability	Capability for using OS, OA, and the Internet
	Information usage capability	Using IT infra, IT methodology, and e-Business
	Information application capability	Application within enterprise, among enterprises, and to customers

3.2 Measurement Types Specialized to Specific Evaluation Areas

The types of indices used in diagnosis and evaluation of levels in existing researches can be divided into two main categories: one is 'ADQ (Answer-Driven Question),' which leads respondents to select correct answers, and the other is 'PFQ (Personality - Fit Question),' in which multiple questions are set up as measurements to understand the characteristics of respondents.

The diagnosis areas of EIPI, however, because it diagnoses the understanding level of enterprise informatization status as well as personal knowledge levels and characteristics, it must contain not only ADQ and PFQ types but also include another type that can diagnose the understanding level. The integrated index system developed in this research contains three main types of indices, shown in table 3.

Table 3. System of measurement types

Question Types	Features	Examples
ADQ (Answer-Driven Question)	-question types for absolute knowledge -answers that can be classified as clearly correct or wrong -objective types	◎ Among the electronic interaction types, what is correct explanation for B2C? 1) electronic business between consumer and consumer 2) electronic business between enterprise and government 3) electronic business between government and consumer 4) electronic business between enterprise and customer
PFQ (Personality-Fit Question)	-question types for personal informatization aptitudes and inclinations -question types in which measurements for inferiority / superiority are set up -objective /aptitude test types	◎ I always help coworkers who are not used to doing jobs involving information systems 1) Definitely No 2) Usually No 3) Sometimes 4) Usually Yes 5) Definitely Yes
UOQ (Understanding Oriented-Question)	-question type for person's understanding of informatization environment in the company -question types that consider environmental factors of respondents (all answers are close to correct answers, but only one or two answers that most closely reflect the person's company environment are considered correct) -Objective types	◎ Select two answers that you think are the most effective results that can be acquired through your company's CRM? 1) Provision of fair service benefits for all customers 2) Prevention of customer secession 3) Strengthening of customer service using new IT technology 4) Maintenance of relationship with bad customers 5) Inducement to Re-Sell

UOQ(Understanding-Oriented Questions) refers to question types that have no right or wrong answers but make the respondents select the answers that fit the respondents' context. This is a newly developed index type that measures the respon-

dents' capability for using and applying information technology that fits their department and company environment, and their understanding level of their company's informatization strategic plans.

These index types reflect the characteristics of measurement indices and are applied separately. Fig. 3 shows the mapping results of the index types among the measurement indices.

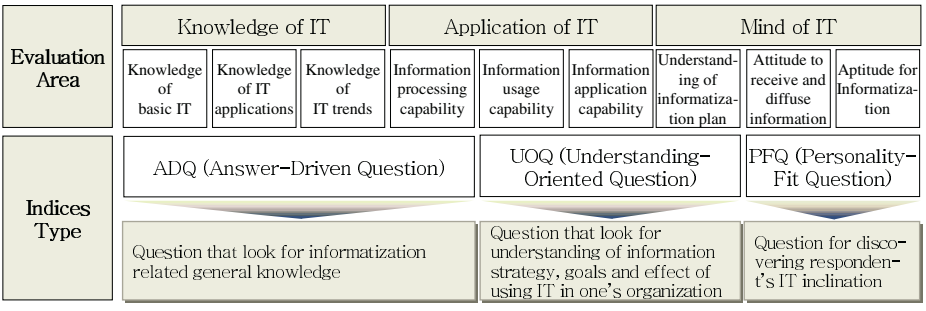


Fig. 3. Relations between evaluation areas and index types

Because the index system in D-EIPI can reflect the characteristics of various areas that need to be considered in measuring personal informatization level, it has the advantage of being able to include not only the quantitatively measurable personal knowledge level but also the qualitative personal inclinations and understanding level as its objects of diagnosis.

4 Case Study – Korean Motors Company Case

D-EIPI was applied to 18,898 headquarter staff members of 'H' company, a representative Korean motor company, which ranks 180 worldwide and is one of Fortune 500 (2003). Executives and regular employees were all included in the 18,898 potential respondents, and among them, 17,246 people responded, showing a high percentage of participation at 91.7%. The questionnaire for measuring personal informatization levels contained a total of 33 questions for mind of IT, 27 questions for knowledge of information technology, and 18 questions for use of information technology, and the diagnosis was carried out 16 times in real-time using a web-based online system.

Regression analysis was carried out on the area scores for the 17,246 staff members to verify the cause-and-effect relations between three evaluation areas and nine sub-categories, and Fig 4 shows that the cause-and-effect relation model for evaluation areas proposed by D-EIPI is proper. When the Pearson Correlation values marked on each link in Fig. 4 are examined, we can see that they generally show high correlations (over 0.4), and their significance levels were all 0.000, which show them to be highly significant. When Durbin-Watson statistic values were analyzed as verification statistic values to verify the auto-correlation tendencies, they were generally close to 2 (minimum 1.830, maximum 1.985), meaning auto-correlation can be neglected.

Knowledge of IT trends is shown to have relatively low correlations to other areas because it is optional knowledge separate from the information technology about to be realized or being used in the respondents' environments.

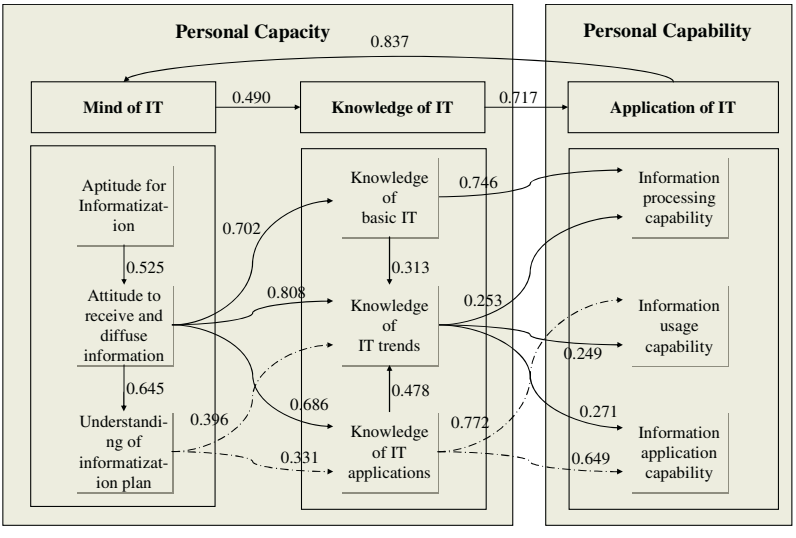


Fig. 4. Correlations among areas of D-EIPI

Accordingly, a high mind level of IT becomes a motivation for increasing the IT knowledge level, which ultimately results in a high rate of information usage by organization members.

The results below in table 4, where the scores of the IT usage areas from the model that was applied to the H company employees are divided into 10% grading steps and the averages for mind and knowledge of IT are organized for each grade step, show all the more clearly the cause-and-effect relations between the areas.

Table 4. IT usage levels compared with the other areas (10% grade step comparison)

Percentage	Application of IT	Knowledge of IT	Mind of IT
low 10%	36.9	49.2	32.7
10%~20%	50.3	57.5	44.4
20%~30%	55.4	61.2	49.6
30%~40%	59.8	62.0	53.1
40%~50%	63.1	64.8	56.7
50%~60%	66.4	65.5	59.6
60%~70%	69.9	67.2	62.1
70%~80%	73.6	67.8	65.5
80%~90%	77.7	70.8	69.5
high 10%	85.2	73.3	76.6

Also, a high level of usage can serve as a momentum for each individual to take part again in informatization activities, thus increasing the level of informatization aptitude, and eventually accomplish a maturity flow that cycles from mind, knowledge, Application to mind again.

Table 5. IT usage levels compared with informatization aptitude (10% grade step comparison)

Percentage	Application of IT	Informatization Aptitude
low 10%	36.9	36.2
10%~20%	50.3	46.6
20%~30%	55.4	50.3
30%~40%	59.8	53.3
40%~50%	63.1	56.3
50%~60%	66.4	58.1
60%~70%	69.9	59.9
70%~80%	73.6	61.2
80%~90%	77.7	65
high 10%	85.2	68.8

5 Conclusions and Future Research

In this paper, the importance of personal informatization level in current enterprise environments that are based on information technology was emphasized and the D-EIPI (Dynamic Evaluation Indices of Personal Informatization) model was developed to systematically measure it. When the model was applied to 18,000 staff members of 'H Co.' in Korea, one of Fortune 500 and global motor company, the case verified our research model and gave us various discussion points.

However, this research is insufficient in making an unified presentation that integrates similar evaluation and measurement results to an analytical model. In addition, because personal informatization level signifies one's level as a member of an enterprise, it needs to start from a skill-set analysis which reflects the characteristics of job positions and functions in the work processes. However, since it is realistically impossible to reflect in detail the work characteristics of 18,898 workers in a case application, such consideration was at least reflected in the way we focused on the general knowledge workers in the EUC environment, including some of the manufacturing staff.

Accordingly, future researchers should construct evaluation models that can supplement these problems, and present level evaluation models that can both reflect the characteristics of job positions and functions, and inclusively evaluate their levels. In addition, research is needed regarding level advancing roadmaps that can help with management after the evaluation results.

References

1. Mathis & Jackson: Human Resource Management. South-Western College Publishing (2000)
2. Torkzadeh & Lee: Measures of perceived end-user computing skills. Information & Management (2002)
3. Hannu Kivijarvi, Timo Saarinen: Investment in information systems and the financial performance of the firm. Information & Management (1995)
4. McCoy, Randall W.: Computer Competency for the 21st Century Information Systems Educator. Information Technology, Learning, and Performance Journal (2001)
5. Yoon: International Journal of Electronic Business. Development of an evaluation system of personal e-business competency and maturity levels (2004)
6. Floyd J. Fowler, Jr.: Improving Survey Questions –Design and Evaluation. Sage Publications (1995)
7. Floyd J. Fowler, Jr.: Survey Research Methods 3rd edition. Sage Publications (2001)

Critical Attributes of Organizational Culture Promoting Successful KM Implementation

Heejun Park

School of Engineering, Yonsei University, 134 Shinchon-Dong, Seodaemun-Gu, Seoul,120-749 Rep. of Korea
h.park@yonsei.ac.kr

Abstract. Many organizations are implementing Knowledge Management (KM) technologies to promote knowledge sharing. An extensive review of recent articles and journals about such implementations reveals that one of the main barriers to implementation of KM technology is the absence of an organizational culture that supports knowledge sharing. The purpose of this research is to explore the possible relationship between the successful implementation of knowledge management technology and specific organizational culture attributes. The OCP and the KMTP instruments were used to identify and rank the most critical organizational culture attributes of promoting knowledge sharing and KM technology implementation successes. Data were collected from twenty six US organizations involved in a KM effort.

1 The Role of Organizational Culture in KM Efforts

An extensive review of recent articles and journals about KM implementation reveals that one of the main barriers to implementation of KM technology is the absence of an organizational culture that promotes knowledge sharing [8]. The result from a recent survey conducted by the Knowledge Management Review demonstrates the main challenges KM practitioners faced when launching their KM initiative. The two main challenges are "Encouraging cultural adoption of KM" mentioned by 37.8% of the respondents and "Encouraging people to share" mentioned by 27.7%. "Managing information" was only mentioned by 8.1% of the respondents which indicates that Information Technology (IT) is far to be one of the main barriers to KM initiative success [7]. It seems that currently the IT tools designed to facilitate knowledge creation, capture, storage and distribution are available (even though no vendor currently offers an integrated enterprise wide KM solution) but the efficient use and acceptance of those tools are constrained by organizational culture [4]. After having primarily focused efforts on IT, practitioners are now realizing the importance of the "soft" aspect of KM initiatives [4].

There is a general agreement that organizational culture supporting knowledge sharing must be present or nurtured in order to succeed with a KM initiative. However, few academic researches have been conducted defining the key organizational culture attributes that support more effective utilization of KM technologies and knowledge sharing. The purpose of this research is to define these key cultural attrib-

O. Gervasi et al. (Eds.): ICCSA 2005, LNCS 3482, pp. 1316 – 1325, 2005.

utes. Once defined, one can measure them within organizational cultures and focus a cultural change initiative on these values. The purpose of this paper is not to describe how to implement this cultural change but one needs to be aware that the development of the social infrastructure which supports knowledge sharing is a much sticker and more contingent affair whereas the acquisition of technology is a relatively speedy process [2]. Chances of success are low if the culture is strong, if the purpose of the culture change is not well understood and not well accepted by employees. A strong communication campaign as well as the use of story telling can help to facilitate such change.

2 Organizational Culture Attributes and Their Measurement

There is almost an agreement among researchers concerning the definition and core components of organization culture. Unfortunately, this agreement is not so strong when we look at how to measure organizational culture. Qualitative and quantitative methods are complementary approaches to the study and assessment of organizational process and attributes. The advantages of qualitative methods include the use of focal unit's own terms to describe itself, the intensive and in-depth information that can be obtained about a unit, and the amenability of the method for exploratory research on

Table 1. 44 Attributes of the OCP

Trust	Problem Solving	Demanding of employee
Flexibility	Being exact	Supportive of employees
Adaptability	Team oriented work	Having a good reputation
Stability	Decisiveness	Sharing information freely
Predictability	Being competitive	Socially responsible
Being innovative	Being aggressive	Being different from others
Compliance	Being result oriented	Security of employment
Experimentation	Fairness	Praised good performance
Risk taking	Informality	Fitting in at work
Being careful	Tolerant of failure	Confront conflict directly
Freedom of action	Taking initiative	Develop friends at work
Rule oriented	Being thoughtful	Enthusiasm for the job
Attention to detail	Being easy going	Working closely with others
Take advantage of opportunity	Respect for individuals' right	Being calm
High expectation for performance	Low level of conflict encouraged	

issues and processes about which little information exists [3]. The advantages of quantitative methods include the ease of cross-sectional assessments and comparisons, the replicability of the assessment in different units and by other researchers or

organizational development professionals, and common articulated frame of reference for interpreting the data. Although both methods share the potential for producing cumulative bodies of information for assessment and theory testing, quantitative approach may be more practical for purposes of analyzing data-based change in organizations [3].

One of the quantitative methods, Organizational Culture Profile (OCP) developed by Harper [6] was used for this research to obtain a global perception of the culture of an organization. The OCP contains 44 attribute statements as shown in Table 1. that can generically capture individual and organizational attributes. The set of attribute statements was developed on the basis of an extensive review of academic and practitioner-oriented writings on organizational attributes and culture [6]. One aspect of this review was to identify a comprehensive set of attributes that could be used to characterize organizations. An attempt was made to find items that (1) could be used to describe any organization, (2) would not be equally characteristic of all organizations, and (3) would be easy to understand [6]. Respondents were asked to sort the 44 items into 8 categories, ranging from most to least desirable or from most to least characteristic, and to put a specified number of statements on each category. Items judged to be less characteristic or uncharacteristic were placed into middle categories [6]. While sorting the deck, the respondents were asked how to describe the culture of a focal organization. To develop a profile of an organization's culture, respondents familiar with the organization were instructed to sort the 44 attributes according to the extent to which the items were characteristic of the organization. With such a procedure, separate groups of individuals can be used to assess a firm's culture.

3 Research Questions and Hypotheses

Research in the field of Knowledge Management reveals that companies are adopting more knowledge management (KM) technologies to maximize the benefit of KM than ever, but they don't take full advantage of them. Is successful implementation of knowledge management not just a combination of new technology, but also organizational culture? If so, which cultural attributes do have positive or negative correlation with the successful KM technology implementation and knowledge sharing?

In determining the relationships between organizational culture attributes and a successful implementation of knowledge management technology, two basic hypotheses will be developed and tested.

Hypothesis I:
H0: There is no positive correlation between *the successful implementation of knowledge management technology* and cultural attributes (*trust, sharing information freely, working closely with others,* or *developing friends at work*) for that organization.
H1: There is a positive correlation between *the successful implementation of knowledge management technology* and cultural attributes (*trust, sharing information freely, working closely with others* or *developing friends at work*) for that organization.

Hypothesis II:
H0: There is no positive correlation between *the high return on knowledge management technology investment in sharing key talent and knowledge* and cultural attrib-

utes (*trust, sharing information freely, working closely with others,* or *developing friends at work*) for that organization.
H1: There is a positive correlation between the high return on knowledge management technology investment in sharing key talent and knowledge and cultural attributes (trust, sharing information freely, working closely with others, or developing friends at work) for that organization.

4 Research Methodology

For this research, the Knowledge Management Technology Profile (KMTP) was developed to assess the success of knowledge management technology implementation and effective knowledge sharing by modifying the Information Technology Investment Performance (ITIP) survey instrument developed by National Research Council [9]. The ITIP was developed to assess and understand patterns of behavior that could help explain why some organizations were, or were not, realizing greater payoffs from the investment in information technology [9]. Methods of determining success on knowledge management technology implementation were researched and it was decided to use a modification to the ITIP survey instrument.

To gauge the momentum of the KM movement, International Data Corp. and Knowledge Management Magazine undertook an extensive electronic survey of U.S. user organizations and individuals familiar with KM [5]. The results of the study demonstrating the most important reasons for adapting KM and the most common challenges to implementing KM were used in modification to ITIP. Once individual KMTP surveys were grouped by organization, overall KMTP score was determined by averaging the responses to each of the nine questions (from the question 1 to 9 in KMTP survey instrument) and summing the average of each question. This gives each organization a single KMTP score, indicating its success in implementation of KM technology.

As mentioned in section 2, the Organizational Culture Profile (OCP), the survey instrument developed by Harper, was used in investigating person-culture fit. The OCP uses the 8-category 44-item Q-sort scale with distribution {3-5-7-7-7-7-5-3}, which meets the general Q-sort distribution decisions based on symmetry of distribution, the number of judgment categories, and the essential shape of the symmetrical distribution.

For the purpose of this research, reliability is not a leading concern because of the changing nature of both organizational culture and the way knowledge management technology is utilized across an organization. This research presents only a snapshot of the organization under study and the employees' feeling and perceptions about organizational culture and the implementation of knowledge management technology. An organization is a dynamic entity; conditions surrounding the operation of the business are constantly changing and thus the results from a reliable test instrument would be expected to vary in reflection of those changing conditions [1].

The questionnaire Organizational Culture Profile (OCP) and the Information Technology Investment Performance (ITIP), slightly modified for this research, have been validated by many researchers in their previous researches.

5 Findings

5.1 Discussion of Data

The purpose of this research has been to determine the correlation, if any, between organizational culture attributes and the successful implementation of knowledge management technology. Data used to test the two hypotheses derived for this research were obtained from 227 respondents from the Organizational Culture Profile (OCP) survey instruments and 67 respondents from the Knowledge Management

Table 2. Summary of Participating Organization

Org.	Industry Type	Sample Size of KMTP	Sample Size of OCP
Org. A	Consulting	3	13
Org. B	Software Development	3	7
Org. C	Financial/Banking/ Accounting	3	9
Org. D	Consulting	3	10
Org. E	Manufacturing	3	8
Org. F	Financial/Banking/ Accounting	4	10
Org. G	IT/ Telecommunication	2	7
Org. H	Government	3	13
Org. I	Consulting	3	9
Org. J	IT/ Telecommunication	3	11
Org. K	Software Development	4	8
Org. L	Consulting	3	22
Org. M	Software Development	3	7
Org. N	Government	3	7
Org. O	IT/ Telecommunication	3	11
Org. P	Consulting	3	11
Org. Q	Software Development	2	9
Org. R	Education	2	4
Org. S	Financial/Banking/ Accounting	3	18
Org. T	Consulting	3	7
Org. U	Software Development	3	5
Org. V	IT/ Telecommunication	1	3
Org. W	Consulting	1	5
Org. X	Consulting	1	4
Org. Y	Financial/Banking/ Accounting	1	5
Org. Z	Financial/Banking/ Accounting	1	4

Technology Profile (KMTP) survey instruments representing 26 separate organizations. A total of 1060 OCP survey instruments and 212 KMTP survey instruments were distributed across 44 organizations. The OCP survey instruments were distributed to employees within the organization regardless of employees' function and level. The KMTP survey instruments were distributed to managers who were in a position to be knowledgeable about knowledge management technology across the organizations. A total of 236 OCP survey instruments were completed and returned from 27 organizations with the response rate of 22.3 percent. A total of 67 KMTP survey instruments were completed and returned from 26 organizations with the response rate of 31.6 percent. One organization that returned only the OCP survey instruments was excluded out of sample organizations. Table 2. provides detailed information as to the number of respondents to the OCP and the KMTP survey instruments from each of 26 participating organizations and the industry types of these organizations. The alphabet (A to Z) was assigned to each of 26 organizations randomly to protect confidentiality of participating organizations.

Table 3. Correlation between OCP Cultural Attributes and KMTP Score

OCP Attributes	Correlation	OCP Attributes	Correlation
Sharing information freely	0.83	Security of employment	0.01
Working closely with others	0.78	Low level of conflict encouraged	-0.05
Team oriented work	0.69	Being careful	-0.08
Trust	0.69	Socially responsible	-0.09
Fairness	0.63	Stability	-0.09
Enthusiasm for the job	0.63	Confront conflict directly	-0.09
Autonomy	0.47	Fitting in at work	-0.13
Flexibility	0.45	Respect for individual's right	-0.14
Supportive of employees	0.44	Being different from others	-0.14
Tolerance of failure	0.44	High expectations for perform-	-0.17
Rule orientation	0.41	Informality	-0.18
Praised good performance	0.37	Being innovative	-0.18
Experimentation	0.33	Being result oriented	-0.24
Demanding of employees	0.32	Predictability	-0.33
Take advantage of opportunity	0.31	Taking initiative	-0.34
Having a good reputation	0.31	Being easy going	-0.42
Being exact	0.29	Compliance	-0.46
Decisiveness	0.28	Risk taking	-0.49
Problem solving	0.17	Attention to detail	-0.53
Adaptability	0.16	Being competitive	-0.63
Developing friends at work	0.09	Being aggressive	-0.68
Being thoughtful	0.04	Being calm	-0.79

5.2 Data Analysis

The Pearson Product-Moment correlation coefficient was considered as a method of determining linear relationship between two quantitative variables measured in interval scales – organizational culture and the successful implementation of knowledge management technology. However, nonparametric alternative to Pearson Product-Moment correlation, Spearman's correlation coefficient, was used with replacing the data values for each variable by ranks because the variables are not normally distributed. The fact that variables are not normally distributed is due to the sample size.

Relationship between KMTP Score and OCP Cultural Attributes. The correlations between 44 OCP cultural attributes and the KMTP success score were examined using Spearman's correlation coefficients (see Table 3.). This non-parametric

Table 4. Correlation between KMTP Score for Knowledge Sharing and OCP Cultural Attributes

OCP Attributes	Correlation	OCP Attributes	Correlation
Team oriented work	0.72	Respect for individual's right	0.15
Working closely with others	0.68	Enthusiasm for the job	0.10
Sharing information freely	0.62	Tolerant of failure	0.06
Trust	0.61	Security of employment	0.01
Supportive of employees	0.58	Demanding of employee	-0.09
Take advantage of opportunity	0.52	Being different from others	-0.14
Flexibility	0.46	Low level of conflict encour-	-0.15
Confront conflict directly	0.44	Risk taking	-0.24
Autonomy	0.44	Being careful	-0.28
Having a good reputation	0.41	Taking initiative	-0.34
Fairness	0.38	Being aggressive	-0.34
Being innovative	0.37	Rule oriented	-0.38
Developing friends at work	0.35	Being exact	-0.38
Adaptability	0.31	Being easy going	-0.39
Experimentation	0.31	Being result oriented	-0.41
Fitting in at work	0.29	Being competitive	-0.41
Praised good performance	0.27	Predictability	-0.43
Being thoughtful	0.24	Decisiveness	-0.44
Problem Solving	0.20	Attention to detail	-0.48
Socially responsible	0.19	Stability	-0.61
Informality	0.17	Compliance	-0.61
High expectation for performance	0.15	Being calm	-0.75

correlation analysis reveals a number of cultural attributes having moderate to high positive correlation with the KMTP success score. These attributes include *sharing information freely*, *working closely with others, team oriented work, trust, fairness,* and *enthusiasm for the job*. A number of cultural attributes having a moderate to high negative correlation with the KMTP score were also identified. These attributes include *being calm, being aggressive,* and *being competitive.*

Relationship between Successful Knowledge Sharing and OCP Cultural Attributes. From the KMTP survey instruments grouped by organization, a portion of the assessment included the respondents' judgment on KM technology investment in sharing key talent and knowledge within the respective organizations. The average KMTP score for knowledge sharing for each of participating organizations was tallied and normalized. A non-parametric correlation analysis was used to determine the correlation between the score of successful implementation of KM technology on knowledge sharing and the 44 OCP attributes. Table 4. summarizes the correlation between these two variables produced by the Spearman's Correlation analysis.

Analysis of Hypotheses. The Spearman's Rank Correlation coefficient (Rho) was used to determine the relationship between two quantitative variables measured in interval scale with replacing the data values for each variable by ranks because the variables are not normally distributed. The Pearson Product-Moment correlation could be used with the sample size larger than 30 if the variables are approximately normally distributed. However, the sample size of this research (n=26) is not sufficiently large to use the Pearson Product-Moment correlation coefficient. The hypotheses were tested based on the findings from the correlation analysis with 99% confidence interval.

Table 5. Testing Hypothesis I

OCP Attributes	Correlation	*t* - value
Trust	0.69	$t : 3.45 > t_{.005,\ 25} : 2.79$
Sharing information freely	0.83	$t : 4.15 > t_{.005,\ 25} : 2.79$
Working closely with others	0.78	$t : 3.90 > t_{.005,\ 25} : 2.79$
Developing friends at work	0.09	$t : 0.45 < t_{.005,\ 25} : 2.79$

The research hypothesis I postulates that organizations indicating a higher overall success for knowledge management technology implementations, would find that employees rank attributes such as *trust, sharing information freely, working closely with others,* or *developing friends at work* more positively in their assessment of organizational culture attributes than employees within companies whose knowledge management technology implementations indicate a lower overall success. The *t* – values calculated against Spearman's Correlation coefficients of attributes *trust* ($r = 0.69$), *sharing information freely* ($r = 0.83$), and *working closely with others* ($r = 0.78$) found from the data analysis are sufficient to reject null hypothesis (see Table 5.).

Table 6. Testing Hypothesis II

OCP Attributes	Correlation	*t* - value
Trust	0.61	t : 3.05 > $t_{.005, 25}$: 2.79
Sharing information freely	0.62	t : 3.10 > $t_{.005, 25}$: 2.79
Working closely with others	0.68	t : 3.15 > $t_{.005, 25}$: 2.79
Developing friends at work	0.35	t : 1.75 < $t_{.005, 25}$: 2.79

The hypothesis II postulates organization indicating a high return on knowledge management technology investment in sharing key talent and knowledge, would find that employees rank attributes such as *trust, sharing information freely, working closely with others,* or *developing friends at work* more positively in their assessment of organizational culture attributes than employees within companies whose investment in knowledge management technology indicate less return on sharing key talent and knowledge. The *t* – values calculated against Spearman's Correlation coefficients of attributes *trust* (r = 0.61), *sharing information freely* (r = 0.62), and *working closely with others* (r = 0.68) found from the data analysis are sufficient to reject null hypothesis (see Table 6.).

6 Conclusions and Recommendations

The results of the data analysis revealed sufficient evidence to establish a correlation between cultural attributes and the successful implementation of knowledge management technology and knowledge sharing. Before an organization launches a knowledge management technology initiative, it should deal with cultural issues. The success of KM technology implementation is mediated by human behavior. The research identifies cultural attributes, which have moderate to high positive correlation with the success of KM technology implementation such *as sharing information freely, working closely with others, team-oriented work, trust, fairness,* and *enthusiasm.*

Many organizations are actually implementing KM strategies and technologies that are giving them real benefits in terms of knowledge sharing. The results of this research which support the supplemental hypothesis II indicate that a high return on knowledge management technology investment in sharing knowledge has moderate high correlation with cultural attributes such as *team-oriented work, working closely with other, sharing information freely, trust,* and *supportive of employees.*

The findings of this research could help KM researchers and practitioners to develop a better understanding of the role of organizational culture in the successful implementation of KM technology and knowledge sharing initiatives. The findings provide some key cultural attributes that practitioners will be able to focus on and to pay particularly attention to during cultural change initiatives.

References

1. Anastasi, A.: The Concept of Validity in the Interpretation of Test Scores, *Educational and Psychological Measurement,* Spring (1950) 67-78
2. Bruss, Lois R.: Ten steps to achieve KM success, *KMWorld*, Vol. 8, Iss. 4 (1999) 13-16
3. Cooke, R and Rousseau, D.: Behavioral Norms and Expectations: A Quantitative Approach to the Assessment of Organizational Culture, *Group and Organizational Studies*, Vol. 13 (1988) 245-273
4. Davenport, Thomas, and Prusak, Laurence: Working Knowledge. How organizations manage what they know, Harvard Business School Press, MA (1998)
5. Dyer, Greg: KM Crosses the Chasm: IDC State of the Market Survey, *Knowledge Management*, March (2000) 50-54
6. Harper, George H.: Assessing Information Technology Success as a Function of Organizational Culture, Ph.D. diss., Dept. of Industrial and Systems Engineering and Engineering Management, University of Alabama (2000)
7. KM Review survey reveals the challenges faced by KM practitioners, Knowledge Management Review, Vol. 4, Iss. 5 (2001)
8. KPMG Consulting: *Knowledge Management Research Report*, KPMG (2000)
9. National research Council, Computer Science and Telecommunications Board: *Information Technology in the Service Society*, National Academy Press, Washington, DC (1994)

Author Index