

Studies in Computational Intelligence 595

Grzegorz Borowik
Zenon Chaczko
Witold Jacak
Tadeusz Łuba *Editors*

Computational Intelligence and Efficiency in Engineering Systems

 Springer

Studies in Computational Intelligence

Volume 595

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series “Studies in Computational Intelligence” (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

More information about this series at <http://www.springer.com/series/7092>

Grzegorz Borowik · Zenon Chaczko
Witold Jacak · Tadeusz Łuba
Editors

Computational Intelligence and Efficiency in Engineering Systems

 Springer

Editors

Grzegorz Borowik
Warsaw University of Technology
Warsaw
Poland

Zenon Chaczko
University of Technology
Sydney, NSW
Australia

Witold Jacak
University of Applied Sciences Upper
Austria
Hagenberg
Austria

Tadeusz Łuba
Warsaw University of Technology
Warsaw
Poland

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-319-15719-1 ISBN 978-3-319-15720-7 (eBook)
DOI 10.1007/978-3-319-15720-7

Library of Congress Control Number: 2015932240

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Foreword

As the quantity of data increases and data processing hardware and related processing algorithms become increasingly efficient, also the scope of the field of Computational Intelligence is broadening. This parallels rapid extensions of heuristic and meta-heuristic computer science and engineering methods that are inspired by nature. These methods can tackle such demanding areas of inquiry as data mining, knowledge discovery, semantic databases, big data and data analytics... to name a few.

Along with the development of tools for producing, processing and storage of data, the technology makes it possible to process increasingly more complex information in many diverse forms. The forms include multimedia, streamed videos including spatial and acoustic channels, massive spatial data and geospatial maps, time flow diagrams, social network connections and communication patterns, including their frequency spectra and time series consisting of social and commercial interactions.

The emerging time/frequency data consisting of events, sequences and their analog or discrete descriptors typically base on classic concepts and transforms of J.B. Fourier, G. Boole, C. Shannon and others. However, augmented with computational intelligence techniques, they allow for increasingly deep insights into data. To meet the various challenges of data mining and knowledge discovery, researchers continue to investigate and use complex models and computational algorithms including Boolean reasoning and fast Boolean computation, data ontologies, description logic, grammatical optimization, fuzzy logic, data fusion, evolutionary approaches and other heuristic methods.

Applications and theoretical extensions presented in this book are inspired by both these classic and modern techniques. They include evolutionary computation for environmental monitoring, digital patterns for heritage and data preservation, bioinformatics, intelligent warehousing and logistics, business analytics, cognitive radio and software-defined networks, multimedia and image processing, smart grids, cloud computing and internet of things, wireless sensor networks and environmental monitoring.

The editors have carefully selected the best papers presented at the APCASE 2014 conference to provide an insightful and critical review of issues related to intelligent computing, covering the state of the art of various techniques and real-world applications. For easier reference, the extended versions of papers originally presented at the APCASE 2014 were grouped into four main parts related to: computational models and knowledge discovery; communications networks and cloud computing; computer-based systems; and data-oriented and software-intensive systems.

This carefully edited and reviewed volume addresses the increasingly popular demand for seeking more clarity in the data that we are immersed in. It should be of interest to scientists, Internet companies' researchers, information technology professionals, or even ordinary users who are looking for nuggets of information across the vast universe of today's Internet.

I sincerely hope that this book will be a valuable reading for both researchers and students in the area of computer science, telecommunications and electronics.

Louisville, October 2014

Dr. Jacek M. Żurada
Professor of Electrical and Computer Engineering
University of Louisville, Kentucky, USA
Foreign Member of Polish Academy of Sciences
Past President of IEEE Computational Intelligence Society
IEEE Life Fellow

Preface

The content of the book includes high quality contributions to APCASE 2014, the second Asia-Pacific Conference on the Computer Aided System Engineering, February 10–12, 2014 in South Kuta, Bali, Indonesia. APCASE 2014 was organized by the University of Technology, Sydney (Australia), Binus University of Technology (Indonesia), Warsaw University of Technology (Poland) and the University of Applied Sciences in Hagenberg (Austria). The aim of APCASE series of conferences is to provide a highly prestigious venue for academics, system engineering and applied science researchers as well as practitioners in the Asia-Pacific region.

The proposed volume includes an excellent choice of extended versions of papers presented at the APCASE 2014 conference. The book is made of four main parts that cover data-oriented engineering science research in a wide range of applications:

- Computational Models and Knowledge Discovery
- Communications, Networks, and Cloud Computing
- Computer-Based Systems
- Data-Oriented and Software-Intensive Systems

The first part presents recent advances in computational models and knowledge discovery. This part also covers heuristic computational models and model-driven system design. These approaches are highly applicable for problems which cannot be easily solved by deterministic methods, due to the dimensionality, complexity or specificity issues. The issues discussed in this part of the book include: Boolean reasoning and fast Boolean computation, data ontologies, description logic, accelerated and simulation-based optimization techniques, implicit solution spaces and dynamic solution space reduction, neural network methods, iterative back projection methods, identification and classification, location of objects and events, symbolic regression and search strategies for grammatical optimization. The second part discusses most recent advances in communication and networks. In this part, innovative computational models of data transmission, as well as, practices and

implementation issues related to sensor network software infrastructure; reliability in multistage interconnection networks; authentication, authorization and the task workflow in the cloud infrastructure. The third part presents most recent developments, challenges and practical applications related to computer-aided system optimization and design principles. A range of application fields and case studies is presented, including: tracking moving targets and image processing; control, data fusion and monitoring of biosignals; electronic manufacturing, multi-core processors and smart grid technologies. The fourth and the final part presents an overview of the most recent advances in the domain of data-oriented and software-intensive systems. Such areas as: intelligent warehousing, simulation of supply-chain logistic, steganographic image processing, augmented reality solutions, e-commerce applications, financial markets and Internet banking systems are covered.

The organization of the book allows highlighting the tremendous and steadily growing role of data-oriented and software intensive systems in almost all domains of human activity. Additionally, such organization enables authors to provide a comprehensive review of new challenges these systems are facing when dealing with increasingly more complex solutions for solving data explorations.

This book offers excellent examples of the intelligent ubiquitous computation, as well as recent advances in systems engineering and informatics. The content represents state-of-the-art foundations for researchers in the domain of modern computation, computer science, system engineering and networking, with many examples that are set in the industrial application context.

The readers will greatly benefit from acquiring knowledge of the advanced methods and applications in computational intelligence on what and how various engineering problems and challenges can be approached and resolved in several domains. They will learn various methods and techniques that could be applied in order to solve these problems. One of the most important benefits that potential readers will gain is not only good understanding of what the major challenges are but also the most practical and efficient solutions to address them.

There are very few books that currently could offer similar content, depth and practicality of the methodologies and techniques contained in the book. Additionally, the publication offers a very convenient entry for researchers and engineers who intend to work in the discussed research domains.

The book is mainly targeted at scientists, engineers and IT specialists in the fields of computational intelligence, telecommunication, control engineering, artificial intelligence, signal processing, software engineering, electrical engineering, mechanics, robotics, soft computing and many others who have interests and needs to understand the theory and practical applications of computational intelligence and heuristic methodologies. Many students at various educational levels (undergraduate, graduate and postgraduate) could benefit from this book, as it offers a

consistent material as well as a well-researched bibliography. From this perspective, the book can be used as a textbook in any course in the field of computational intelligence, computer system engineering, telecommunications, software engineering and IT.

Warsaw, October 2014
Sydney
Hagenberg

Grzegorz Borowik
Zenon Chaczko
Witold Jacak
Tadeusz Łuba

Contents

Part I Computational Models and Knowledge Discovery

1	Evolutionary Feature Optimization and Classification for Monitoring Floating Objects	3
	Anup Kale and Zenon Chaczko	
2	Alternative Approaches for Fast Boolean Calculations Using the GPU	17
	Bernd Steinbach and Matthias Werner	
3	Technique for Transformation of DL Knowledge Base to Boolean Representation	33
	Grzegorz Borowik and Dariusz Nogalski	
4	Digital Patterns for Heritage and Data Preservation Standards	47
	Lucia Carrion Gordon and Zenon Chaczko	
5	Methods for Genealogy and Building Block Analysis in Genetic Programming	61
	Bogdan Burlacu, Michael Affenzeller, Stephan Winkler, Michael Kommenda and Gabriel Kronberger	
6	Multi-Population Genetic Programming with Data Migration for Symbolic Regression	75
	Michael Kommenda, Michael Affenzeller, Gabriel Kronberger, Bogdan Burlacu and Stephan Winkler	

7 Search Strategies for Grammatical Optimization Problems—Alternatives to Grammar-Guided Genetic Programming 89
Gabriel Kronberger and Michael Kommenda

8 Identification and Classification of Objects and Motions in Microscopy Images of Biological Samples Using Heuristic Algorithms. 103
Stephan M. Winkler, Susanne Schaller, Daniela Borgmann, Lisa Obritzberger, Viktoria Dorfer, Christian Haider, Sandra Mayr, Peter Lanzerstorfer, Claudia Loimayr, Simone Hennerbichler-Lugscheider, Andrea Lindenmair, Heinz Redl, Michael Affenzeller, Julian Weghuber and Jaroslav Jacak

Part II Communications, Networks, and Cloud Computing

9 Comparison of TCP/IP Routing Versus OpenFlow Table and Implementation of Intelligent Computational Model to Provide Autonomous Behavior 121
Ameen Banjar, Pakawat Papatwibul and Robin Braun

10 Designing Biomimetic-Inspired Middleware for Anticipative Sensor-Actor Networks 143
Christopher Chiu and Zenon Chaczko

11 Empirical Analysis of Terminal Reliability in Multistage Interconnection Networks 157
Nur Arzilawati Md Yunus and Mohamed Othman

12 Authorial, Adaptive Method of Users’ Authentication and Authorization. 171
Robert Sekulski and Marek Woda

13 Estimation of Quality of Service in Stochastic Workflow Schedules 185
Michal Wosko and Jan Nikodem

14 Multi-Party System Authentication for Cloud Infrastructure by Implementing QKD 195
Roszelinda Khalid, Zuriati Ahmad Zukarnain, Zurina Mohd Hanapi and Mohamad Afendee Mohamed

15 SmartCloud Orchestrator—the First Implementation for Education in the World at WrUT 209
 Jerzy Greblicki, Jerzy Kotowski, Mariusz Ochla and Jacek Oko

16 Cloud Computing—Effect of Evolutionary Algorithm on Load Balancing 217
 Shahrzad Aslanzadeh, Zenon Chaczko and Christopher Chiu

Part III Computer-Based Systems

17 Crossed Linear Arrays Using Doppler Radar Beamforming for Detecting Single Moving Targets 229
 Jiajia Shi and Robin Braun

18 Image Construction Using Low Cost Airborne Beamforming 243
 Jiajia Shi and Robin Braun

19 Control of Hand Prosthesis Using Fusion of Biosignals and Information from Prosthesis Sensors 259
 Andrzej Wolczowski and Marek Kurzynski

20 Transcutaneous Bladder Spectroscopy: Computer Aided Near Infrared Monitoring of Physiologic Function 275
 Andrew Macnab, Lynn Stothers, Babak Shadgan and Behnam Molavi

21 Designing and Manufacturing Quartz Crystal Oscillators 293
 Benfano Soewito

22 Evaluation of Cache Coherence Mechanisms for Multicore Processors 307
 Malik Al-Manasia and Zenon Chaczko

23 Aggregator Operation for EVs in Korean Smart Grids Testbed 321
 Wang-Cheol Song and Zhong Ming Huang

Part IV Data-Oriented and Software-Intensive Systems

24 Bio-informatics with Genetic Steganography Technique 333
 Raniyah Wazirali, Zenon Chaczko and Lucia Carrion

25 Augmented Reality and the Adapted of Smart Grid Monitoring for Educational Enhancement 347
Zenon Chaczko, Wael Alenazy and Amy Tran

26 Analysis of E-Commerce User Behavior of Indonesian Students: A Preliminary Study of Adaptive E-Commerce 365
Rianto, Lukito Edi Nugroho and P. Insap Santosa

27 Optimizing Financial Markets in C# .NET 377
Billy Leung, Zenon Chaczko and Jan Nikodem

28 The Model of Customer Trust for Internet Banking Adoption 399
Shidrokh Goudarzi, Wan H. Hassan, Mir Ali Rezazadeh Bae and S.A. Soleymani

29 Robust Storage Assignment in Warehouses with Correlated Demand 415
Monika Kofler, Andreas Beham, Stefan Wagner and Michael Affenzeller

30 Concise Supply-Chain Simulation Optimization for Large Scale Logistic Networks 429
Erik Pitzer and Gabriel Kronberger

Part I
Computational Models and Knowledge
Discovery

Chapter 1

Evolutionary Feature Optimization and Classification for Monitoring Floating Objects

Anup Kale and Zenon Chaczko

Abstract Water surfaces are polluted due to various man-made and natural pollutants. In urban areas, natural water sources including rivers, lakes and creeks are the biggest collectors of such contaminants. Monitoring of water sources can help to investigate many of details relating to the types of litter and their origin. Usually two principle methods are applied for this type of applications, which include either a use of in-situ sensors or monitoring by computer vision methods. Sensory approach can detect detailed properties of a water including salinity and chemical composition. Whereas, a camera based detection helps to monitor visible substances like floating or immersed objects in a transparent water. Current computer vision systems require an application specific computational models to address a variability introduced due to the environmental fluctuations. Hence, a computer vision algorithm is proposed to detect and classify floating objects in various environmental irregularities. This method uses an evolutionary algorithmic principles to learn inconsistencies in the patterns by using a historical data of river pollution. A proof of the concept is built and validated using a real life data of pollutants. The experimental results clearly indicate the advantages of proposed scheme over the other benchmark methods used for addressing the similar problem.

1.1 Introduction

Monitoring of water surfaces is becoming a daily need due to the amount of waste accumulated in urban rivers, lakes and creeks [1]. Urban water streams are often used in various ways by the people living around, this includes transport, domestic use and recreation. Due to frequent usage and presence of human population living around there is always a high probability to receive the man-made pollution. This pollution

A. Kale (✉) · Z. Chaczko
Faculty of Engineering and IT, University of Technology Sydney, 15 Broadway,
Ultimo, NSW, 2007, Australia
e-mail: Anup.V.Kale@student.uts.edu.au

Z. Chaczko
e-mail: Zenon.Chaczko@uts.edu.au

© Springer International Publishing Switzerland 2015
G. Borowik et al. (eds.), *Computational Intelligence and Efficiency
in Engineering Systems*, Studies in Computational Intelligence 595,
DOI 10.1007/978-3-319-15720-7_1

may include plastic containers, plates, polythene bags, papers and other materials which can float and are not biodegradable. Hence, existence of such foreign materials can create long term implications to the rivers. One of the most affected species due to such pollution is marine life. Plant and animals living in rivers has their own eco-cycle and may have interdependence on each other. If a pollution creates a certain effect of a particular marine specie then it quite possible that other as well may get affected. Thus, care needs to be taken to avoid such possibilities in order to maintain the balance in the lifecycle existing in a particular river. Another important aspect under consideration is a human-health and general environmental concerns. River banks can often attract local community and tourists due to its attractiveness and usefulness for community. In many case community residing near river is regularly in contact with the river and surroundings due to jogging/walking paths, cycling tracks, fishing, boating and in some cases other special recreation. If polluted river surfaces are not maintained then over the period of time it is possible to either have an effect on the people using this environment or reduction in the usage itself. In modern world many developed countries use rivers in the urban areas as a showcase as their environment friendliness and even an occasional presence of floating pollution can create a serious damage to the reputation of a particular river or city. Thus, even from human life perspective it is very important to monitor those water surfaces which are continuously in touch with high volume of human population.

Attempts to monitor rivers using in situ sensors and computer vision are made in the past. Wireless sensors can monitor detailed features [2] and parameters of the water and a computer vision [3] can provide a higher level perception of the floating objects. Various computer vision studies related with monitoring water surfaces include oil-slick monitoring [4] and marine life monitoring [5]. Commonly used algorithms may include segmentation based techniques, probabilistic approaches and heuristic methods. So far very limited investigation is carried out to address the problem specifically relating to natural water in urban areas. Whereas, there are many techniques which needs detailed investigation to find their merit for this particular problem. A feature based algorithmic approach is very popular technique in computer vision applications. Dimensionality reduction of imagery by using a feature representation reduces performance concerns and enhances accuracy. But this approach is still under investigation and many studies are conducted to optimize or select features from an original feature set extracted from the image. It is believed that first version of feature-set extracted from any image may include many redundant features. So, investigation is conducted to reduce these redundant features which may be common to all feature-sets in a particular multiclass image dataset. Genetic Algorithms are very useful for such a purpose and is a common tool for feature subset optimization. Another important merit of GA for a particular problem under consideration here is ability to deal with highly variable datasets. Two very common implementations include filter and wrapper methods, where filter uses unclassified data and the other used classified data. Use of the prelabelled or previously classified historical data is usually categorized under a supervised machine learning approach whereas use of unlabelled raw data provides an unsupervised approach. For both filter and wrapper techniques the fitness function is usually a classification accuracy or number of

features in the optimized feature set. Many variants of GA based methods are applied to several different types of problems and are discussed in details in the next section of this paper.

For adapting the GA based method specifically to the computer vision application it needs various alterations and specific modeling of the problem. Here, we propose a method for floating pollution detection on water surfaces using computer vision. The proposed technique uses a supervised learning algorithm for object detection and classification. Various computational stages in the algorithm include feature extraction, optimization and classification. A Matlab based prototype was built to prove the concept and real-life imagery was acquired to validate the prototype. The results clearly indicate merits our approach and its advantages over past research studies in-line with this body of knowledge. Following sections discuss background context of computer vision and feature based detection approach, proposed method and most recent results respectively.

1.2 Object Detection Techniques Using Feature Selection

Feature based detection techniques are very important from performance and accuracy point of view. In general a feature is a representative of bigger size of data/information and some of such features can carry meaningful information from a bigger perspective in a particular problem. In this section we will discuss image features and their types, feature selection methods using genetic algorithm and its various implementations.

1.2.1 Feature Selection an Overview

Feature Selection is a technique to select meaningful collection of features and omit redundant ones from a larger pool. It is important computing operation in many applications to reduce the noisy data and make the overall collection more meaningful. Typically, in this type of problem the historical feature set with large volume of features is searched for a consistency and those features which are relevant in all feature instances are selected as an optimized feature subset. The relevance of these features can be assessed against one or more out of the objective, sample datasets, probabilities of instance spaces and class labels, complexity measurement, incremental usefulness and entropic relevance. Feature Selection Algorithms or FSAs can be characterized on the basis of either on search organization or generation of successors or evaluation measures. The search organization characteristic is a general method to for exploring the hypothesis. Whereas, the generation of successors is a system to propose the expected irregularities and the evaluation measures are the functions to measure the successor candidates with some ideal hypothesis. Feature Selection process may have three different schemes which are:

- a. Embedded Scheme: Inducer having own feature selection algorithm
- b. Filter Scheme: Here the feature selection takes place in an unsupervised manner and before the induction step.
- c. Wrapper: In this case the features are periodically classified to assess their relevance against the objective.

Depending on the application every scheme has its own merits and drawbacks. As an example the filter scheme can be very good approach when the desired classification results are not known and has its effect on the accuracy of the results. In case of wrapper selection approach it gives very good accuracy since it uses the historical and known objectives.

1.2.2 Genetic Algorithm for Feature Selection

Genetic Algorithm is a search and optimization method inspired from Darwinian principle of evolution i.e. Survival of the fittest. In this algorithm multiple solution candidates are generated randomly and assessed using a certain objective or a fitness function. Every element in the population is encoded in a certain way which can be string of bits or numbers. If this fitness value is at the desired level or any other termination criterion is met the algorithm is terminated. Else, the process continues with the regeneration operations called crossover and mutation. In case of the crossover operation the one or more bit/s is/are swapped within the pairs of original population generated and a new population of siblings is created. Whereas, mutation operation randomly changes bits of a very smaller portion of the population and maintains the diversity in it. This cycle of fitness check and regeneration is continued till the termination criterion is met. The termination criterion is a condition to terminate the algorithm which can be either fitness level or time consumed. This routine of operation in the GA provides many advantages including the automated, randomized and parallel approach to the search of information a very large pool of data. But nonetheless, it also needs modifications for making it usable and efficient for a particular application. Making the GA application specific may need a problem specific modelling of the chromosomes, hybridization with the classical methods and use of special or additional operators. Ferri et al. discuss need of hybridization and population diversity to make feature selection more efficient [6]. Yang and Honvar [7] discuss multi-criteria optimization formulation to the feature selection problem. Similarly, Oliveria et al. [8] apply modified wrapper based multi-criteria algorithm for reducing the feature set. Penkopf and O'Leary [9] provide an example of encoding in GA which can improve the process of feature selection. In case of Sun et al. [10] their method uses Hybrid GA implementation for image processing applications. A study conducted by Oh et al. gives an example of hybridization of GA where the local search method fine tunes the search by GA [11]. This work stresses the need of embedding a domain or an application specific knowledge into the formulation of GA based feature selection scheme. Lin et al. implemented GA with silhouette

classification of multiple tumor classes [12]. Zhu and Guan [13] suggest modular GA for feature classification problem. Here, less important features are detected for reducing the redundant information and improving the accuracy of the machine learning process. As per authors, this method also reduce the computational cost significantly. Lac and Stacey suggest Multi-objective GA and Pareto optimality as a solution for accuracy improvement and feature set reduction [14]. Zhuo et al. applied GA to parallelize the search operation which involves searching of spectral bands, features and SVM kernels [15]. Other practical examples include food and industrial application with very similar approach in previously discussed cases here.

All implementations of GA discussed here draw some parallels which are:

- a. Most of the implementations use either filter (unsupervised) or a wrapper (supervised) scheme.
- b. The classification is usually done using SVM, 1-NN or Neural networks.
- c. Most of the applications consider reducing feature subset and improving accuracy of the classifier as a part of their objective function.

Even so, all above examples do not address the problem we are trying to solve, i.e. the problem having a moving optima or the datasets acquired from a variable environmental conditions. Thus, as discussed before it needs special consideration and a detailed look at GA to investigate the possible modifications. Next section of the paper discusses the modified version of GA which assumes embedding additional operation and a different approach towards encoding to achieve the adaptively.

1.3 Supervised Feature Classification

Based on the background study and problem specific requirements we propose a detection algorithm. It is based on the supervised machine learning principles and uses a GA for feature optimization. The process for detection has following sequence of computational operations:

1.3.1 Genetic Algorithm for Feature Selection

1.3.1.1 Training Mode

1. Extract feature matrix which includes shape descriptors, texture features and interest points from a labeled historical image data
2. Apply genetic algorithm for optimization of the feature space
3. Store final pool of population after GA optimization cycle
4. Find repetitively present (call static) and rarely present (call dynamic) features
5. Test various percentage combinations of static and dynamic feature members. I.e. 90–10, 80–20 etc. for determining the object classification accuracy
6. Store best fit model as a setting data for the working mode of the algorithm

1.3.1.2 Work Mode

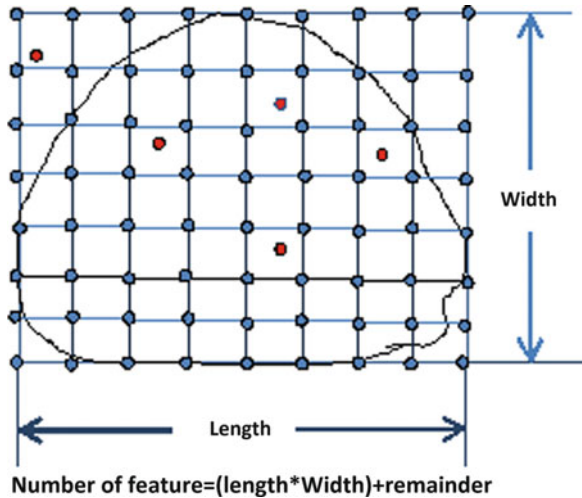
1. Extract features from live video stream
2. Select features as per training mode result
3. Apply classification to the selected feature set

Following part of this section we discuss some of the important computational processes of the algorithm in detail.

1.3.2 Feature Extraction

A feature in the context of computer vision application can be defined as an entity which carries valuable information. One or more features can help to define objects present in the image. Most commonly used features can be either in spatial domain (i.e. points, edges, shapes, etc.) or in a spectral domain (i.e. spectral descriptors for colour, shape, images, etc.). From real time performance perspective and for reducing the computational cost we decided to process interest points. In our proposed feature extraction method, we use colour pixels sampled proportional to the length and width of the region interest acquired using the blob detection algorithm. We also add shape features [16] in the spatial domain which include area, centroid, fullness, major and minor axis lengths, etc. The overall feature extraction process is described below (Figs. 1.1, 1.2 and 1.3):

Fig. 1.1 Interest point sampling



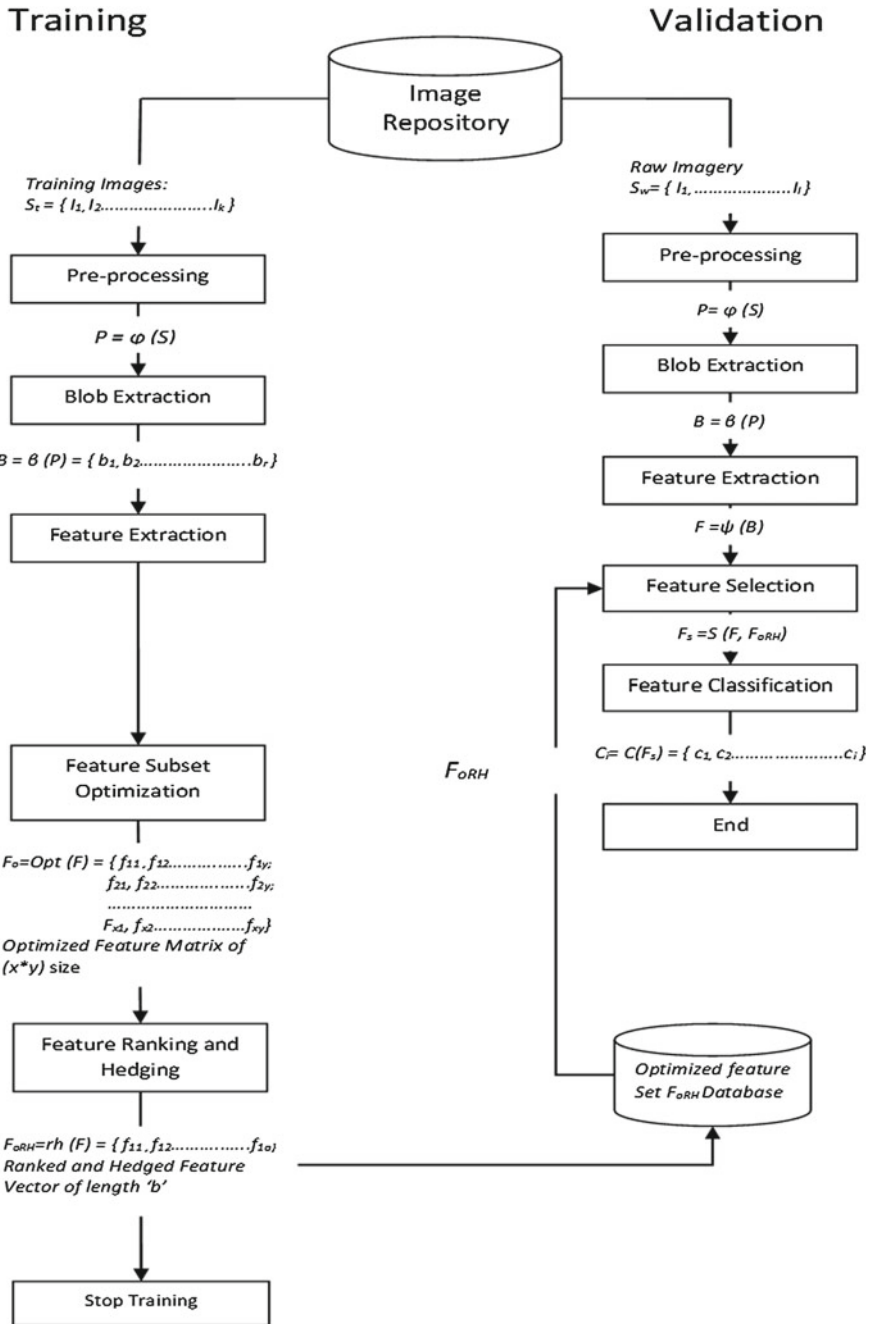


Fig. 1.2 Flow diagram of detection process

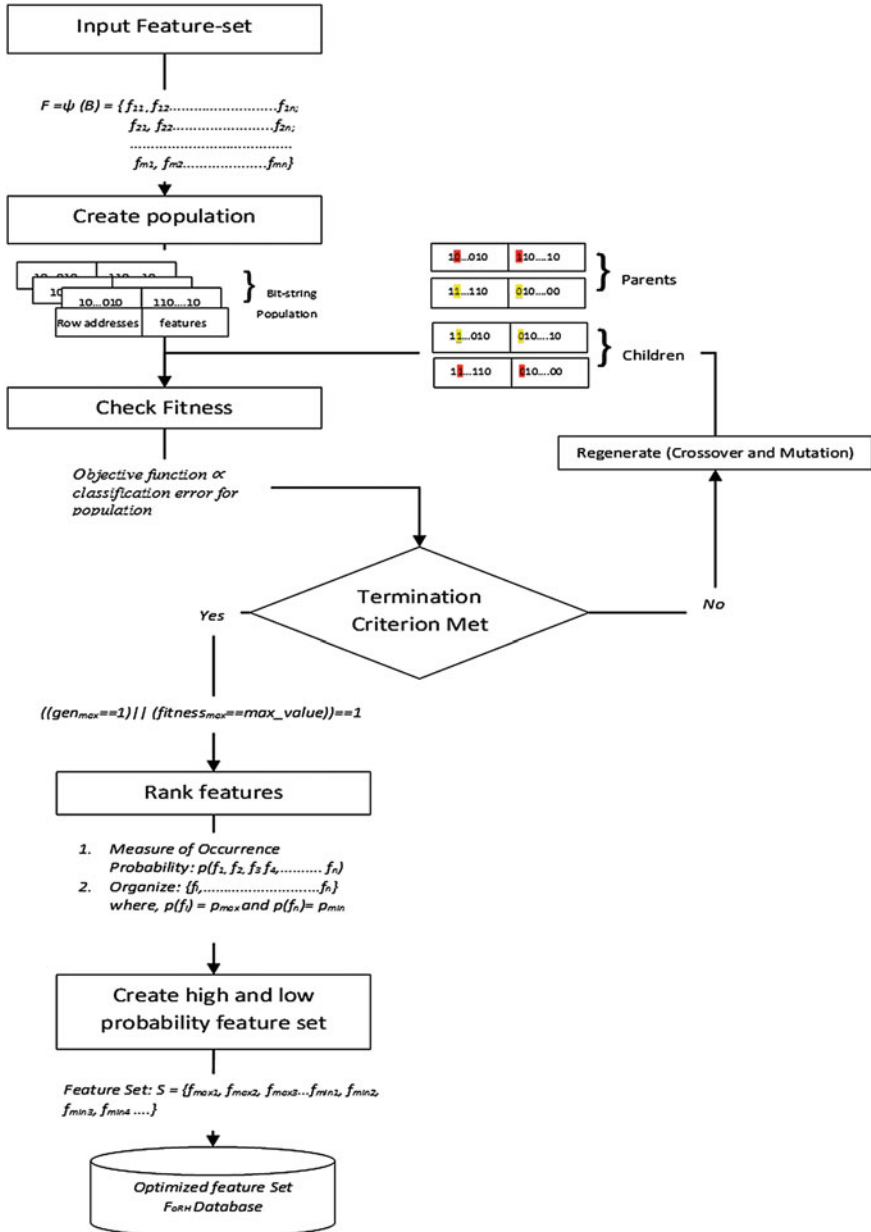


Fig. 1.3 Feature subset optimization process

1.3.2.1 Feature Extraction Process

1. Detect Blobs using chain code algorithm [17] or any conventional method
2. Extract 17 texture/geometrical features: Area, centroid, fullness, major and minor axis, etc.
3. Based on blob dimensions locate it on the colour image
4. Calculate length and width of bounding boxes
5. Sample points proportional dimensions of the blob
6. Arrange extracted points in a matrix form

1.3.2.2 Pixel Sampling Algorithm

1. Based on blob dimensions and locations, make a bounding box on the colour image
2. Calculate length and width of a bounding box
3. Sample points proportional length and width of the blob
4. Remaining samples i.e. (Total samples – (length*width)) are sampled randomly

1.3.3 Feature Subset Optimization

The feature set extracted in the previous step includes plenty of redundant features and has an effect classification error and computational cost. As discussed previously, we propose a method based on a Genetic Algorithm or GA for feature subset optimization. For adapting the GA for our problem we also propose application specific encoding and an additional operator to fine tune the GA results. This subsection discusses the proposed feature optimization technique in detail.

- a. Chromosome Encoding: We propose an encoding which assumes bidirectional search in the feature matrix. Two directions of search include rows and columns of the feature matrix to explore feature space thoroughly. Every chromosome gene represents a row number or feature number. Diagram represents the chromosome encoding.
- b. Crossover and Mutation: The crossover and mutation operations are operations to create children population by swapping genes at fixed and random locations. In case of crossover we propose swapping at row alleles and feature alleles at multiple points to create a new diversified population. In case of mutation the swapping will happen at a random location within an allele.
- c. Fitness Check: We propose using a traditional parameter for the fitness check which is an error in the feature classification. Here, we exclude the other parameters including the number of features because we are proposing the predetermined length which will much lesser than the extracted one.
- d. Selection: We propose a tournament selection approach to select most fit population for the regeneration operations (Fig. 1.4).

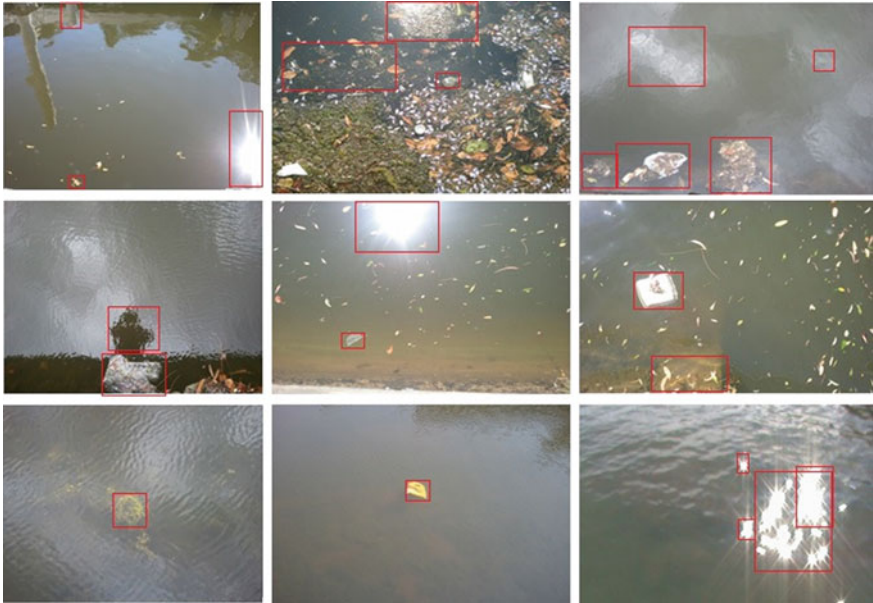


Fig. 1.4 Image samples with detected blobs

e. Termination Criterion: The termination criterion in this case will be number of generations so that every time enough time period is spent before fitness saturation (Figs. 1.5 and 1.6).

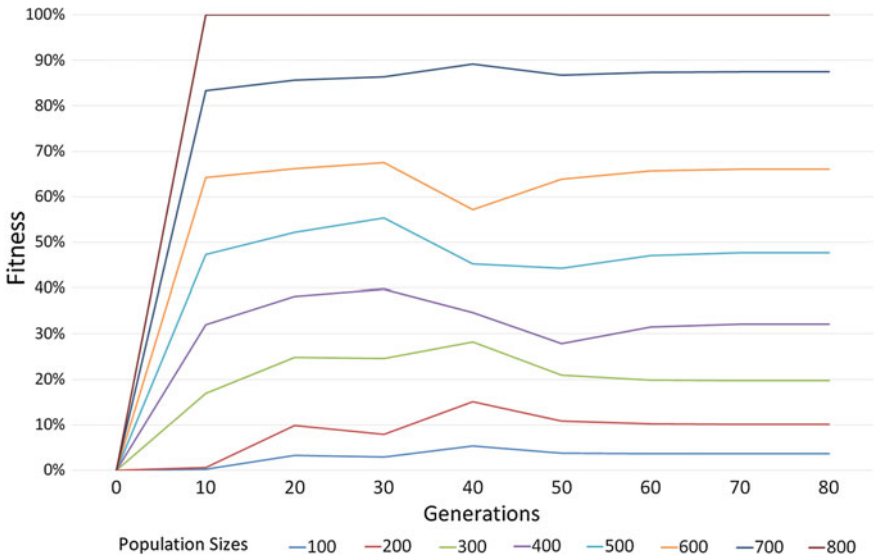


Fig. 1.5 Population size/fitness relationship

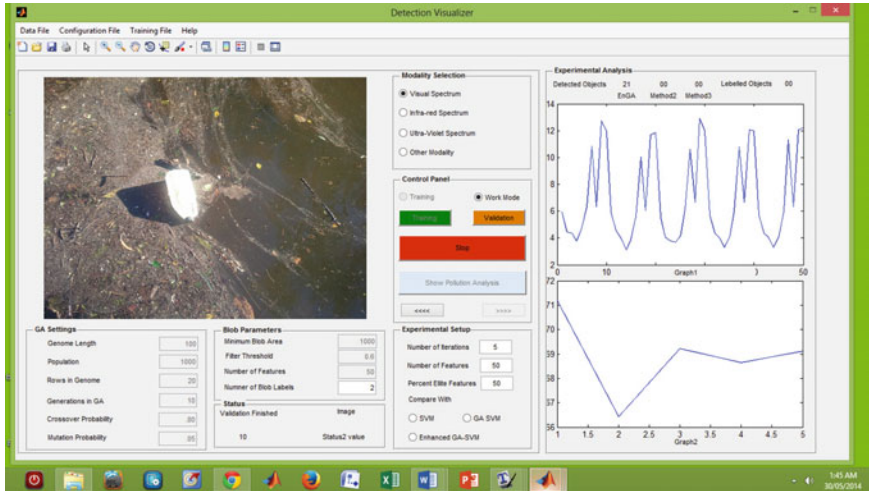


Fig. 1.6 Experimental platform

1.3.4 Feature Classification

Since, focus of this work is on the feature optimization and selection process we propose a most commonly used classifier for such applications i.e. an SVM classifier. The proposed classifier provides following advantages [18]:

- a. Choice to select different kernels to make classification more flexible
- b. Suitable for nonlinear feature combinations
- c. Robustness against partially biased training sets

1.4 Experimental Work

Method:

1. Real Life Data-Sample Collection
2. Select small subset of samples addressing various environmental conditions and presence of various objects as a training data
3. Label the training data with object labels
4. Run the training mode of algorithmic and find the optimized and ranked subset
5. Run the work mode on remaining data samples and select the features found in training mode
6. Classify features and display results (Tables 1.1 and 1.2)

Table 1.1 Accuracy comparison with Benchmark methods

Object type	SVM (%)	GA-SVM (%)	En-GA-SVM (%)
Bottles	53–56	70–78	82–85
Marine plants	53–57	71–76	80–85

Table 1.2 Feature ranking and classification accuracy relationship

High rank features (%)	Low rank features (%)	Classification error (%)
50	50	35–38
60	40	42–47
70	30	54–58
80	20	64–71
90	10	82–85

During the experimental work, a Matlab based implementation was validated by supplying 250 real-life images acquired from a river surface in Sydney region. Algorithm was tested from qualitative and quantitative perspectives. The quality of detection was determined by observing accuracy over the range of variable datasets. The quantitative results were determined by testing the detection for several number datasets. Initial results clearly show better detection accuracy of our approach over traditional classifier based or GA-classifier based solutions (Fig. 1.7).

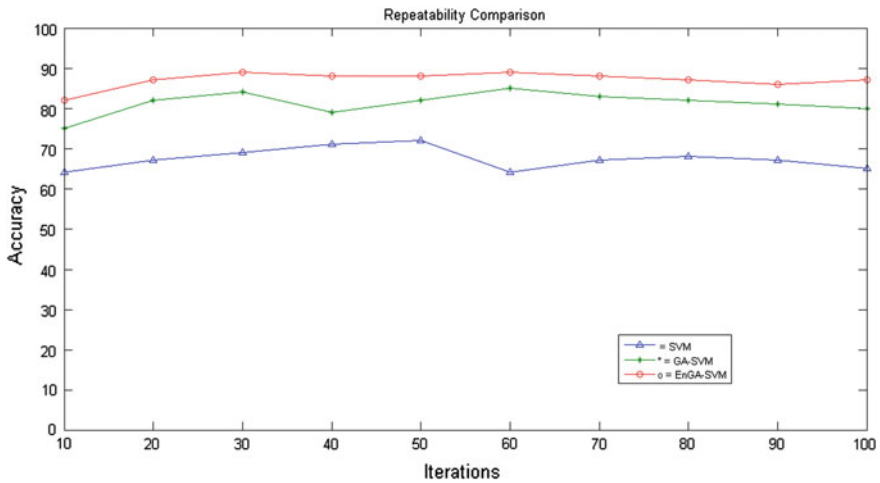


Fig. 1.7 Repeatability comparison of algorithms

1.5 Conclusion

Based on the problem-space studied a method is proposed to serve the need to monitoring of river water surfaces. While achieving this goal, one of the most important consideration is this method should work in dynamic environmental conditions. The proposed feature based detection method proves comparatively efficient and accurate than previous methods. Reduction of highly scalable data allows to reduce curse of dimensionality. It also allows to detect and classify large number of objects very effectively. The ranking scheme presented here provides adaptivity with respect to environmental variations. Use of high rank features provide stability to the algorithm in case of environmental changes. Sampling of variable shapes makes the proposed technique more versatile for large range sizes and shapes of the objects. Bi-directional encoding by including rows and columns in the chromosome enhances search in the larger space. Thus a curse of premature convergence can be minimized with a new encoding approach. Comparative analysis between previously applied techniques with our method shows its superior results repetitively. Depending on the size of the image in megapixels the magnitude of the extracted feature set can be varied. Thus an adaptive and more accurate method is suggested here. Further work in enhancing the capabilities of this method is in progress.

References

1. Sikder, M.T., Kihara, Y., Yasuda, M., Yustiawati, Mihara, Y., Tanaka, S., Odgerel, D., Mijiddorj, B., Syawal, S.M., Hosokawa, T., Saito, T., Kurasaki, M.: River water pollution in developed and developing countries: judge and assessment of physicochemical characteristics and selected dissolved metal concentration. *Clean Soil Air Water* **41**(1), 60–68 (2013). doi:[10.1002/clen.201100320](https://doi.org/10.1002/clen.201100320)
2. Hart, J.K., Martinez, K.: Environmental sensor networks: a revolution in the earth system science? *Earth-Sci. Rev.* **78**(3–4), 177–191 (2006). doi:[10.1016/j.earscirev.2006.05.001](https://doi.org/10.1016/j.earscirev.2006.05.001)
3. Jacobs, N., Burgin, W., Fridrich, N.: The global network of outdoor webcams: properties and applications. In: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS'09, pp. 111–120 (2009), doi:[10.1145/1653771.16537891](https://doi.org/10.1145/1653771.16537891)
4. Fingas, M.F., Brown, C.E.: Review of oil spill remote sensing. *Spill Sci. Technol. Bull.* **4**(4), 199–208 (1997)
5. Plotnik, A.M., Rock, S.M.: Quantification of cyclic motion of marine animals from computer vision. In: *OCEANS'02MTS/IEEE*, vol. 3, pp. 1575–1581 (2002), doi:[10.1109/OCEANS.2002.1191870](https://doi.org/10.1109/OCEANS.2002.1191870)
6. Ferri, F.J., Pudil, P., Hatef, M., Kittler, J.: Comparative study of techniques for large-scale feature selection. *Mach. Intell. Pattern Recognit.* **16**, 403–413 (1994)
7. Yang, J., Honavar, V.: Feature subset selection using a genetic algorithm. *Feature Extraction, Construction and Selection*. Springer, New York (1998)
8. Oliveira, L.S., Benahmed, N., Sabourin, R., Bortolozzi, F., Suen, C.Y.: Feature subset selection using genetic algorithms for handwritten digit recognition. In: *IEEE Proceedings of XIV Brazilian Symposium on Computer Graphics and Image Processing*, pp. 362–369 (2001)
9. Pernkopf, F., O'Leary, P.: Feature selection for classification using genetic algorithms with a novel encoding. *Computer Analysis of Images and Patterns*. Springer, Berlin (2001)

10. Sun, Z., Bebis, G., Miller, R.: Object detection using feature subset selection. *Pattern Recognit.* **37**(11), 2165–2176 (2004)
11. Oh, I.-S., Lee, J.-S., Moon, B.-R.: Hybrid genetic algorithms for feature selection. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(11), 1424–1437 (2004)
12. Lin, T.C., Liu, R.S., Chen, S.Y., LiU, C.C., Chen, C.Y.: Genetic algorithms and silhouette measures applied to microarray data classification. In: *APBC*, pp. 229–238 (2005)
13. Zhu, F., Guan, S.: Feature selection for modular GA-based classification. *Appl. Soft Comput.* **4**(4), 381–393 (2004)
14. Lac, H.C., Stacey D.A.: Feature subset selection via multi-objective genetic algorithm. In: *IEEE International Joint Conference on Neural Networks*, vol. 3 (2005)
15. Zhuo, L., Zheng, J., Li, X., Wang, F., Ai, B., Qian, J.: A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine. In: *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images*, pp. 71471J–71471J. *International Society for Optics and Photonics* (2008)
16. Jurie, F., Schmid, C.: Scale-invariant shape features for recognition of object categories. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2 (2004)
17. Xiaolong, D., Khorram, S.: A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geosci. Remote Sens.* **37**(5), 2351–2362 (1999)
18. Campbell, C.: Kernel methods: a survey of current techniques. *Neurocomputing* **48**(1), 63–84 (2002)
19. Koza, J.R.: Survey of genetic algorithms and genetic programming. In: *WESCON/’95, Conference Record, Microelectronics Communications Technology Producing Quality Products Mobile and Portable Power Emerging Technologies*, p. 589 (1995)
20. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: a comprehensive study. *Int. J. Comput. Vis.* **73**(2), 213–238 (2007)
21. Kenji, K., Rendell, L.A.: A practical approach to feature selection. In: *Proceedings of the Ninth International Workshop on Machine Learning*. Morgan Kaufmann Publishers Inc. (1992)
22. Kale, A., Chaczko, Z.: Supervised feature classification for pollution monitoring on river water surface. In: *2nd Asia-Pacific Conference on Computer Aided System Engineering—APCASE* (2014)

Chapter 2

Alternative Approaches for Fast Boolean Calculations Using the GPU

Bernd Steinbach and Matthias Werner

Abstract The growing number of Boolean variables requires very efficient approaches to solve the given tasks. We explore the utilization of the GPU for fast parallel Boolean calculations in this chapter. Hundreds of processor cores of the GPU offer a significant potential for improvements. Constraints in their application may restrict the achievable speedup. This chapter gives a taxonomy about possible approaches to solve a problem using a computer. We select one problem from the Boolean domain and summarize alternative approaches for utilizing the GPU. It will be shown that the calculation time could be reduced by several orders of magnitudes for the selected Unate Covering Problem.

2.1 Introduction

The technological progress in micro- and nano-electronics leads to both a strong extension of applications and growing requirements for the design of digital systems. Boolean functions are the main instrument for their description. It is well known that the number of function values of a Boolean function exponentially grows depending on the number of Boolean variables.

An important source for improvements to solve Boolean tasks is the utilization of many computer cores for parallel computations. Today's processors contain a small number of cores in the *Central Processing Unit* (CPU), but significantly more cores are available on the *Graphics Processing Unit* (GPU). Hence, the utilization of the GPU to solve exponentially complex tasks of the Boolean domain is an important challenge for scientists and engineers.

There are several approaches for utilizing the GPU in the Boolean domain. In this chapter, we give a taxonomy to classify these approaches. Due to the restricted

B. Steinbach (✉) · M. Werner
Institute of Computer Science, Freiberg University of Mining and Technology,
09596 Freiberg, Germany
e-mail: steinb@informatik.tu-freiberg.de

M. Werner
e-mail: werner3@mailserver.tu-freiberg.de

© Springer International Publishing Switzerland 2015
G. Borowik et al. (eds.), *Computational Intelligence and Efficiency
in Engineering Systems*, Studies in Computational Intelligence 595,
DOI 10.1007/978-3-319-15720-7_2

space, we select a single Boolean problem, explore different solution methods based on the mentioned taxonomy, and compare both the necessary effort and the benefit achieved. The results of these comparisons can guide scientists and engineers who want to speed up other Boolean problems using a GPU.

2.2 The Explored Boolean Problem: Unate Covering

We select *the Unate Covering Problem* (UCP) of given *Petrick Functions* as object of our exploration. This problem has on the one hand an exponential complexity; on the other hand it has a high practical significance; e.g., it is needed for circuit design [3] and data mining [2]. A *Petrick Function* $P(\mathbf{p})$ depends on non-negated variables p_i within a conjunctive form as shown in the following example:

$$\begin{aligned} P(\mathbf{p}) = & (p_4 \vee p_5 \vee p_6 \vee p_8) \wedge (p_2 \vee p_3 \vee p_4 \vee p_7 \vee p_8) \wedge \\ & (p_1 \vee p_3 \vee p_4 \vee p_7 \vee p_8) \wedge (p_1 \vee p_4 \vee p_5 \vee p_7 \vee p_8) \wedge \\ & (p_1 \vee p_2 \vee p_5 \vee p_6) \wedge (p_4 \vee p_5 \vee p_6 \vee p_7 \vee p_8) \wedge \\ & (p_1 \vee p_4 \vee p_5 \vee p_6 \vee p_7 \vee p_8) \wedge (p_4 \vee p_6 \vee p_7) = 1. \end{aligned} \quad (2.1)$$

The aim of the Unate Covering Problem consists in finding a subset of variables p_i such that values 1 of these variables determine the value 1 of the given Petrick Function $P(\mathbf{p})$. A minimal solution is a subset of variables p_i which cannot be reduced without losing the covering of all disjunctions (also called clauses). The 12 minimal solutions of 2 or 3 variables p_i which satisfy the equation $P(\mathbf{p}) = 1$ given above, are:

$$\begin{aligned} & p_1 p_4 \vee p_2 p_4 \vee p_4 p_5 \vee p_4 p_6 \vee p_1 p_2 p_6 \vee p_1 p_3 p_6 \vee \\ & p_3 p_5 p_6 \vee p_6 p_7 \vee p_6 p_8 \vee p_1 p_7 p_8 \vee p_5 p_7 \vee p_2 p_7 p_8 = 1. \end{aligned} \quad (2.2)$$

Exact minimal solutions of the Unate Covering Problem are minimal solutions consisting of the smallest number of variables. Selected from the set of 12 minimal solutions, the 7 *exact minimal solutions* of 2 variables p_i are wanted:

$$p_1 p_4 \vee p_2 p_4 \vee p_4 p_5 \vee p_4 p_6 \vee p_6 p_7 \vee p_6 p_8 \vee p_5 p_7 = 1. \quad (2.3)$$

It is our aim to find *all* exact minimal solutions of a given Petrick Function.

2.3 Advantages and Drawbacks of the GPU

GPUs were basically developed to accelerate the graphical representation of wanted data on the screen. Hence, GPUs naturally have to transform geometry and texture data to colored pixels. The mathematical basis for such transformations is the multiplication of matrices. Such tasks can be efficiently solved in parallel on many processor cores, because most of the computations are independent of each other

due to the data locality. GPUs have been heavily optimized for such tasks over the years. This has been achieved by throughput-oriented, many-core architectures with hundreds of compute cores. In contrast, modern CPUs have a multi-core architecture (e.g., 8 cores) where each core has to execute single instructions as fast as possible.

The main advantage of the GPU is the much higher number of cores in comparison with the CPU. Therefore, it is an interesting challenge to port a very computation-intensive task from the CPU to the GPU.

Drawbacks of the GPU are caused by the differences in the architecture and programming paradigm. Due to Amdahl's law [1], the theoretically maximal speedup is determined by the sequential part of a program even when an infinite number of processor cores can be utilized. Consequently, significant speedups can be reached on the GPU only for programs with a large parallel part.

The following analogy emphasizes differences between sequential and parallel programming. Sequential programming on a single CPU core is like employing a single, skillful and fast workman. Instructions are executed step by step on single data (see *Single Instruction, Single Data*, SISD, [5]). Parallel and concurrent programming on the GPU is like managing many, but slow workers. They should not obstruct each other, even though they have to share scarce resources. If some of the workers have to wait for resources, others stand in to conceal the latencies. The efficiency of such a team depends on the programmer who faces much more challenges than a sequential programmer. The GPU programmer is responsible for cache coherency, optimal work balance and utilization of memory hierarchies—e.g., coalesced memory access is crucial on a GPU.

A GPU consists of several, concurrently acting *Multiprocessors* (MPs). Each of these MPs is equipped with shared memory, very fast thread-private registers, load and store units and many compute cores which work in parallel. On Nvidia GPUs the workers are lightweight threads, where each 32 threads are grouped in a warp. Instructions are executed by warps based on the *Single Instruction, Multiple Data* principle, (SIMD) [5]. Generally, 32 threads of a warp execute the same instruction at the same time on different data. Warp instructions are serialized by the hardware, if divergent control flows or memory conflicts are present.

The PCIe bus transfers the data between the main memory of the CPU and the device memory of the GPU. This data transfer is restricted by the bandwidth of PCIe bus of, e.g., 8 GB/s for PCIe v2.x. The GPU itself has internal memory layers designed by decreasing both the size and the latency.

Valuable hints regarding optimal algorithms and programs on Nvidia GPUs are given in [19].

2.4 Classification of Approaches for the Utilization of GPUs

Figure 2.1 shows a taxonomy which can be used to classify how a certain task can be solved on a computer. Firstly, the utilized device for computation must be selected. The device, which mainly contributes to solve the problem, can be the CPU or

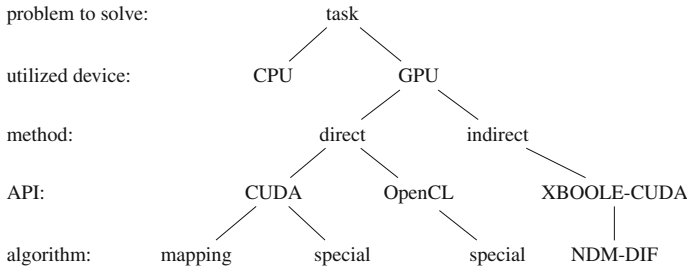


Fig. 2.1 Classification of the explored approaches

the GPU. Due to the focus of this chapter, we explore the further taxonomy only for the GPU. The CPU is only used to control the GPU and as reference for the speedup reached on the GPU. For that reason we skip the classification details for the CPU.

Secondly, we decide about the method how the device (GPU) is utilized to solve the task. The direct method faces the programmer with the technical details of the GPU and significantly increases the effort to implement the needed program. The indirect method simplifies the programming effort because general domain-specific operations wrap all details utilizing the GPU.

The third level of the suggested taxonomy deals with the *Application Programming Interface* (API). Such an API provides a set of data types and functions needed to serve a certain purpose, which is the utilization of the GPU in the studied case. The *Compute Unified Device Architecture* (CUDA) [4] is an API provided by Nvidia. Alternatively, the API *Open Computing Language* (OpenCL) [6] can be used to utilize GPUs not only from Nvidia. OpenCL is an API that allows access to heterogeneous platforms consisting of GPUs of different producers, CPU cores, and even *Digital Signal Processors* (DSPs) or *Field-Programmable Gate Arrays* (FPGAs) in the used hardware configuration.

The final question of this taxonomy asks for the algorithm to use. Typically, as for the explored problem, there are many different algorithms [16]. We select two algorithms for CUDA and refer to [14, 17] where further CUDA-algorithms are evaluated which solve the studied problem. Our decision for these two algorithms is based on the ability of the GPU. Knowing that the GPU is heavily optimized for matrix multiplication, in [8] an algorithm was developed that maps the UCP to the multiplication of matrices. Alternatively, we can utilize the properties of the UCP in a special algorithm for the GPU [13]. For comparison of CUDA and OpenCL the same special algorithm was implemented using OpenCL [7]. The domain-specific API XBOOLE-CUDA [18] provides, among others, the operations *Negation according to De Morgan* (NDM) and the set operation *Difference* (DIF) which can be utilized to solve the Unate Covering Problem with significantly less effort for the implementation of the algorithm.

2.5 Direct Utilization of the GPU for Solving the Unate Covering Problem

2.5.1 Matrix-Multiplication Using CUDA

The Petrick Function $P(\mathbf{p})$ is mapped to the matrix P of n rows and k columns. The rows are associated top down to the variables p_i using an increasing order: p_1, \dots, p_n . Each column of the matrix P represents one clause of $P(\mathbf{p})$ where a value 1 indicates the existence of the associated variable in the clause.

All vectors of the Boolean space are assigned top down to the matrix A of $l = 2^n$ rows and n columns. The elements of the result matrix $R = A \times P$ of $l = 2^n$ rows and k columns can be calculated as usual:

$$R[r, c] = \sum_{i=1}^n A[r, i] \cdot P[i, c]. \quad (2.4)$$

The value of $R[r, c]$ indicates how many variables of the row r of A cover the clause associated to the column c of P . Hence, if $R[r, c] = 0$ then the row r of A is no solution of the Unate Covering Problem. The evaluation of the matrix R detects all exact minimal solutions. Figure 2.2 shows an example of this approach. The rows of bold numbers in R indicate valid covers. The set $\{(101), (111)\}$ of associated rows of the matrix A contains the row (101) with the smallest number of values 1 and determines the only exact minimal cover $p_1 p_3$.

A subset of variables p_i is a valid cover when all clauses are covered. It is not needed to know how often each clause is covered, therefore the algebraic matrix multiplication (2.4) can be substituted by the Boolean matrix multiplication (2.5):

$$RB[r, c] = \bigvee_{i=1}^n A[r, i] \wedge P[i, c]. \quad (2.5)$$

Fig. 2.2 Exact minimal cover $p_1 p_3$ of $P(\mathbf{p})$ found by matrix multiplication

$$P(\mathbf{p}) = (p_1) \wedge (p_1 \vee p_2) \wedge (p_2 \vee p_3) \wedge (p_3)$$

$$R = A \times P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 1 & 2 & 1 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{2} & \mathbf{1} \end{bmatrix}$$

The advantage of the Boolean matrix multiplication (2.5) in comparison to (2.4) is that a simpler data type reduces the memory for the matrices. However, the number of rows in the matrices A and R exponentially grows depending on the number of variables p_i and restricts this approach to small instances of the UCP.

2.5.2 Ordered Restricted Vector Evaluation Using CUDA

The achievable speedup does not only depend on the number of used processor cores, but also on the implemented algorithm. The ordered restricted vector evaluation is a very powerful algorithm for the Unate Covering Problem. Some intermediate steps help to understand this approach.

Searching for more powerful algorithms we compare the given Petrick Function, e.g. (2.1), with the expression of the corresponding exact minimal cover (2.3). This comparison shows the transformation from the given conjunctive form into an equivalent minimal disjunctive form. This transformation is realized by the distributive law:

$$(a \vee b) \wedge (c \vee d) = a c \vee a d \vee b c \vee b d. \quad (2.6)$$

The application of the absorption law:

$$a \vee a b = a \quad (2.7)$$

reduces the calculated form to conjunctions of the minimal cover. The exact minimal cover can be found by counting the number of variables in the conjunctions and the selection of the conjunctions having a minimal number of variables. It was shown in [16] that a repeated application of (2.7) after the calculation of (2.6) for the intermediate result and a single clause reduces the runtime by a factor of more than 10^4 . Therefore, we use this significantly improved algorithm as basis for all further comparisons.

The theoretical basis of another approach is shown in (2.8). Two consecutive negations do not change $P(\mathbf{p})$. The inner negation can be executed in constant time according to *De Morgan's Law* (NDM). The outer negation must be executed as *Complement* operation (CPL).

$$\begin{aligned} P(\mathbf{p}) &= 1 \\ \overline{\overline{P(\mathbf{p})}} &= 1 \\ \overline{NDM(P(\mathbf{p}))} &= 1 \\ CPL(NDM(P(\mathbf{p}))) &= 1 \end{aligned} \quad (2.8)$$

A proof in [14] shows that an algorithm based on (2.8) solves the UCP. Using the XBOOLE operations NDM and CPL on a single CPU core, the time to solve an Unate Covering Problem could be reduced by a factor of almost 10^5 .

The CPL(NDM(P))-approach achieved a strong improvement. This algorithm needs almost all time for the calculation of the complement operation that must take

into account all 2^n elements of the Boolean space \mathbb{B}^n . The wanted exact minimal solutions are not distributed over the whole Boolean space, but are characterized by a fixed number of values 1. Hence, the division of the Boolean space in certain subspaces is a starting point for further improvements.

Definition 2.1 The function $f(\mathbf{p})$ is symmetric with regard to two variables $\{p_i, p_j\}$ if

$$f(p_i, p_j, \mathbf{p}_0) = f(p_j, p_i, \mathbf{p}_0). \tag{2.9}$$

Definition 2.2 The function $S_i(\mathbf{p})$ is a symmetric function that is symmetric with regard to each pair of variables. The index i indicates the number of non-negated variables in their conjunctions.

There are $n + 1$ symmetric functions $S_i(\mathbf{p})$ in each Boolean space \mathbb{B}^n . For $n = 4$ we have, for instance,

$$\begin{aligned} S_0(p_1, p_2, p_3, p_4) &= \bar{p}_1 \bar{p}_2 \bar{p}_3 \bar{p}_4 \\ S_1(p_1, p_2, p_3, p_4) &= p_1 \bar{p}_2 \bar{p}_3 \bar{p}_4 \vee \bar{p}_1 p_2 \bar{p}_3 \bar{p}_4 \vee \bar{p}_1 \bar{p}_2 p_3 \bar{p}_4 \vee \bar{p}_1 \bar{p}_2 \bar{p}_3 p_4 \\ S_2(p_1, p_2, p_3, p_4) &= p_1 p_2 \bar{p}_3 \bar{p}_4 \vee p_1 \bar{p}_2 p_3 \bar{p}_4 \vee p_1 \bar{p}_2 \bar{p}_3 p_4 \vee \\ &\quad \bar{p}_1 p_2 p_3 \bar{p}_4 \vee \bar{p}_1 p_2 \bar{p}_3 p_4 \vee \bar{p}_1 \bar{p}_2 p_3 p_4 \\ S_3(p_1, p_2, p_3, p_4) &= p_1 p_2 p_3 \bar{p}_4 \vee p_1 p_2 \bar{p}_3 p_4 \vee p_1 \bar{p}_2 p_3 p_4 \vee \bar{p}_1 p_2 p_3 p_4 \\ S_4(p_1, p_2, p_3, p_4) &= p_1 p_2 p_3 p_4. \end{aligned}$$

Figure 2.3 shows how the five symmetric functions $S_i(\mathbf{p})$ structure the Boolean space \mathbb{B}^n where edges connect elements which differ in one position. The following theorems guide us to a more powerful algorithm for the UCP.

Theorem 2.1 For each Boolean space of n variables it holds:

$$\bigvee_{i=0}^n S_i(\mathbf{p}) = 1. \tag{2.10}$$

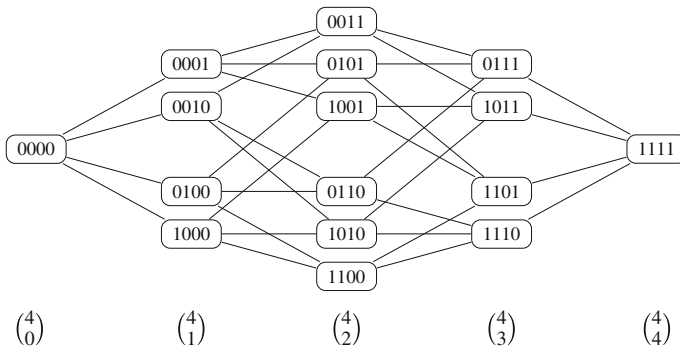


Fig. 2.3 The Boolean space \mathbb{B}^4 structured with regard to the number of values 1

Theorem 2.2 For each Petrick Function $P(\mathbf{p}) \neq 0$ of n variables it holds:

$$P(\mathbf{p}) \wedge S_0(\mathbf{p}) = 0. \quad (2.11)$$

Theorem 2.3 For each Petrick Function $P(\mathbf{p}) \neq 0$ of n variables it holds:

$$P(\mathbf{p}) \wedge S_n(\mathbf{p}) = S_n(\mathbf{p}). \quad (2.12)$$

Theorem 2.1 directly follows from Definition 2.2. Theorem 2.2 holds because $P(\mathbf{p}) \neq 0$ cannot be covered without any non-negated variable p_i . Theorem 2.3 holds because a $P(\mathbf{p}) \neq 0$ does not include any negated variable p_i .

Theorem 2.1 is the key to split the calculation of the complete complement into $n + 1$ difference operations ($DIF(f, g) = f \wedge \bar{g}$) between $S_i(\mathbf{p})$ and $NDM(P(\mathbf{p}))$:

$$\begin{aligned} P(\mathbf{p}) &= 1, \\ \overline{\overline{P(\mathbf{p})}} &= \overline{NDM(P(\mathbf{p}))} = 1 \wedge \overline{NDM(P(\mathbf{p}))} = 1, \\ \bigvee_{i=0}^n [S_i(\mathbf{p})] \wedge \overline{NDM(P(\mathbf{p}))} &= \bigvee_{i=0}^n [S_i(\mathbf{p}) \wedge \overline{NDM(P(\mathbf{p}))}] = 1, \\ \bigvee_{i=0}^n DIF(S_i(\mathbf{p}), NDM(P(\mathbf{p}))) &= 1. \end{aligned} \quad (2.13)$$

The evaluation of the symmetric functions S_i can be organized in increasing order. In the visualization of a Boolean space as shown in Fig. 2.3, this ordered procedure evaluates the columns of binary vectors from the left to the right. Due to Theorem 2.2 the evaluation of S_0 can be skipped. Due to the increasing order of the evaluated symmetric functions S_i , the first non-empty solution set of the difference operation of S_i contains all wanted exact minimal solutions of the UCP. Hence, all other difference operations of S_j , $j > i$, can be skipped. A precise representation of this approach is given in Algorithm 1. This algorithm terminates due to Theorem 2.3.

Algorithm 1: Solve the UCP by Ordered Evaluation of Symmetric Functions

Input: Petrick Function $P(\mathbf{p}) = d_1(\mathbf{p}) \vee \dots \vee d_k(\mathbf{p})$

Output: all exact minimal solutions $AEMS$ of $P(\mathbf{p}) = 1$

```

1 begin
2    $AEMS \leftarrow \emptyset$ 
3    $i \leftarrow 1$ 
4   while  $AEMS = \emptyset$  do
5     generate  $S_i(\mathbf{p})$ 
6      $AEMS \leftarrow DIF(S_i(\mathbf{p}), NDM(P(\mathbf{p})))$ 
7      $i \leftarrow i + 1$ 

```

The symmetric function $S_i(\mathbf{p})$ is generated in line 4 of Algorithm 1 and contains all permutations of i values 1 within an n -bit vector. Due to this method of generation

we call it permutation vector pv . Similarly, each clause of the Petrick Function $P(\mathbf{p})$ can be represented as an n -bit binary vector where a value 1 in the position j indicates the appearance of the variable p_j in the clause. We call the vector of all clauses of a Petrick Function $P(\mathbf{p})$ clause vector cv .

The operations in line 6 of Algorithm 1 require the comparison of each element of pv with each element of cv . The necessary and sufficient condition that the n -bit binary vector $pv[j]$ covers all $cv.elements$ clauses of the Petrick Function $P(\mathbf{p})$ is:

$$\forall cv[k] \in cv : \quad pv[j] \wedge cv[k] \neq 0. \quad (2.14)$$

The \forall -quantifier of (2.14) provides one more possibility to restrict the computation effort. The vector $pv[j]$ is no solution if the result of $pv[j] \wedge cv[k] = 0$ for one value of k . Hence, all further evaluations for such a vector $pv[j]$ can be skipped.

The calculation of (2.14) can be realized in parallel on the GPU for the different permutation vectors $pv[j]$ as shown in Algorithm 2.

Using CUDA [4, 19] we have implemented a program [13, 17] in which Algorithm 2 is used to realize the main step 6 of Algorithm 1. The number of permutations $\binom{n}{i}$ can achieve such a large value that memory conflicts occur. We avoid this problem in our implementation by splitting $S_i(\mathbf{p})$ into slices of a fixed maximal value. These slices are sequentially evaluated on the GPU.

Algorithm 2: $sv = \text{BDIF_kernel}(pv, cv)$ for the GPU

Input: the permutation vector pv of a symmetric function $S_i(\mathbf{p})$ and the clause vector cv of the Petrick Function $P(\mathbf{p})$

Output: solution vector $sv(\mathbf{p})$ which holds $P(\mathbf{p}) = 1$

```

1 begin
2   for  $ic \leftarrow 0, ic < cv.elements, ic \leftarrow ic + 1$  do
3     if  $pv.vector[ip] \wedge cv.vector[ic] = 0$  then
4       break
5   if  $ic = cv.elements$  then
6      $sv.vector[is] \leftarrow pv.vector[ip]$  ▷ add solution
7      $is \leftarrow is + 1$ 

```

2.5.3 Ordered Restricted Vector Evaluation Using OpenCL

The APIs CUDA [4, 19] and OpenCL [7] require different implementation details but realize the same paradigm for parallel programs. Therefore, the powerful algorithm of the previous subsection can be implemented using OpenCL. The execution of both the CUDA and the OpenCL implementation of the same algorithm on the same GPU allows a comparison of the influence of these APIs to the needed runtime.

Using Algorithm 2 to realize the main step 6 of Algorithm 1, an OpenCL program was implemented in [7]. The advantage of this OpenCL implementation is that it

can be used for different GPUs and even for multi-core CPUs. Experiments of [7] confirm the flexibility of OpenCL with regard to the utilization of different hardware resources.

A more efficient data management leads for small benchmarks to faster calculations in comparison with the CUDA implementation of [13, 17]. However, for the largest benchmark of 32 Boolean variables and 1024 clauses in the Unate Covering Problem the OpenCL implementation needs approximately twice the time of the CUDA implementation using the same GPU Tesla C2070. This result confirms the advantage of a well-developed special API for a restricted set of GPUs.

2.6 Indirect Utilization of the GPU for Solving the Unate Covering Problem

2.6.1 Implementation of XBOOLE Using CUDA

XBOOLE is a library of more than 100 functions which can be used within programs written in C or C++ to solve a wide field of Boolean problems [9, 15]. In order to make this chapter self-contained, we give a very brief introduction to XBOOLE.

XBOOLE uses the dash element ($-$) to express the combination of the Boolean values 0 and 1. A ternary vector with d dash elements represents 2^d binary vectors. In this way the needed memory to store sets of binary vectors and the time for their computation can be reduced exponentially. The *Ternary Vector List* (TVL) is the main data structure. An orthogonal TVL does not contain any binary vector in more than one ternary vector. Most of the XBOOLE-operations compute an orthogonal TVL. Each non-orthogonal TVL can be transformed into an orthogonal TVL using the XBOOLE-operation ORTH.

A TVL can be considered as a set of binary vectors. The set operations:

CPL	the Complement \bar{A} ,
ISC	the Intersection $A \cap B$,
UNI	the Union $A \cup B$,
DIF	the Difference $A \setminus B$,
SYD	the Symmetrical Difference $A \Delta B$, and
CSD	the Complement of the Symmetrical Difference $A \bar{\Delta} B$

can be used to compute needed new sets of binary vectors.

Using the form attribute a TVL is able to represent a Boolean function by each of the four basic forms [10]:

Disjunctive Form (D),
Conjunctive Form (K),
Antivalence Form (A), or
Equivalence Form (E).

A benefit of an orthogonal TVL is that such a TVL in ODA form can be used in both D- or A-form. Dual properties are valid for the OKE-form.

Boolean operations between functions are directly realized by the introduced set operations. In case of an ODA-form we have the following association:

CPL negation (\bar{f}),
 UNI disjunction ($f \vee g$),
 ISC conjunction ($f \wedge g$),
 DIF difference ($f \wedge \bar{g}$),
 SYD antivalence ($f \oplus g$), and
 CSD equivalence ($f \odot g$).

The Boolean Differential Calculus [9, 11, 15] extends the Boolean Algebra by operations which evaluate certain changes of Boolean values.

XBOOLE directly provides all k -fold derivative operations:

DERK the k -fold derivative,
 MINK the k -fold minimum,
 MAXK the k -fold maximum;

and all vectorial derivative operations:

DERV the vectorial derivative,
 MINV the vectorial minimum,
 MAXV the vectorial maximum.

The portable source code of XBOOLE is used to provide programming libraries for several types of CPUs and versions of operating systems. The time to solve complex Boolean problems can be reduced when time-consuming operations of XBOOLE are executed on the GPU.

Following this idea, a compatible library XBOOLE-CUDA was implemented in [18] using CUDA. In this way all recent and future applications benefit from the speedup of XBOOLE-CUDA and the simple implementation of Boolean algorithms on the high domain-specific level. XBOOLE-CUDA provides the same operations as XBOOLE. Hence, the migration from an XBOOLE-program to an XBOOLE-CUDA-program simply requires the replacement of the library. XBOOLE-CUDA-operations decide by the size of the TVLs whether CPU is used for small TVLs or the GPU accelerates the calculation for large TVLs. Additional operations allow the programmer to customize such decisions.

The speedup achieved by XBOOLE-CUDA strongly depends on the executed operation and the size of the data. In best case a speedup of more than $13 * 10^3$ was realized using a special brute force algorithm. Table 2.1 summarizes the arithmetic average (\emptyset), the standard deviation (σ), as well as the measured minimal and maximal speedup of XBOOLE-CUDA-operations on a graphics board GTX 470 in comparison to the same XBOOLE-operations running on a CPU Intel i7 3.06 GHz.

Table 2.1 Speedups of XBOOLE-CUDA using the GPU GTX 470 (448 Cores)

Operation	\varnothing	σ	Minimum	Maximum
CPL()	5.04×	2.21×	2.11×	8.11×
ISC()	54.90×	7.48×	43.54×	67.32×
UNI()	34.25×	15.57×	13.83×	61.44×
DIF()	19.63×	15.03×	2.22×	43.15×
SYD()	17.61×	5.40×	11.02×	24.74×
CSD()	197.58×	287.55×	3.47×	843.54×
DERK()	29.95×	10.26×	19.90×	52.63×
MINK()	68.78×	30.31×	41.06×	136.75×
MAXK()	56.24×	16.80×	42.25×	95.97×
DERV()	18.84×	2.63×	16.24×	24.01×
MINV()	77.42×	22.93×	55.63×	115.05×
MAXV()	43.19×	6.36×	32.69×	52.09×
ORTH()(xm3)	448.48×	728.29×	21.20×	2196.97×

2.6.2 Difference-Operations of XBOOLE-CUDA

Algorithm 3 shows the simple implementation of the unate covering problem using XBOOLE-CUDA. The theoretical basis for this algorithm is Algorithm 1.

The Petrick Function $P(\mathbf{p})$ is given as object 1 of the memory list in the file `petInput.sdt`. A file of the type `sdt` contains all data belonging to a recent XBOOLE-system. Such files are used to exchange XBOOLE-systems between programs. The LDS-operation in line 2 of Algorithm 3 reads the file `petInput.sdt`. XBOOLE stores all objects in a special box-system and uses pointers for their access. The GET_ML-operation in line 3 of Algorithm 3 delivers the pointer `ctvl` to the TVL of all clauses of $P(\mathbf{p})$.

The NDM-operation in line 6 of Algorithm 1 calculates in each swap of the while-loop the same result. Hence, in Algorithm 3 this NDM-operation is moved into line 4 outside of the while-loop. The wanted vectors of the exact minimal solution must be stored in a separate TVL. The EMPTY operation in line 5 prepares an empty list having all variables of the Petrick Function $P(\mathbf{p})$. The final step for preparation is the initialization of the variable `min_cover` which indicates the index of the symmetric function which must be evaluated.

The XBOOLE-operation NTV in line 7 calculates the number of ternary vectors of the solution-TVL. The loop in lines 7 to 9 terminates when the number of ternary vectors of the solution-TVL is greater than 0. The function `generate_permutations()` in line 8 is implemented in C and uses the XBOOLE-operation SDATV for storing the generated permutation vectors into the TVL with the access pointer `ptvl`.

The main work to solve the Unate Covering Problem is realized by the DIF-operation in line 9. XBOOLE-CUDA uses the CPU for small TVLs of permutation vectors. Time-consuming DIF-operations of larger TVLs of permutation vectors are executed on the GPU. Extremely large TVLs are split into slices so that no memory problems occur. The solution-TVL is registered in the XBOOLE-memory list as object number 2 using the XBOOLE-operation PUT_ML in line 10. The STS operation stores all data of the extended XBOOLE-system in the file `petSolution.std` for later evaluation.

Algorithm 3: `petSolution.std = xb_cuda_ndm_dif_p(petInput.std)`

Input: file `petInput.std` that contains all clauses of of the Petrick Function $P(\mathbf{p})$ as object 1

Output: file `petSolution.std` that additionally contains all exact minimal solutions of $P(\mathbf{p}) = 1$ as object 2

```

1 begin
2   LDS(petInput.std)                                ▷ load task file
3   GET_ML(1, &ctvl)                                  ▷ select P
4   NDM(&ctvl, &ctvl)                                ▷ negate P according to De Morgan's Law
5   EMPTY(&ctvl, &stvl)                              ▷ prepare solution TVL
6   min_cover ← 1
7   while (NTV(&stvl) = 0) do
8     generate_permutations(min_cover++, &ptvl)
9     DIF(&ptvl, &ctvl, &stvl)                        ▷ main task (CPU or GPU)
10  PUT_ML(2, &stl)
11  STS(petSolution.std)                             ▷ store solution file

```

2.7 Experimental Results

The mapping of the UCP to the matrix multiplication is explained in Sect. 2.5.1. The implementation of this mapping-approach in [8] achieves in the best case a speedup of 3.427 (time on GPU: 236.532 ms; time on CPU: 810.594 ms) using the 64 cores of the GPU GeForce 9600 GT in comparison to a single core of the CPU Intel i7 940 (2.93 GHz) solving the benchmark of 16 variables and 256 clauses. The evaluation of all 2^n Boolean vectors restricts this approach to such very small UCPs.

The achieved strong improvement requires much larger benchmarks for time measurement. We used a Petrick Function of 32 variables p_i and 1024 clauses for comparison. All experiments utilized one core of the CPU Intel Xeon X5650 (2.67 GHz) or the 448 cores of a GPU Tesla C2070. Table 2.2 shows the experimental results measured on this hardware. The speedups achieved on the GPU are calculated using the reference introduced in Sect. 2.5.2.

Table 2.2 Benchmark results: Petrick Function of 32 variables p_i and 1024 clauses

Device	Method	API	Algorithm	Time in ms	Speedup	Implementation effort
GPU	Direct	CUDA	2 within 1	20.849	1.2×10^{11}	Expensive
GPU	Direct	OpenCL	2 within 1	43.804	5.7×10^{10}	Expensive
GPU	Indirect	XBOOLE_CUDA	3	213.600	1.2×10^{10}	Minor
CPU	Indirect	XBOOLE	3	786.967	3.2×10^9	Minor

2.8 Conclusions

The large number of cores of a GPU is an important source to reduce the time for hard Boolean problems. There are two ways for utilizing the GPU:

1. the direct implementation using a given API;
2. the indirect implementation utilizing a domain-specific basic software, such as XBOOLE.

The first way provides unrestricted possibilities utilizing the properties of the GPU, but requires an expensive effort for the implementation. We showed that the results of this way significantly depends on the selected algorithm in a range of 3.4 for matrix multiplication to 1.2×10^{11} using CUDA and the algorithm *Ordered Restricted Vector Evaluation*. The implementation of the same algorithm using OpenCL reduces to speedup to one half but extends the usable devices.

The second way allows us to utilize the power of the GPU without exploring all details of GPU-programming. A much simpler XBOOLE-program requires only 37.5 times more time in comparison to the fastest CUDA implementation. The GPU ported library XBOOLE_CUDA needs only 10 times more time in comparison to the fastest CUDA implementation, but is more than 10^{10} times faster than the reference implementation.

References

1. Amdahl, G.M.: Validity of the single processor approach to achieving large scale computing capabilities. In: Proceedings of the April 18–20, 1967, Spring Joint Computer Conference, pp. 483–485. AMC, New York (1967). doi:[10.1145/1465482.1465560](https://doi.org/10.1145/1465482.1465560)
2. Borowik, G.: Data mining approach for decision and classification systems using logic synthesis algorithms. In: Klempous, R., Nikodem, J., Jacak, W., Chaczko, Z. (eds.) Proceedings of the Advanced Methods and Applications in Computational Intelligence, Springer International Publishing, pp. 3–23, ISBN: 9783319014357. (2014). doi:[10.1007/978-3-319-01436-4_1](https://doi.org/10.1007/978-3-319-01436-4_1)
3. Cordone, R., Ferrandi, F., Sciuto, D., Wolfer Calvo, R.: An efficient heuristic approach to solve the unate covering problem. In: Proceedings of the Conference on Design, Automation and Test in Europe, Paris, France, pp. 364–371 (2000)
4. Farber, R.: CUDA Application Design and Development. pp. 1–336, Elsevier LTD, Oxford (2011). ISBN 0123884268

5. Flynn, M.J.: Some computer organizations and their effectiveness. In: IEEE Transactions Computer, vol. 21(9), IEEE Computer Society, Washington, DC, USA, ISSN 0018-9340, pp. 948-960 (1972). doi:[10.1109/TC.1972.5009071](https://doi.org/10.1109/TC.1972.5009071)
6. Gaster, B., Howes, L., Kaeli, D.R., Mistry, P., Schaa, D.: Heterogeneous Computing with OpenCL. Elsevier Science and Technology, pp. 1-296, ISBN 978-0123877666. (2011)
7. Grehl, S.: Vergleich von Implementierungen des Unate Covering Problems mit OpenCL und CUDA. Freiberg University of Mining and Technology, Project-Thesis (2013)
8. Paul, E., Steinbach, B., and Perkowski, M.: Application of CUDA in the Boolean domain for the unate covering problem. In: Steinbach, B. (ed.) Boolean Problems, Proceedings of the 9th International Workshops on Boolean Problems, Freiberg University of Mining and Technology, Freiberg, 16-17 September 2010, pp. 133-142 (2010). ISBN 978-3-86012-404-8
9. Posthoff, Ch., Steinbach, B.: Logic Functions and Equations—Binary Models for Computer Science. Springer, Dordrecht, The Netherlands (2004)
10. Steinbach, B., Posthoff, Ch.: An Extended Theory of Boolean Normal Forms. In: Proceedings of the 6th Annual Hawaii International Conference on Statistics, Mathematics and Related Fields, Honolulu, Hawaii, pp. 1124-1139 (2007)
11. Steinbach, B., Posthoff, Ch.: Boolean differential calculus-theory and applications. In: Journal of Computational and Theoretical Nanoscience, American Scientific Publishers, Valencia, California, USA, ISSN 1546-1955, vol. 7, no. 6, pp. 933-981 (2010)
12. Steinbach, B., Werner, M.: Fast boolean calculations using the GPU. In: Chaczko, Z., Gaol, F.L., Chiu C. (eds.) Proceedings of the 2nd Asia-Pacific Conference on Computer Aided System Engineering APCASE 2014, Book of Extended Abstracts, Bali Dynasty Resort, Bali, Indonesia, 10-12, pp. 86-89, ISBN 978-0-9924518-0-6 February 2014
13. Steinbach, B. and Posthoff, Ch.: Fast calculation of exact minimal unate coverings on both the CPU and the GPU. In: Roberto Moreno-Díaz, Franz Pichler and Alexis Quesada-Arencibia: Proceedings of the Computer Aided Systems Theory—EUROCAST 2013, 14th International Conference, Las Palmas de Gran Canaria, Spain, February 2013, Revised Selected Papers, Part II, Lecture Notes in Computer Science vol. 8112, Springer, pp. 234-241, ISBN: 978-1-612-08292-9, (2013). doi:[10.1007/978-3-642-53862-9_30](https://doi.org/10.1007/978-3-642-53862-9_30)
14. Steinbach, B. and Posthoff, Ch.: Improvements of the construction of exact minimal covers of boolean functions. In: Roberto Moreno-Díaz, Franz Pichler and Alexis Quesada-Arencibia: Proceedings of the Computer Aided Systems Theory—EUROCAST 2011, 13th International Conference, Las Palmas de Gran Canaria, Spain, February 6-11, 2011, Revised Selected Papers, Part II, Lecture Notes in Computer Science Volume 6928, Springer, pp. 272-279, ISBN: 978-3-642-27578-4, (2012). doi:[10.1007/978-3-642-27579-1_35](https://doi.org/10.1007/978-3-642-27579-1_35)
15. Steinbach, B., Posthoff, Ch.: Logic Functions and Equations-Examples and Exercises. Springer Science + Business Media B.V. (2009)
16. Steinbach, B., Posthoff, Ch.: Parallel Solution of Covering Problems— Super-Linear Speedup on a Small Set of Cores. GSTF International Journal on Computing, Global Science and Technology Forum (GSTF), Singapore, ISSN: 2010-2283, vol. 1, No. 2, pp. 113-122 (2011)
17. Steinbach, B., Posthoff, Ch.: Sources and obstacles for parallelization —a comprehensive exploration of the unate covering problem using both CPU and GPU. In: Astola, J., Kameyama, M., Lukac M., and Stankovi R. S. (eds.): GPU Computing with Applications in Digital Logic. Tampere International Center for Signal Processing. TICSP series # 62, Tampere 2012, pp. 63-96, ISBN 978-952-15-2920-7, ISSN 1456-2774
18. Werner, M.: Parallelisierung von XBOOLE-Operationen mit CUDA. Freiberg University of Mining and Technology, Master-Thesis (2014)
19. Wilt, N.: The CUDA Handbook: A Comprehensive Guide to GPU Programming. ISBN: 9780133261509, Pearson Education (2013)

Chapter 3

Technique for Transformation of DL Knowledge Base to Boolean Representation

Grzegorz Borowik and Dariusz Nogalski

Abstract Knowledge Cartography (KC) allows for fast answering of Description Logic (DL) knowledge base queries, but requires expensive preprocessing to represent knowledge in internal representation, i.e., the algorithm for computation of map of concepts as binary signatures is exponential time (however, for taxonomies—as many practical cases have shown—it is at most quadratic time). Preprocessing is already part of DL reasoning process and some computations are pre-calculated before user issues query to knowledge base. Using another method—Tableaux, no knowledge preprocessing is performed, however, all reasoning is done after user issues query. That’s why KC is faster than Tableaux during query answering. The chapter focuses on preprocessing issue for KC. It mainly considers the research on efficient generation of binary signatures and signatures rebuilding by employing the methods for logic synthesis. It has been confirmed that logic synthesis Complement algorithm is efficient when applied to the construction of the map of concepts. The research has shown that strategy of construction should be adjusted depending on ontology size. For smaller ontologies—the non-recursive approach should be used, on the contrary—for larger ontologies—recursive approach with bi-partitioning of the ontology graph. The recursive procedure indicated good scaling for large taxonomies. Another observation was that Complement algorithm works faster for non-sorted CNFs.

Keywords Knowledge representation · Description logics · Knowledge cartography · Logic synthesis · Boolean function complement

G. Borowik (✉)

Institute of Telecommunications, Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warsaw, Poland
e-mail: G.Borowik@tele.pw.edu.pl

D. Nogalski

Division for Cooperation with International Organizations
and the Uniformed Services Solutions, Asseco Poland S.A.,
Jerozolimskie street 123A, 02-017 Warsaw, Poland
e-mail: dariusz.nogalski@gmail.com

© Springer International Publishing Switzerland 2015

G. Borowik et al. (eds.), *Computational Intelligence and Efficiency in Engineering Systems*, Studies in Computational Intelligence 595,
DOI 10.1007/978-3-319-15720-7_3

3.1 Introduction

Knowledge Representation (KR) is a field of Artificial Intelligence which focuses on inventing and developing of languages and formalisms that allow capturing knowledge about the world. It typically depicts objects, object relations, types of objects, and rules/axioms that describe laws and constraints existing within the world. Another important element of KR is automated reasoning where a generic purpose algorithm (a reasoning algorithm) allows solving complex problems stated in KR language. Therefore, a combined term—Knowledge Representation and Reasoning (KRR)—is used.

For reasoning purposes programming paradigms apply, where declarative and imperative computer languages can be distinguished. Imperative language allows specifying algorithm that solves a problem (e.g. Java procedure which is a sequence of steps to derive a solution). The other, declarative language (e.g. Prolog) allows specifying user goal, and subsequently, generic reasoning algorithm finds the solution to satisfy the goal [14].

For solving complex decision problems on which traditional solutions fail (e.g. imperative languages) the KRR can be applied. Thus, KRR is a building block of many expert systems and has many applications ranging from medical diagnosis, knowledge management [1], classification systems, system mediation [16], information fusion [15] to autonomous agents.

One of the KR formalisms to represent Knowledge Bases (KB) is Description Logic (DL) [1]. It is a subset of First Order Logic (FOL), limited to two types of predicates (unary and binary). The unary predicates allow representing concepts (types of objects) and the binary predicates represent relationship among objects. As opposed to a generic set of statements in FOL, DL has decidable decision procedures that make it applicable in computer systems. Thus, DL is a family of languages where each language is determined by the set of allowable logical constructors. Some constructors, however, qualify reasoning to computations of non-polynomial complexity [1].

Knowledge Reasoning in Description Logic is based on the Open World Assumption (OWA) [21] where DL knowledge base consists of two parts: TBox (terminology box) and ABox (assertion box). TBox is a set of axioms—logical formulas that describe relationships between objects. ABox is a set of assertions—logical formulas which describe memberships of particular individuals and binary relationships between particular individuals.

Example. “Family” ontology.

The small example ontology describes family.

TBox represents axioms within the family universe. In this universe, only individuals of type Person (T1) were assumed. It has been stated that each individual Person can be either a Man or a Woman (T2). Additionally, a constraint has been made that there does not exist an individual that is a Man and a Woman at the same time (T3). Noticeably, all elements of the type HealthyChild are also type of

HealthyPerson (T6). The HappyParent is defined as somebody who is a Person and has at least one HealthyChild (T9).

TBox:

- (T1) $\neg Person \equiv \perp$
- (T2) $Person \equiv Woman \sqcup Man$
- (T3) $Woman \sqcap Man \equiv \perp$
- (T4) $Woman \sqsubseteq Person$
- (T5) $Man \sqsubseteq Person$
- (T6) $HealthyChild \sqsubseteq HealthyPerson$
- (T7) $HealthyPerson \sqsubseteq Person$
- (T8) $\exists hasChild.HealthyChild \sqsubseteq Person$
- (T9) $HappyParent \equiv (Man \sqcup Woman) \sqcap \exists hasChild.HealthyChild$
- (T10) $HappyParent \sqcap HealthyChild \equiv \perp$

ABox introduces individuals existing within the universe. Individual John is a Man (A2) and he is in the relation of “hasChild” with Agness (A4) who is a HealthyPerson (A3).

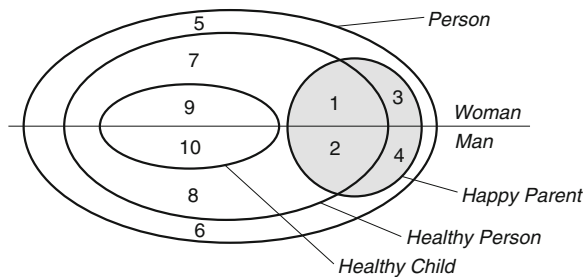
ABox:

- (A1) $Woman(Agness)$
- (A2) $Man(John)$
- (A3) $HealthyPerson(Agness)$
- (A4) $hasChild(John, Agness)$

For the ABox with respect to the TBox the reasoning system should deduce that *John is a HappyParent*. □

Knowledge Cartography [7] reduces each DL problem to Boolean logic algebra over binary vectors. The idea is based on the fact that DL TBox concepts are interpreted as sets over domain of disclosure. Then, relationships among the sets (intersection, disjointness, etc.) can be shown in the Venn diagram—map of concepts (Fig. 3.1), where each TBox concept, represented as a set, is a sum of atomic regions (1–10) of the map. This observation allows defining a binary vector for each

Fig. 3.1 Map of concepts for “Family” ontology (in the form of Venn diagram)



concept. Thus, DL TBox operations, i.e., sum, intersection, etc., can be reduced to respective operations: disjunction, conjunction, etc.; over binary vectors—called signatures.

3.2 Significance

DL is one of the leading (and formally proved) logical formalisms used in the semantic web initiative [27]. It allows defining and developing new generation of languages for describing and interpreting the content of the Internet [29]. In the future, such languages will make possible “understanding” web content by automated software agents. DL supports building knowledge bases [1], in particular, health care knowledge bases [25], battlefield knowledge bases [17]. DL may also be applied in the area of system mediation [16], military information fusion [15], and many others.

Most commonly used method in DL reasoning is the First Order Logic—Tableaux. Using Tableaux no knowledge preprocessing is done and all reasoning is executed after user issues a query.

This chapter addresses, however, an alternative approach to Description Logic reasoning, namely Knowledge Cartography. KC employs preprocessing to convert knowledge into internal representation—a map of concepts, and allows fast answering DL knowledge base queries. Thus, KC is faster than Tableaux, since many DL tasks are reduced to binary vector operations. Nevertheless, the preprocessing phase—generation of map of concepts—is still the part of DL reasoning process and requires expensive computations.

The algorithm for computation of the map of concepts as binary signatures is exponential time as well as signature length grows exponentially. Thus, for some ontologies the length of the concept signature may equal 2^k , where k is the number of concepts in the ontology (assuming an ontology with 50 concepts, the size of the signature may be 2^{50}).

In practical knowledge bases, however, the signature can be computed. For example, for ontologies that are of taxonomy type the signature length grows linearly with the number of concepts [23]. It has been shown in Sect. 3.4 that for examined taxonomies the size of the signature equals $k + 1$, for ontologies of k concepts. Since many practical knowledge bases are taxonomies with some non-taxonomical relations among concepts the Knowledge Cartography method may still be exploited.

The recent research studies [5] have shown that the above described problems can be solved applying the ideas and methods that have proven to be very efficient and effective in the design of digital circuits and logic synthesis [3, 4, 6]. Moreover, it has already been shown that these methods can successfully be adopted to solve typical data problems in data mining [2–5].

Most of the logic synthesis methods were originally developed in the Fifties and Sixties for designing switching circuits, e.g. an evaluation mechanism for checker playing programs [22]. Recently, the machine-learning community has adopted such

approaches [20]. For example, Michie [13] used a manual decomposition of the problem and an expert-assisted selection of examples to construct rules for the concepts in the hierarchy. In comparison with standard decision tree induction techniques, structured induction exhibits about the same degree of classification accuracy with the increased transparency and lower complexity of the developed models. Zupan [24] presented a general-purpose approach for machine-learning. According to this approach, attributes are transformed into new concepts in an iterative manner and create a hierarchy of concepts. In 2003, Lang [12] has suggested using a bi-decomposition and shows its applicability in data mining.

The recent research studies done by Borowik and Łuba [2–4] have shown that many data mining problems result not only from the deficiencies of the tools and computation procedures, but to a large extent are caused by the immature and insuitable algorithms. It has also been shown that the cause of low efficiency of data mining algorithms lies in the commonly used basic mathematical model of transformation of logical formulas [11], which exploits only the fundamental Boolean properties. The use of another model [4], so far used in logic synthesis of digital circuits, can significantly improve the analysis of data by accelerating the computations by a factor of several thousands. For example, the implementation of the Boolean complementation algorithm computes all the reducts for Dermatology database [28] in less than 4 min, whereas Rough Set Exploration System [26] cannot provide the result because of the memory limit when used for this purpose.

These issues highlight the importance of logic synthesis not only in digital design but also in Data Mining and Knowledge Discovery.

A characteristic universal component that links all these issues is the algorithm for transformation of a Boolean function from Conjunctive Normal Form (CNF) to Disjunctive Normal Form (DNF). It can be executed efficiently by applying the algorithms for Boolean function expansion and Boolean function complementation, being part of the Espresso logic synthesis system [6]. This transformation applies to many methods of logic synthesis, e.g. argument reduction, minimization of logic functions, etc., as well as data mining, e.g. feature extraction, discretization, rule induction, etc. [3].

Current knowledge reasoning algorithms for processing the knowledge bases are very efficient for strongly taxonomical structures only (subsumption hierarchy as main relationship between concepts) [7]. Such assumption is very strict and limits ontology expressiveness. As has been shown in [7] the Description Logic reasoning can be reduced to Boolean Logic. The described in the chapter research tries to answer to what extent the efficient algorithm of complementation of Boolean function [4, 6] can be applied to represent large ontologies. Preliminary experiments [5] show that we should be able to process DL ontologies available in practice without strict assumption on taxonomical structure of the ontologies.

3.3 Methodology

Our experimental ontologies were defined using Ontology Web Language [29]. Additionally, Java application capable of reading OWL ontologies using OWL API has been developed.

The procedure transforms ontology DL axiom set into Conjunction Normal Form propositional formula [23] (Table 3.1). These axioms may be class equivalence axioms or class subsumption axioms. Each axiom is represented as one or more rows in CNF matrix (Table 3.2). When the CNF matrix is prepared, we run Complement algorithm [3, 6] to compute Disjunctive Normal Form matrix, which is in fact a map of concepts.

We have designed two procedures of map generation:

- non-recursive—a procedure that creates single CNF matrix for the whole axiom set, and for such CNF matrix generates single DNF matrix (map of concepts)
- recursive—a procedure that splits axiom set recursively, subsequently creates component CNF matrices, runs Complement to compute component DNF matrices, and finally multiplies component DNF matrices to obtain resultant DNF matrix (map of concepts).

To solve the problem of CNF to DNF transformation, the algorithm for Boolean function complementation, being part of the Espresso logic synthesis system, has been applied [3, 6]. The Espresso system has been run with the option “*-Dminterms*” with CNF as input represented in Berkeley’s format [6] (Fig. 3.2a). The computation yields the uncompressed result containing all minterms of considered function, as presented in Fig. 3.2b. Noticeably, each row represents appropriate atomic region presented in Fig. 3.1 and each column represents appropriate signature for each concept.

Table 3.1 Transformation from DL to the propositional formulas (“Family” ontology)

DL formula	Propositional formula
$\neg Person \equiv \perp$	$\neg Z_{Pe} \rightarrow 0$
$Person \equiv Woman \sqcup Man$	$Z_{Pe} \rightarrow Z_W \vee Z_M$ $Z_{Pe} \leftarrow Z_W \vee Z_M$
$Woman \sqcap Man \equiv \perp$	$Z_W \wedge Z_M \rightarrow 0$
$Woman \sqsubseteq Person$	$Z_W \rightarrow Z_{Pe}$
$Man \sqsubseteq Person$	$Z_M \rightarrow Z_{Pe}$
$HealthyChild \sqsubseteq HealthyPerson$	$Z_{HeCh} \rightarrow Z_{HePe}$
$HealthyPerson \sqsubseteq Person$	$Z_{HePe} \rightarrow Z_{Pe}$
$\exists hasChild.HealthyChild \sqsubseteq Person$	$Z_{Ex} \rightarrow Z_{Pe}$
$HappyParent \equiv$ $(Man \sqcup Woman) \sqcap \exists hasChild.HealthyChild$	$Z_{HaPa} \rightarrow (Z_M \vee Z_W) \wedge Z_{Ex}$ $Z_{HaPa} \leftarrow (Z_M \vee Z_W) \wedge Z_{Ex}$
$HappyParent \sqcap HealthyChild \equiv \perp$	$Z_{HaPa} \wedge Z_{HeCh} \rightarrow 0$

Table 3.2 Transformation to Boolean algebra (“Family” ontology)

Propositional formula	Z_{Pe}	Z_W	Z_M	Z_{HePe}	Z_{HeCh}	Z_{Ex}	Z_{HaPa}
$\neg Z_{Pe} \rightarrow 0$	0	–	–	–	–	–	–
$Z_{Pe} \rightarrow Z_W \vee Z_M$	1	0	0	–	–	–	–
$Z_{Pe} \leftarrow Z_W \vee Z_M$	0	–	1	–	–	–	–
	0	1	–	–	–	–	–
$Z_W \wedge Z_M \rightarrow 0$	–	1	1	–	–	–	–
$Z_W \rightarrow Z_{Pe}$	0	1	–	–	–	–	–
$Z_M \rightarrow Z_{Pe}$	0	–	1	–	–	–	–
$Z_{HeCh} \rightarrow Z_{HePe}$	–	–	–	0	1	–	–
$Z_{HePe} \rightarrow Z_{Pe}$	0	–	–	1	–	–	–
$Z_{Ex} \rightarrow Z_{Pe}$	0	–	–	–	–	1	–
$Z_{HaPa} \rightarrow (Z_M \vee Z_W) \wedge Z_{Ex}$	–	–	–	–	–	0	1
	–	0	0	–	–	–	1
$Z_{HaPa} \leftarrow (Z_M \vee Z_W) \wedge Z_{Ex}$	–	–	1	–	–	1	0
	–	1	–	–	–	1	0
$Z_{HaPa} \wedge Z_{HeCh} \rightarrow 0$	–	–	–	–	1	–	1

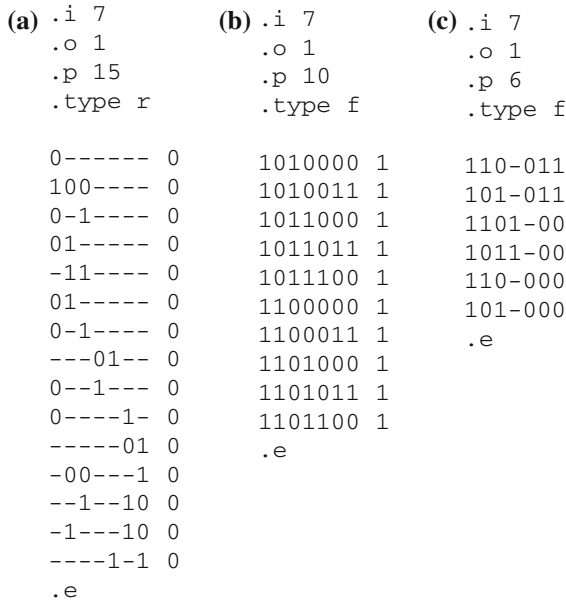


Fig. 3.2 Espresso input and output data for “Family” ontology: **a** CNF, **b** uncompressed DNF, **c** compressed DNF

The non-recursive procedure using Complement algorithm generates uncompressed DNF directly out of CNF. Recursive procedure, however, applies different approach. CNF matrix is splitted into smaller components using bipartite graph partitioning. For each CNF component, appropriate DNF component is computed in the compressed form (Fig. 3.2c). The resultant compressed DNFs are then multiplied recursively together yielding final compressed DNF. To compute the map of concepts the final DNF is decompressed.

The recursive procedure is based on the observation that knowledge base axioms and concepts may be represented by bipartite graph [23] of two kinds of vertexes: (a) axiom vertex, and (b) concept (variable) vertex. The relationship between axiom and concept vertex has meaning “axiom uses concept in its definition”. Such observation is the basis for axiom graph partitioning procedure. In general case (for any ontology), such bi-partitioning problem is NP-complete. But in case of taxonomies, efficient hierarchical graph partitioning procedure can be used, e.g. METIS [10].

3.4 Experiments and Results

Three test cases have been performed. The first test addressed a non-recursive procedure and examined the impact of taxonomy structure on computation time of the map of concepts. In the experiment, a deep and shallow taxonomy trees have been built and computation time analyzed. The second test addressed a non-recursive procedure and examined the problem of sorting the axioms in the axiom set. CNF columns and rows have been sorted and randomly permuted to analyze the impact on the computation time. The third test showed the improvement that brings the recursive procedure to overcome limitations of non-recursive procedure. The analyzed ontologies were mainly taxonomies, but the computation problems for non-taxonomies have also been highlighted.

The **first test** (Fig. 3.3) used non-recursive procedure and was performed for ontologies that are artificial taxonomies (no real life examples) with different division factors (e.g. for division factor = 5 means that each class within taxonomy has five direct disjoint subclasses). We have compared ontologies of different division factors demonstrating the impact of ontology structure on computation time. Flat ontologies (division factor = 5) reach the computational limits later than deep ontologies (division factor = 2). It is due to the fact that flat ontologies provide proportionally more information about disjointness of concepts with respect to total number of axioms than deep ontologies (see factor AT_D/CNF_p in table (Fig. 3.3)). Moreover, information about disjointness reduces signature length.

The non-recursive procedure, however, has serious disadvantages. Even for small taxonomies (approx. 100–200 concepts) the computational limit is quickly reached (exponential time explosion). The discussed later recursive procedure brings large improvements to that problem. As observed by Waloszek [23], it’s due to the fact that for taxonomies, variables within two partitions form two disjoint sets, i.e., variable from one partition doesn’t appear in the other partition (except the cut axiom set).

id	concepts (CNF variables)	division factor (subclass.)	AT_D/ AT_S				DNF gen. time	
			CNF_p	AT_D	CNF_p	AT_S	DNF_p	
2.102	102	2	151	51	0.34	100	103	00:06.000
2.112	112	2	166	56	0.34	110	113	00:25.620
2.122	122	2	181	61	0.34	120	123	01:05.620
2.130	130	2	193	65	0.34	128	131	04:42.340
2.142	142	2	211	71	0.34	140	143	18:06.970
3.120	120	3	237	120	0.51	117	121	00:02.830
3.135	135	3	267	135	0.51	132	136	00:18.280
3.153	153	3	303	153	0.5	150	154	02:07.320
3.162	162	3	321	162	0.5	159	163	06:43.780
3.174	174	3	345	174	0.5	171	175	14:42.750
4.124	124	4	306	186	0.61	120	125	00:00.490
4.144	144	4	356	216	0.61	140	145	00:02.840
4.164	164	4	406	246	0.61	160	165	00:15.460
4.172	172	4	426	258	0.61	168	173	00:24.280
4.184	184	4	456	276	0.61	180	185	00:51.930
4.204	204	4	506	306	0.6	200	205	04:37.710
4.216	216	4	536	324	0.6	212	217	11:15.600
5.155	155	5	460	310	0.67	150	156	00:00.910
5.175	175	5	520	350	0.67	170	176	00:03.190
5.190	190	5	565	380	0.67	185	191	00:06.160
5.205	205	5	610	410	0.67	200	206	00:20.820
5.225	225	5	670	450	0.67	220	226	01:10.820
5.250	250	5	745	500	0.67	245	251	04:33.400
5.280	280	5	835	560	0.67	275	281	29:45.040

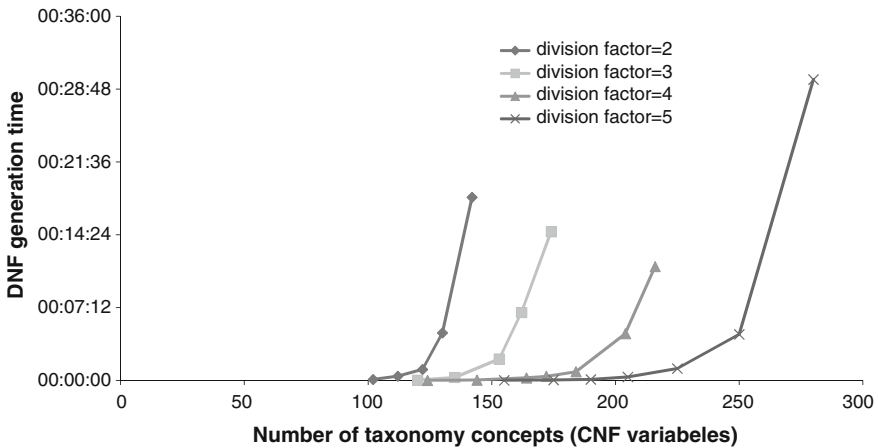


Fig. 3.3 Generation time for taxonomies with different division factor (non-recursive procedure)

id	concepts (CNF variabl.)		div. factor (subcl.)	AT_D/ CNF_p AT_S DNF_p					DNF gen. time
	id	proc.							
3.162	162	sorted	3	321	162	0.5	159	163	06:43.780
3.162	162	shuffled	3	321	162	0.5	159	163	00:29.490
4.204	204	sorted	4	506	306	0.6	200	205	04:37.710
4.204	204	shuffled	4	506	306	0.6	200	205	00:56.260
4.216	216	sorted	4	536	324	0.6	212	217	11:16.000
4.216	216	shuffled	4	536	324	0.6	212	217	03:04.410
5.250	250	sorted	5	745	500	0.67	245	251	04:33.000
5.250	250	shuffled	5	745	500	0.67	245	251	00:41.110
5.280	280	sorted	5	835	560	0.67	275	281	29:45.040
5.280	280	shuffled	5	835	560	0.67	275	281	14:11.230

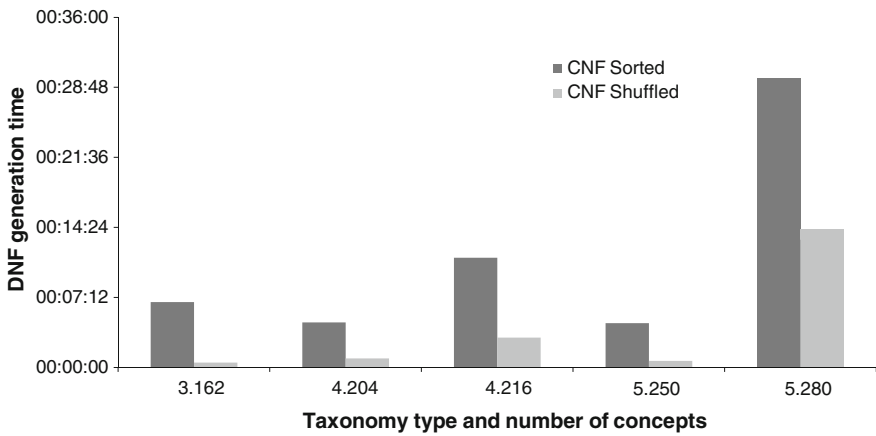


Fig. 3.4 Generation time and comparison of sorted versus not-sorted procedure (non-recursive)

The **second test** used non-recursive procedure (Fig. 3.4) and has been performed to demonstrate that proper order of CNF elements influences computation time. The CNF matrices for analyzed artificial ontologies are sparse matrices containing '0', '1' and '-' elements (symbol '-' is a "don't care" value [6, 8]). In the sorted approach the following steps have been made: (a) CNF rows have been treated as strings and sorted (b) rows have been sorted in the following order: disjoint axioms rows, subsequently equivalence, and finally subsumption. As a result of CNF sorting, values '0' or '1' lay mostly in the matrix diagonal, and the rest of the matrix cells are filled with '-'. In the non-sorted version of the matrix, CNF rows and columns are randomly shuffled. Non-recursive procedure yielded better performance for shuffled CNF matrices than sorted ones. It may be explained by the fact that in shuffled CNF, disjointness axioms are spread around the whole matrix that significantly impacts the Complement algorithm yielding DNF.

id	concepts (CNF variables)		division factor (subclasses)	DNF_p	DNF gen. time	DNF gen. time (log)
	procedure	DNF_p			time	(log)
2.122	122	non-recursive	2	123	00:01:06	1.883
2.122	122	recursive	2	123	00:00:03	0.540
2.142	142	non-recursive	2	143	00:18:06	3.099
2.142	142	recursive	2	143	00:00:04	0.665
2.254	254	non-recursive	2	aborted	aborted	aborted
2.254	254	recursive	2	255	00:00:07	0.908
5.155	155	non-recursive	5	156	00:00:01	0.063
5.155	155	recursive	5	156	00:00:01	0.063
5.225	225	non-recursive	5	226	00:01:12	1.920
5.225	225	recursive	5	226	00:01:09	1.902
5.280	280	non-recursive	5	281	00:29:45	3.315
5.280	280	recursive	5	281	00:01:15	1.938
5.505	505	non-recursive	5	aborted	aborted	aborted
5.505	505	recursive	5	506	00:04:00	2.443

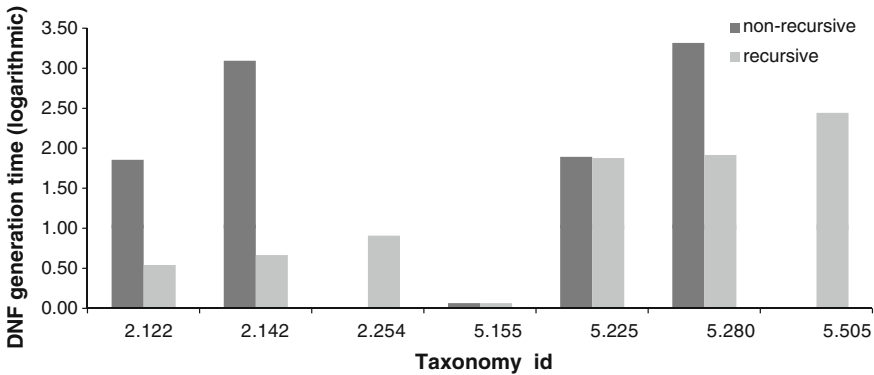


Fig. 3.5 Generation time and comparison of recursive versus non-recursive procedure

The **third test** (Fig. 3.5) has been performed to demonstrate the scalability of recursive procedure as opposed to non-recursive procedure. The non-recursive procedure shows poor scaling, that has already been demonstrated in the first test (Fig. 3.3). The non-recursive procedure can run taxonomical ontologies containing 100–300 concepts depending on the division factor. As shown in table (Fig. 3.5) two computations of non-recursive procedure have been aborted by the user due to long computation time (more than 30 min). In large cases, depending on division factor, computation time is the order of magnitude lower in favor of recursive procedure (e.g. taxonomy id: 2.142, 2.254, 5.280, and 5.505).

3.5 Discussion and Conclusions

Knowledge Cartography [7] may still be used in practice even identified limitations of non-taxonomies where large signatures are generated. As pointed out by Waloszek [23], KC is especially efficient for taxonomical ontologies of large size, or non-taxonomical ontologies with moderate number of concepts and very large number of individuals.

In the chapter has been shown that:

- Non-recursive procedure can only be exploited for small ontologies. For these ontologies, it is much faster than recursive procedure. It is due to complex graph operations in recursive procedures. However, non-recursive procedure cannot handle large taxonomies (depends on division factor) and some way of divide-and-conquer procedure is required.
- Similarly to TreeFusion algorithm [23], the recursive procedure applying the Complement algorithm benefits from good scalability (for large taxonomical ontologies). It's due to the fact that both methods use similar partitioning procedure with fast graph partitioning.
- It has been confirmed that Knowledge Cartography using Complement algorithm manifests itself as a good method for ontologies that are taxonomies.

We also noticed that:

- TreeFusion algorithm [23] does not generate the most efficient compressed DNF matrixes. Proposed Complement algorithm generates matrixes with lesser number of rows. The intuition is that our recursive procedure benefits from optimal compressed DNFs, since DNF multiplications have to execute less number of operations.
- TreeFusion benefits from the fact that examined ontologies are taxonomical [23]. That situation allows applying simple variable sorting that reduces complexity of fusing component trees. However, for ontologies that are not strict taxonomical such sorting cannot be applied. Our hypothesis is that Complement algorithm demonstrates better performance for ontologies that are not strictly taxonomical.

We would like to make a remark that both suppositions require further examinations.

References

1. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F.: The Description Logic Handbook: Theory, Implementation and Applications, 2nd edn. Cambridge University Press, Cambridge (2007)
2. Borowik, G.: Boolean function complementation based algorithm for data discretization. In: Moreno-Díaz, R., Pichler, F., Quesada-Arencibia, A. (eds.) Computer Aided Systems Theory—EUROCAST 2013, Part II. LNCS, vol. 8112, pp. 218–225. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-53862-9_28](https://doi.org/10.1007/978-3-642-53862-9_28)

3. Borowik, G.: Data mining approach for decision and classification systems using logic synthesis algorithms. In: Klempos, R., Nikodem, J., Jacak, W., Chaczko, Z. (eds.) *Advanced Methods and Applications in Computational Intelligence*, Chap. 1. Topics in Intelligent Engineering and Informatics, pp. 3–23. Springer International Publishing (2014). doi:[10.1007/978-3-319-01436-4_1](https://doi.org/10.1007/978-3-319-01436-4_1)
4. Borowik, G., Łuba, T.: Fast algorithm of attribute reduction based on the complementation of Boolean function. In: Klempos, R., Nikodem, J., Jacak, W., Chaczko, Z. (eds.) *Advanced Methods and Applications in Computational Intelligence*, Chap. 2. Topics in Intelligent Engineering and Informatics, pp. 25–41. Springer International Publishing (2014). doi:[10.1007/978-3-319-01436-4_2](https://doi.org/10.1007/978-3-319-01436-4_2)
5. Borowik, G., Nogalski, D.: Efficient technique for transformation from concepts to Boolean algebra. In: Chaczko, Z., Gaol, F.L., Pichler, F., Chiu, C. (eds.) *2nd Asia-Pacific Conference on Computer Aided System Engineering—APCASE 2014*, pp. 24–27. APCASE Foundation, Bali (2014)
6. Brayton, R.K., Hachtel, G.D., McMullen, C.T., Sangiovanni-Vincentelli, A.: *Logic Minimization Algorithms for VLSI Synthesis*. Kluwer Academic Publishers (1984)
7. Goczyła, K., Grabowska, T., Waloszek, W., Zawadzki, M.: The knowledge cartography—a new approach to reasoning over description logics ontologies. In: Wiedermann, J., Tel, G., Pokorný, J., Bieliková, M., Štuller, J. (eds.) *SOFSEM 2006: Theory and Practice of Computer Science. Lecture Notes in Computer Science*, vol. 3831, pp. 293–302. Springer, Berlin (2006). doi:[10.1007/11611257_27](https://doi.org/10.1007/11611257_27)
8. Grzymała-Busse, J.W.: LERS—a system for learning from examples based on rough sets. In: Słowiński, R. (ed.) *Intelligent Decision Support— Handbook of Application and Advanced of the Rough Sets Theory, Theory and Decision Library*, vol. 11, pp. 3–18. Springer, Netherlands (1992). doi:[10.1007/978-94-015-7975-9_1](https://doi.org/10.1007/978-94-015-7975-9_1)
9. Grzymała-Busse, J.W.: LERS—a data mining system. In: Maimon, O., Rokach, L. (eds.) *Data Mining and Knowledge Discovery Handbook*, pp. 1347–1351. Springer (2005). doi:[10.1007/0-387-25465-X_65](https://doi.org/10.1007/0-387-25465-X_65)
10. Karypis, G., Kumar, V.: Parallel multilevel k-way partitioning scheme for irregular graphs. (1999)
11. Komorowski, J., Pawlak, Z., Polkowski, L., Skowron, A.: *Rough Sets: A Tutorial*. (1999)
12. Lang, C., Steinbach, B.: Bi-decomposition of function sets in multiple-valued logic for circuit design and data mining. *Artif. Intell. Rev.* **20**(3–4), 233–267 (2003). doi:[10.1023/B:AIRE.0000006608.31990.cd](https://doi.org/10.1023/B:AIRE.0000006608.31990.cd)
13. Michie, D.: Problem decomposition and the learning of skills. In: Lavrac, N., Wrobel, S. (eds.) *Machine Learning: ECML-95. Lecture Notes in Computer Science*, vol. 912, pp. 17–31. Springer, Berlin Heidelberg (1995). doi:[10.1007/3-540-59286-5_46](https://doi.org/10.1007/3-540-59286-5_46)
14. Nilsson, U., Matuszyński, J.: *Logic, Programming and Prolog*. Previously published by John Wiley & Sons Ltd., 2 edn. (2000)
15. Nogalski, D., Chmielewski, M.: Semantic web service discovery and information fusion using OWL-S and SPARQL formal specifications over NATO JC3IEDM and TIDE services. In: Amanowicz, M. (ed.) *Concepts and Implementations for Innovative Military Communications and Information Technologies*, pp. 165–174. Military University of Technology, Warsaw (2010)
16. Nogalski, D., Najgebauer, A.: Semantic mediation of NATO C2 systems based on JC3IEDM and NFFI ontologies. In: *RTO Symposium “Semantic and Domain based Interoperability” RTO-MP-IST-101*. Oslo (2011)
17. Nogalski, D., Wróblewska, A., Szklarz, P., Chmielewski, M.: Battlefield situational awareness—generating usefull ontology based on relational data model MIP JC3IEDM (in Polish). In: Ficoń, K. (ed.) *Transformacja Systemów Dowodzenia*, pp. 177–194. Ośrodek Badawczo-Rozwojowy Centrum Techniki Morskiej (2010)
18. Papadimitriou, C.H.: *Computational Complexity*. Academic Internet Publishers (2007)
19. Pawlak, Z.: *Rough Sets Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Boston (1991)

20. Rokach, L., Maimon, O.: Data mining using decomposition methods. In: Maimon, O., Rokach, L. (eds.) *Data Mining and Knowledge Discovery Handbook*, pp. 981–998. Springer (2010). doi:[10.1007/978-0-387-09823-4_51](https://doi.org/10.1007/978-0-387-09823-4_51)
21. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*, 3rd edn. Prentice Hall, New jersey (2009)
22. Samuel, A.L.: Some studies in machine learning using the game of checkers. II—recent progress. *IBM J. Res. Dev.* **11**(6), 601–617 (1967)
23. Waloszek, W.: Structural representation of description logics ontologies (in Polish). Ph.D. Thesis, Gdańsk University of Technology, Gdańsk (2007)
24. Zupan, B., Bohanec, M., Demsar, J., Bratko, I.: Feature transformation by function decomposition. *IEEE Intell. Syst. Appl.* **13**(2), 38–43 (1998). doi:[10.1109/5254.671090](https://doi.org/10.1109/5254.671090)
25. NCBO: SNOMED Clinical Terms. <http://bioportal.bioontology.org/ontologies/SNOMEDCT> (2014). Accessed May 2014
26. RSES—Rough Set Exploration System. <http://logic.mimuw.edu.pl/~rses/>. Accessed May 2014
27. SW: W3C Semantic Web Activity. <http://www.w3.org/2001/sw/>. Accessed May 2014
28. UC Irvine machine learning repository. <http://archive.ics.uci.edu/ml/>. Accessed May 2014
29. W3C: OWL Web Ontology Language. <http://www.w3.org/TR/owl-features/>. Release date Feb 2004

Chapter 4

Digital Patterns for Heritage and Data Preservation Standards

Lucia Carrion Gordon and Zenon Chaczko

Abstract This research covers the digital preservation concept and the terms around this definition. According to this contextualization it is related with the management of data in the different fields but with the same issues and concerns. The appropriate way to keep information through the time is with preservation parameters. The managing of massive amounts of critical data involves designing, modeling, processing and implementation of accurate system. The methods for the treatment of data have to consider two dimensions that this chapter has focused: access dimension and cognitive dimension. Both of them have relevance to get results because at the same time, ensure the correct data preservation. On the other side data management has to consider digital resources and digital artefacts and how they affect the process of preservation. In this chapter there are approaches with other authors referencing current and future studies. This chapter has been organized as follow. First in the introduction we can find ideas about the types of data in the different fields nowadays. Second, the main part, the proposal of methods in reference with other authors. Finally the conclusions and future projects based on the findings. This study focus on the development of framework, beginning with the methodology and projected through the application of specific patterns such a model for sustainable solution.

Keywords Digital preservation · Knowledge management · Heritage · Data management · Digital patterns

4.1 Introduction

Through the time, the information and how it can be organised and classified show the challenge in the Era of Data Management. But increasingly our possessions and communications are no longer material, now they are digital and dependent on

L.C. Gordon (✉) · Z. Chaczko
FEIT Faculty of Engineering and IT, University of Technology, NSW, Australia
e-mail: Lucia.CarrionGordon@uts.edu.au

Z. Chaczko
e-mail: Zenon.Chaczko@uts.edu.au

the technology to make them accessible. As a consequence of this big growth, to gather the requirements of data preservation involved the study and analysis inside the enterprises for constructing the profile of the customer with any kind of data. Based on that, the classification of data is important inside the Government Entities, Business Enterprises, Health Field and Cultural Area as a principal sectors.

Increasing regulatory compliance mandates are forcing enterprises to seek new approaches to managing reference data. Sometimes the approach of tracking reference data in spreadsheets and doing manual reconciliation is both timeconsuming and prone to human error. As organizations merge and businesses evolve, reference data must be continually mapped and merged as applications are linked and integrated, accuracy and consistency, realize improved data quality, strategy lets organizations adapt reference data as the business evolves.

A further motivation is to provide and improve data referencing. There are two parts: Physical (related with accessibility) and Cognitive (related with contextualization). The management of artifact have to be considered as a technological component. The main criteria is the challenge of sort out the information or data processed in smart and efficient way. The knowledge management give us the best direction of the processes of standardization. Between the principal requirements of the users are involved the authenticity and reliability. The development of the right process is independent of the type and amount of data. Nowadays the requirements of the appropriate media for the data storage should be solved easily. The needs in storage media, migration, conversion, emulation and management strategies have been focused as a consequence of the appropriate implementation of processes. This research proposed the following parts: definition of the methodology, modeling the process, implementing framework. The strategy and application of policies are important considerations at the moment of applied in case studies. Preservation responsibilities for integrating long-term preservation into planning, administration, system architectures and resource allocation are inherent in the parts of this research.

Other considerations above this topic are the applications of Digital Preservation applied in Cloud Computing architecture. Consequently referencing is a challenge because it means to identify metadata and classify them while the conventional solutions have been very limited. The investigation of scalable solution with the identification of metadata and emphasize the role of methodology and patterns should be useful to provide guidelines for protecting resources from dealing with obsolescence, responsibilities, methods of preservation, cost, and metadata formats. Based on the experience of National Library of Australia, the principal factor that we have to consider in the application of preservation is to provide frameworks related to digital heritage. It is one applications basically with the use of archives. The principal actors in this process of preservation are users and custodians. Knowledge Preservation has been applied in many areas focus on standards, scalable designing, adaptive and survivable network applications.

Furthermore, the reason for using the frameworks is the standardization of the process and the results. It does not mean the uniformity of the data, so the easy access to maintain the originality to ensure the preservation of the information. Preservation strategies and specific software tools for emulation or migration must always be

chosen according to requirements of individual institutions or users. Software tools sometimes guarantee full traceability and documentation of all elements influencing the final result. Utility Analysis and its ability to integrate inhomogeneous criteria sets is used to evaluate different strategies. The framework is often a layered structure indicating what kind of relations can or should be built and how they would interrelate.

4.2 Heritage Concepts

The heritage term is defining as the crucial and central part of the research, we can refer it to “heritage is those items and places that are valued by the community and is conserved and preserved for future generations” [2]. The concept is much wider than historical buildings. It includes items and places with natural heritage significance and Aboriginal heritage significance. “The heritage value of a place is also known as its cultural significance which means its aesthetic, historic, scientific, social or spiritual value for past, present or future generation” [2].

One of the principal keywords in this research is Heritage. UNESCO is one of the Entities that refers in an accurate way to this definition. The expert meeting denied a heritage route as “composed of tangible elements of which the cultural significance comes from exchanges and a multi-dimensional dialogue across countries or regions that illustrate the interaction of movement, along the route, in space and time” [10].

But the question is what is heritage and which parameters defining the artifact or the information as a heritage? The context and the interpretation of data is the answer.

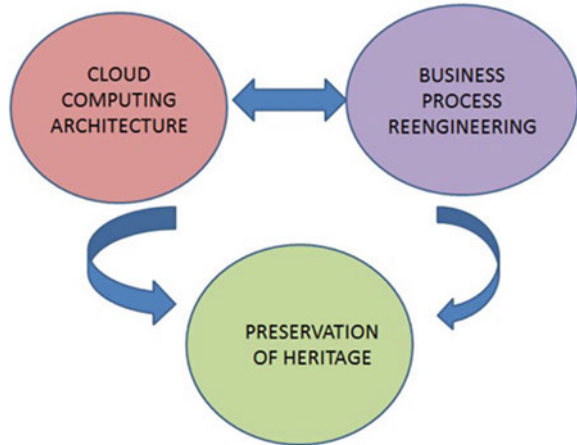
4.3 General Frameworks for Heritage

The development of the preservation framework is related with the value of information. “Value has always been the reason underlying heritage conservation. It is selfevident that no society makes an effort to conserve what it does not value” [2]. The Value of the information is located in the second level of important after the concept of Heritage. If we can define the result that we want, how we can manage and measure the value of that information? The principal ways are the perception, interpretation and contextualization.

4.3.1 *Value Derived from Individual Perceptions*

Depends on the Entity or Enterprise that information is classified, the study can consider the specific perceptions adapted to the needs of the end user. For example, there are geological, archeological, economic, social matters that could be managed by this methodology. These characteristics are considered in the case studies (Fig. 4.1).

Fig. 4.1 Areas of information management



4.3.2 Process of Preservation

The main concern around researching and definition of data preservation is the technique or methodology used. More than using technology as the solution, the principal consideration is how we can apply and verify the whole process. The definition of cycles and steps to follow has been considered as a principal part for the first consolidation of the preservation knowledge. The different perspective for the information management supports the idea to achieve general solution. The following figure shows us which could be the areas that cover the information management. There are three areas: Cloud Computing Architecture, Business Process Reengineering and Preservation of Heritage.

4.4 Architecture Vision

The formatting of data provides the unique result called Digital Age of the information. The Knowledge Management and Ontology are techniques for analyzing information. One concern of digitalization would be the format, standards and migration of the data. It should be solved with the use of Architectural Methodology and with the development of a fast prototype. This requires the definition of the sequential process.

First, should be considering a Framework as a whole front end and for reception of the information in a basic way. Using the Open Group Architectural Framework (TOGAF) [7] it is found to be more suitable as a result of dissemination of the data. Meanwhile, it is associated with an Ontology and Knowledge Management terms to be more specific.

Second, the Methodology with an architectural vision using concepts of Architectural Development Method (ADM). It has been identified in terms of enterprise description for validating information of several types of data.

Third, the conceptualization of patterns for a centralization of the preservation knowledge providing a unique result: the digital age of the data.

The connection is also with the artefacts and the correct use of them. The recovering of the information is another issue has to consider. Between the techniques we used “encapsulation, migration and emulation” [6]. The evaluation is based on “PRIST model defining by Privacy, Rights, Integrity, Security and Trust” [5] across PC considerations related with Physical and Cognitive characteristic of data.

According to the author [3] the correct use of the interface and the exact generation of metadata, are the key considerations to follow around the process of optimal data preservation (Fig. 4.2).

$$TOGAF + ADM + PRIST \tag{4.1}$$

Formula Open Group Architectural Framework (TOGAF) + Architectural Development Method (ADM) [7] + Privacy, Rights, Integrity, Security and Trust (PRIST model) [5].

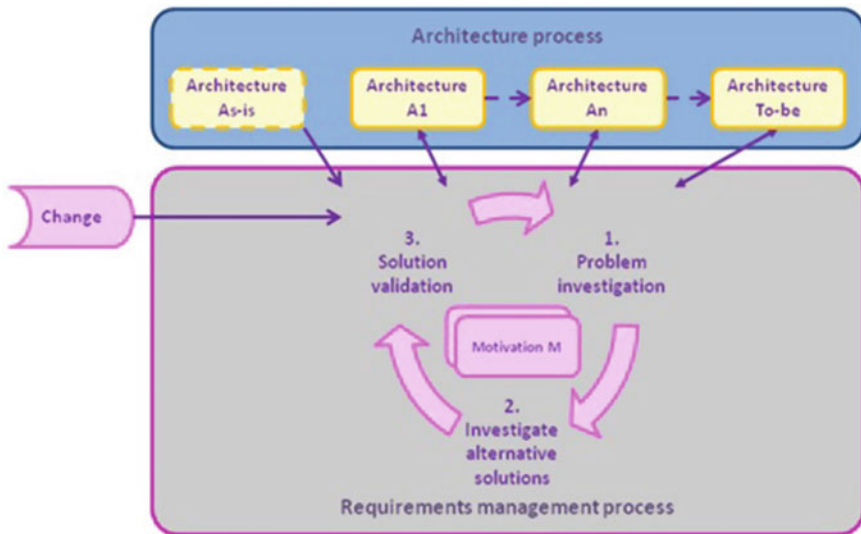


Fig. 4.2 Application of the requirements cycle [7]

4.5 Evaluation of Preservation and Planning Procedure

Nowadays Electronic publications, and a collection of multimedia art are the principal issues that need to preserve their data holdings, simulation models, or studies over time. The other matters that we have to consider is the legal constraints, to guarantee the accessibility and usability over time. Moreover the Digital Preservation is considering as a research discipline.

The two methods related in this study is: migration and emulation. There are often hold very valuable data in complex structures.

Migration is the method of repeated conversion of files or objects. Emulation denotes the duplication of the functionality of systems, be it software, hardware parts, or legacy computer systems as a whole, needed to display, access, or edit a certain document [9].

The integration of inhomogeneous criteria sets is used to evaluate different strategies. For example web archive collections, collections of scientific publications, and electronic multimedia art. The collaborative research program is funding research in different aspects of digital preservation, including collection practices, risk analyses, legal and policy issues, and technology. In Europe two new research projects. Scientific information is born-digital or only available in digital form. At the moment libraries, archive and scientific institutions are primarily dealing with the challenge of long term preservation [9].

The Reference Model for an Open Archival Information System (OAIS) was published 2002 by the Consultative Committee for Space Data Systems (CCSDS). ISO 14721:2003 defines an OAIS as “an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community” [9].

The OAIS model further provides a framework for describing and comparing different long term preservation strategies and techniques. Submission Information Package (SIP) Archival Information Package (AIP) Ingest also extracts descriptive information from the AIPs and coordinates updates to Archival Storage and Data Management Access functions include access control, request coordination, response generation in the form of Dissemination Information Packages (DIPs) and delivery of the responses to consumers. The development of preservation strategies and standards as well as packaging designs and plans (Fig. 4.3).

The facility to integrated into existing archival environments is one of the advantages of this process.

The Preservation Planning is a crucial decision process, depending on both object characteristics as well as institutional requirements. One of the methods used is the Utility Analysis approach developed at the Vienna University of Technology and the Dutch testbed designed by the National Archive of the Netherlands.

Two web archives, one coming from a library context, the other one from an archiving institution, two collections of electronic publications with scientific provenance, coming from three European national libraries and a large collection of born-digital multimedia art [9].

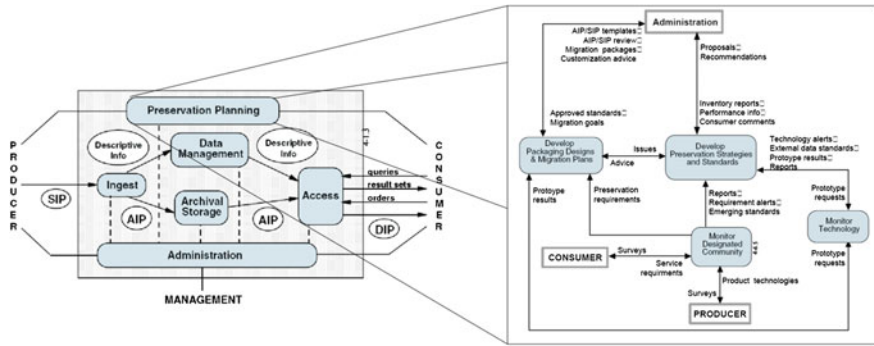


Fig. 4.3 Functional entities of the OAIS reference model [9]

The main idea of this research is related the management of heritage. It could be useful as a cultural case study related directly in terms of preservation. It exposes the experimental methodology and a valid analysis of the results. The reliability and accuracy of data are very strategic points in this article because the knowledge base is wide and complete.

On the other hand is important to differentiate the terms data, information, knowledge and knowledge management is almost as large as a number of authors contributing to the field. “Data is a set of discrete, objective facts”. “Information is data that has been organized or given structure that is, placed in context and thus endowed with meaning”. “Knowledge is information combined with experience, context, interpretation and reflection that is ready to apply to decisions and actions” [8].

4.5.1 Analysis of Capability Maturity Model (CMM)

The CMM measures and manages the improvement in software development processes. Preservation as a process with the heritage follow these steps. There are several levels:

Level 1—Initial, Level 2—Repeatable, Level 3—Defined, Level 4—Managed, Level 5—Optimizing

According to the methods and modelling, there are general approaches that give us the decision making maturity in different level as it shows in the figure below (Figs. 4.4 and 4.5).

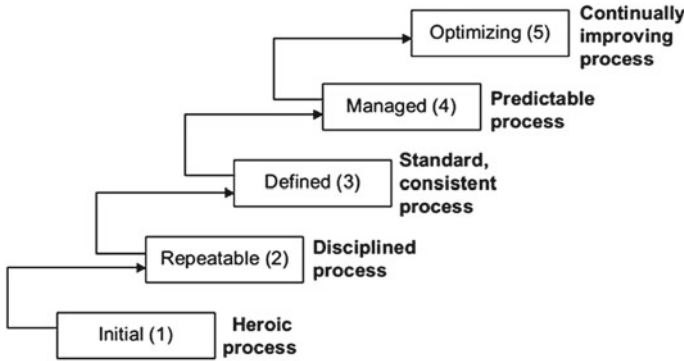
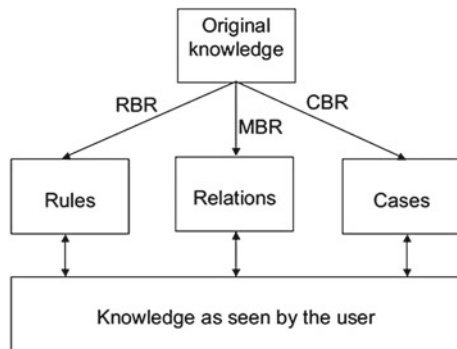


Fig. 4.4 The capability maturity model [8]

Fig. 4.5 General approaches to handling and presenting knowledge [8]



4.5.2 Heritage and Digital Data Preservation Framework

There seems to be a gap in tradition approaches to digital data preservation and planning procedures, as these methods do not integrate digital data preservation with aspects of data heritage which can be in anyway useful. In order to mitigate limitations of processes that use solely data preservation mechanisms, an integrated framework for digital data heritage and preservation is proposed. This new framework offers a consolidated and systematic approach for redesigning how digital data is managed in business enterprises. The methodology involves several process activities and adopts classical Design Patterns for digital data preservation and heritage. Among key activities are: Preparation for re-engineering, Mapping and analysing as-as processes, Designing to-be process, Implementation re-engineered process, Continuous improvement. Among many many Design Patterns were found particularly suitable for the method the following:

- Creational Patterns (Prototype and Singleton)
- Structural Patterns (Adapter, Composite and Proxy) and
- Behavioral Patterns (Visitor and State Machine)

4.5.3 Modelling Heritage and Data Preservation

The new framework’s principal components are represented in 5 + 2 dimensional model. That combines 5 dimensions of: Privacy, Rights, Integrity, Security, Trust that are denoted as PRIST [5] and two additional Cognitive (Authenticity) and Physical (Access) dimensions denoted as CP.

The new model formulas can be expressed as follows:

$$M_A + M_B = 7PC \tag{4.2}$$

$$PRIST + CP = DiADM(DIGITAL_AGE_DATA_MODEL) \tag{4.3}$$

There are three dimensions for the Preservation such as:

Cognitive Dimension: Higher-level knowledge, meaning, situation, understandability, context, interpretability, cataloguing (Fig. 4.6).

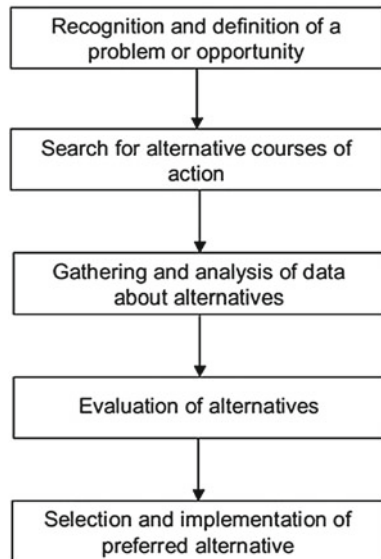
Physical Dimension: Access control, Historical Retrieval, Digital Rights Management, extraction, interface, storage.

Processes: Techniques, Approach.

When representing data in Privacy, Rights, Integrity, Security and Trust dimensions, the proposed model needs to contextualize the interpretation, conservation and preservation of the natural and cultural heritage of the real world artifacts.

Thus, the model needs to incorporate physical and cognitive dimensions to form a unified model we can be called the Digital Age of Data Model or simply DiADM.

Fig. 4.6 Rational economic model of decisionmaking [8]



During the evaluation of the approach the following question need to be addressed to ensure positive outcomes:

1. How is this material to be kept for the future?
2. Does making it accessible involve risks for the right to privacy?
3. Are we going to treat it as heritage?
4. It is possible to encoded contextualization of information to preserve originality?
5. Can all dimensions in this model keep and preserve information without damaging it?

In order to validate the proposed framework, the next step is to generate DiADM model in context, obtain concrete results and evaluate the techniques. The DiADM contextualization will depend on the case studies scenarios.

4.6 Evaluation Methods

Based on the exposed statements, the basic techniques are related with genetic and biological cases of study. The relationship between these terms is given for the behaviour and the treatment of data. There are examples referenced by known authors. The sustainability of the preservation of the information give us the discussion about the appropriate resources infrastructure.

Research by Chandras [4] explains the relationship between the terms for preservation and genetics. Nowadays Biological databases are useful for determining the interaction of biological molecules and process. The financial issues known as the important matters for evaluation of the preservation patterns is useful for testing the methodology presented. The web data concept and the use of data warehouse are the principal tendency for the management of information. The behaviour of BRC Biological Resource Center reflects the best approximation of the concepts. In the context of the studies of the maintaining Data Records [11]. “An operational climate data service must ensure that all climate data are preserved and made available to users. In addition to the climate data; metadata; production software source code; documentation on the data, metadata and data formats; calibration/validation information; and QA information will also be archived. Regular backup of data and the capability to migrate any or all of this information to new media are also important.” The challenges are specialized for deciding to manage the data sets that have to be measured. The activity to generate information and in different types and also integrate the quality for the three dimensions impact.

Based on the weighting factors use of data sets and resources decisions are very important. Research related with the evaluation of a given data set is based upon inputs from three components: the Maturity Matrix, Data Set Activity levels, and Stakeholder Input. The experience related with infrastructure is one of the possibilities to

adopt new kinds of the aggregation of the contents like a case of study of Libraries. The mature Matrix Criteria has the specific considerations given by the term Preservation. According to the author the management of the information is divided in three parts of the Data Set Evaluation Criteria: Science Maturity, Preservation Maturity and Social Impact. The other example is the Safety Net for Scientific Data. There are proposal of how the society and the researches in this case can preserve the information based on Data Archiving Polices Molecular Ecology. This technique is after the study of the basic information and how the databases have to be managed. The considerations with the DNA based microarray technology [1] are given by the treatment of the structured or not structured information.

The schema of XML and how it could be standardized the information is based on the two basic techniques: migration and emulation. This alternative could determine the reliability of the generalized framework and also affect the methodology a patterns that the researching have to establish as a result of the study. The analysis of the sequences of the different behaviours biological and genetic mixes and summarizes the terms as knowledge management and DNA. Otherwise, the patterns are related with this terminology. On the other hand using the Open Group Architectural Framework (TOGAF) [7] the research should base the central part in developing of Standardised Framework. For example Cloud Computing and Service Oriented Architecture (SOA) based on Web services technology designed to assist cultural heritage institutions in the implementation of migration based preservation interventions. SOA delivers a recommendation service and a method to carry out complex format migrations. The recommendation service is supported by three evaluation components that assess the quality of every migration intervention in terms of its performance (Migration Broker), suitability of involved formats (Format Evaluator) and data loss (Object Evaluator). The proposed system has to be able to produce preservation metadata can be used by any enterprise in the Society for documenting preservation interventions and retain objects authenticity. Although it can also be used for other purposes such as comparing file formats or evaluating the performance of conversion applications the best of this technique is Collecting Data for having only one parameter. Metadata conceptualization plays an important role in preservation of digital heritage and archives in the digital objects. The quality as one of the principal issue should be considered like a summary of good authenticity and good reliability. However, as digital objects evolve over time, their associated metadata evolves a consistency issue. Since various functionalities of applications containing digital objects (e.g., digital library, public image repository) are based on metadata, evolving metadata directly affects the quality of such applications. Modern data applications are often large scale (having millions of digital objects) and are constructed by software agents.

4.7 Conclusion

The Framework, Methodology and Architecture are the principal parts that study explains. The digital age of the data is the initiative to standardise the solution for preserve information. The consolidation of data means to have accessibility and understanding of the information during the recovering cycle. It could be the next step of the study, as the validation of the results and method to find information.

The case studies with the genetic and biological references are useful for the conclusion in behaviour of the information. The relationship between many types of data is not relevant when the study is the access of the information. The process of Collecting Data is strategic and should determine the way that it has to be managed, classified and treated.

Preservation patterns are the final step to consider to reach the standardization of the data. Across the theories, the technology should be oriented on the time that they are implemented.

Data must be securely stored and freely available to the research community depends on the context.

Long term sustainability requires adequate and reliable sources of finding data to preserve. The academic-commercial partnership may appear to have potential should corporations become involved in this collaborative effort. To prolonge financial sustainability is vital for data preservation and development of this kind of model.

References

1. Adak, S., et al.: A system for knowledge management in bioinformatics. In: Proceedings of the Eleventh International Conference on Information and Knowledge Management, (2002) pp. 638–641 Virginia, ACM
2. Bathurst, <http://www.bathurst.nsw.gov.au/building-and/heritage/what-is-heritage.html>
3. Castelli, D., et al.: Open knowledge on e-Infrastructures: the BELIEF project Digital Library. IST-Africa (2010)
4. Chandras, C., et al.: Digital preservation—financial sustainability of biological data and material resources. In: 8th IEEE International Conference on BioInformatics and BioEngineering (BIBE) (2008)
5. Challa, S., Gulrez, T., Chaczko, Z., Paranesha, T.N.: Opportunistic information fusion: a new paradigm for next generation networked sensing systems. In: 8th International Conference on Information Fusion, vol. 1, p. 8 (2005)
6. Doyle, J., et al.: Long-term digital preservation: preserving authenticity and usability of 3-D data. *Int. J. Digit. Libr.* **10**(1), 33–47 (2009)
7. Engelsman, W., Jonkers, H., Quartel, D.: ArchiMate® extension for modeling and managing motivation, principles, and requirements in TOGAF®. White Paper, The Open Group (2011)
8. Kaner, M., Karni, R.: A capability maturity model for knowledge-based decisionmaking. *Inf. Knowl. Syst. Manag.* **4**(4), 225–252 (2004)
9. Strodl, S., et al.: How to choose a digital preservation strategy: evaluating a preservation planning procedure. In: Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries. pp. 29–38. Vancouver, ACM (2007)

10. UNESCO. Information Document Glossary of World Heritage Terms. <http://whc.unesco.org/archive/gloss96.htm> (1996)
11. Weaver, R., et al.: Maintaining data records: practical decisions required for data set prioritization, preservation, and access. In:IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2008)

Chapter 5

Methods for Genealogy and Building Block Analysis in Genetic Programming

**Bogdan Burlacu, Michael Affenzeller, Stephan Winkler,
Michael Kommenda and Gabriel Kronberger**

Abstract Genetic programming gradually assembles high-level structures from low-level entities or building blocks. This chapter describes methods for investigating emergent phenomena in genetic programming by looking at a population's collective behavior. It details how these methods can be used to trace genotypic changes across lineages and genealogies. Part of the methodology, we present an algorithm for decomposing arbitrary subtrees from the population to their inherited parts, picking up the changes performed by either crossover or mutation across ancestries. This powerful tool creates new possibilities for future theoretical investigations on evolutionary algorithm behavior concerning building blocks and fitness landscape analysis.

B. Burlacu (✉) · M. Affenzeller · S. Winkler · M. Kommenda · G. Kronberger
Communications and Media, School of Informatics,
University of Applied Sciences Upper Austria, Softwarepark 11,
4232 Hagenberg, Austria
e-mail: bogdan.burlacu@fh-hagenberg.at

M. Affenzeller
e-mail: michael.affenzeller@fh-hagenberg.at

S. Winkler
e-mail: stephan.winkler@fh-hagenberg.at

M. Kommenda
e-mail: michael.kommenda@fh-hagenberg.at

G. Kronberger
e-mail: gabriel.kronberger@fh-hagenberg.at

B. Burlacu · M. Affenzeller · M. Kommenda
Institute for Formal Models and Verification,
Johannes Kepler University Linz, Altenbergerstr. 69,
4040 Linz, Austria

5.1 Introduction

5.1.1 Genetic Programming as an Emergent Process

Emergent phenomena are ubiquitous in nature. In the biological world, complexity emerges through adaptation. The circular relationship between the variation-producing operators and selection can be seen as an interplay between emergents and their basic constituents, leading to increasingly complex patterns and hierarchies.

One of the first accounts of GP as an emergent process [3] defines GP as a weak evolutionary method that employs empirical credit assignment on the solution representations to bias the search towards representations (points in the solution space) that are better suited for solving the problem. Empirical credit assignment allows the dynamics of a system to implicitly determine credit and blame; it is *holistic*, in the sense that credit is assigned to individuals and not their representational components [12]. Thus, it cannot be the representational components themselves that the algorithm manipulates but the more abstract phenotypic properties [3]. Angeline notes that it would be improper to speak of “schemas” in GP, as weak evolutionary methods have no direct correspondence between structures in the genotype and the features of the phenotype. More concretely, selection acts upon (and preserves) phenotypes. It cannot distinguish between two individuals with distinct structures but identical fitness. On the opposite end, recombination operators such as crossover or mutation act on genotypes, but are “blind” to their possible interpretation. These two complementary mechanisms determine the dynamics of GP.

Altenberg [1] contends that adaptation is an emergent property achieved by selection on heritable variation. In GP, because of the indeterminate size, structure and algorithmic properties of the tree-structured representation, both the genetics and the representation can evolve as an emergent property of the dynamics. Thus, GP has an intermediate combination of emergent and specified features. The evolvability of the population, along with the accumulation of junk code represent important aspects concerning representations and their mapping into fitness space (Fig. 5.1).

Bloat occurs under selection pressure with non-injective genotype-phenotype mappings (i.e., when any solution can be expressed in a number of alternative but semantically equivalent ways—different genotypes exhibit identical phenotypic behavior). Therefore, any serious investigation of GP emergent phenomena—bloat, evolvability, building blocks and hyperschema in particular—has to begin from the relationship between the chosen representation and the fitness landscape.

There are many views of GP as an emergent process. In his essay, Banzhaf considers “the variation-selection loop” [4] the driving force for GP emergent phenomena (Fig. 5.2). He outlines the correspondence between emergent properties and GP elements:

- The genes of the artificial genotypes (low level entities) are subject to mutation, recombination and other forces. Their variation provides bottom-up causation: genotypes are mapped to phenotypes which in turn are mapped to fitness values

- Selection is the mechanism for top-down (downward) causation, and fitness is the property selected for, which determines which combinations of genes survive.

Multiple evidence in favor of GP as an emergent process mandates the study of the origin of variation and the evolution of the genotype-phenotype map. Whether or not adaptive changes can be produced depends critically on the genotype-phenotype map [14].

This chapter introduces new methods for investigating the origins of variation through genealogy analysis. A methodology for aggregating information about subtree exchanges across genealogies is presented. The proposed methodology paves the way towards an empirical validation of schema theories or building block hypotheses for GP.

5.1.2 Genetic Programming and Building Blocks

Since its establishment [7], GP has seen much effort towards a theoretical framework able to explain its algorithmic dynamics. Holland’s schema theory [6], regarded by many as a solid mathematical basis for the canonical string-based GA, served as inspiration for similar theorems for GP centered on the concept of “hyper-schemata” (for example in [10, 11]).

Looking back at the schema theorem, we see the requirements for GA to provide good performance [2]:

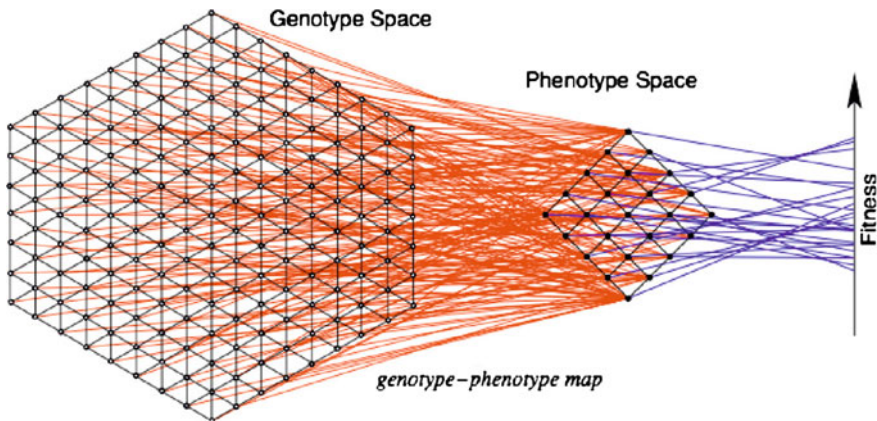
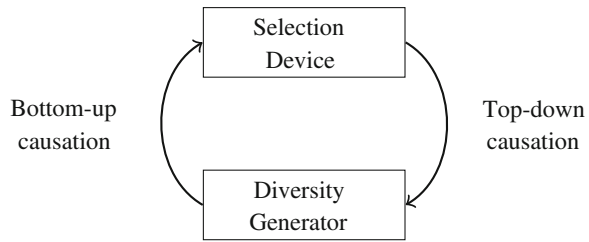


Fig. 5.1 Genotype-phenotype map. Source [13]

Fig. 5.2 Variation—selection loop in GP as an engine for emergence. Source [4]



- The recombination operator determines which schemata are being recombined
- There needs to be a correlation between complementary schemata of high fitness and the fitness distributions of their recombinant offspring in order for the GA to increase the chance of sampling fitter individuals

Hyperschema theories for GP assume the existence of variable-arity schema, such as those defined by [11], in which special symbols are used to designate wildcard nodes and subtrees. Under this assumption, we can consider specific hyperschema instances as “building blocks”, subtrees that contain symbols and operators that are parts of the high quality solutions.

The analysis of genetic inheritance within GP represents a promising starting point for the investigation of hyperschema theories and other important aspects of population dynamics. An idea for identifying important building blocks in the population could be to simply count their occurrence frequency or the impact they have on the fitness of good individuals in the population.

The following sections describe the implementation of the genealogy tracking in HeuristicLab [15] and the specific methods for decomposing and tracing symbolic expressions across individual lineages.

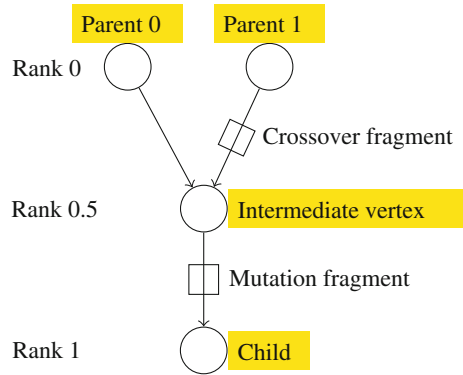
5.2 Genetic Programming Genealogies

Using GP genealogy information, exact methods for analyzing algorithm dynamics can be developed. Tracing gene propagation through the generations and aggregating results concerning size, fitness and population diversity can provide a strong empirical basis for GP behavioral theories. Access to the exact evolutionary trajectories could potentially validate or invalidate existing GP schema theories [9–11] or building block hypotheses [8].

5.2.1 Building the Genealogy Graph

The HeuristicLab implementation stores genetic information as a directed graph in which vertices represent GP individuals and arcs represent their hereditary relation-

Fig. 5.3 “Intermediate” vertex with rank 0.5 added to the genealogy graph in order to save the intermediate genetic exchange between parents (rank 0) before the mutation



ships. Additionally, each arc stores data about the subtrees that were transferred from parents to children (Fig. 5.3). A detailed description of the genealogy graph implementation is given in [5].

The data stored in each arc (going from parent to child) represents genetic information fragments that are passed on from parents to offspring. A *fragment* contains a reference to the swapped subtree (the fragment *Root*), along with two position indices, relative to the preorder enumeration of the tree’s nodes:

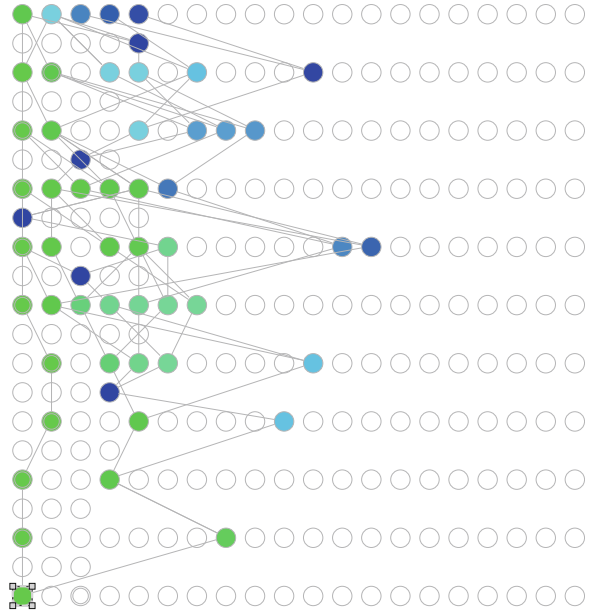
- The first index represents the position of the swapped subtree in the first parent
- The second index represents the position of the swapped subtree in the second parent

Conceptually, a genotypic fragment represents the part that replaces an element of a bigger inherited genetic context. For example, in the case of crossover, the fragment is the subtree from the non-root parent that replaces a subtree in the root-parent to produce the offspring (Fig. 5.5). For mutation, the fragment is considered to be the part that changed, the genotypic difference between parent and child.

The genealogy graph is extended each generation with new vertices representing the offspring that make up the new generation of individuals. In HeuristicLab, each genetic operator can be instrumented so that additional processing can be performed before and after their execution. Through this mechanism, we are able to accurately identify genetic fragments based on the structural comparison between the parent (root parent in the case of crossover) and the child.

Vertices corresponding to each individual are assigned a rank based on the current generation number, so that vertices corresponding to individuals from the initial population have rank 0, their offspring have rank 1, and so on. Additionally, when crossover and mutation act on the same individual consecutively, an “intermediate” vertex is added to the genealogy graph, so that the order in which the genetic information is inherited is preserved.

Fig. 5.4 Sample GP genealogy. Vertices represent individuals and arcs hereditary relationships. *Top-down* successive generations, *left-to-right* individuals color-coded and ordered by decreasing fitness (better fitness → warmer color)



5.2.2 Tracing Genetic Fragments and Building the Fragment Graph

Under the building block hypothesis, we accept that GP solutions are assembled from increasingly complex parts that emerge gradually as by-products of the selection-variation loop. We assume that the genetic fragments we want to trace will contain multiple smaller parts which in themselves are decomposable to constitutive elements as we move backwards through genealogies.

The tracing methodology basically has to solve the following problem: given a whole solution or a part of the solution, can it be decomposed to its constitutive elements from the previous generations? Based on the genealogy graph (Fig. 5.4), any subtree in the population can be traced down to its composing elements. The result is another graph structure called a *fragment graph* which describes the relationship between subtrees and their inherited elements.

Figure 5.5 shows a general picture of how the tracing works for a crossover operation in between two parents. The first parent, called the “root parent” contributes its whole genotype to the creation of the offspring. The second “non-root” parent contributes a part of its genotype which is swapped in place of a randomly selected subtree from the root parent. Thus, what the child inherits is a larger genetic context from the root parent and a genetic fragment from the non-root parent. The number below each node represents its index in the preorder enumeration of the tree nodes.

Preorder traversal of symbolic expression trees has the advantage of allowing subtree inclusion checks via simple arithmetic operations. The general condition for

a subtree A to include a subtree B is: $i_A < i_B < i_A + l_A$, where i_A, i_B are the preorder indices of the given subtrees and l_A, l_B are their lengths. If this condition is not met, then the two subtrees are distinct.

In Fig. 5.5, we see that the fragment received by the child has an index of 4, while in the non-root parent it had an index of 1. Let the whole child be the subtree that we want to trace. Then, its root will have index 0, and it will include the fragment in its structure.

In this situation, the tracing algorithm will visit the root parent and trace the subtree from index 0 (same as in the child), and it will visit the non-root parent and trace the subtree from index 1 (the index of the fragment). Therefore the Fragment Graph will start from the vertex corresponding to the child and branch in two directions corresponding to the two parents, and so on until the beginning of the ancestry is reached.

We describe the subtree-tracing algorithm below. As mentioned before, genetic fragments are identified by comparing the nodes of the parent (in preorder) with the nodes of the child. The indices corresponding to each fragment indicate its position in the child (at the same position we find the replaced subtree in the root-parent), and in the non-root parent.

The problem of subtree tracing basically reduces itself to the problem where a given subtree contains a genetic fragment. We start from the general case, in which we are given a subtree, and a fragment, both belong to the same tree individual. If we denote the fragment index by f_i , the fragment length by f_l , the subtree index by s_i and the subtree length by s_l , we can distinguish four cases to which the tracing problem can be reduced, each described in Table 5.1.

In the four cases above, special attention must be paid to the correct update of indices when the tracing procedure traverses the genealogy graph. In the implementation, an additional index is considered for the fragment: its index in the non-root parent from which it originates. The trace function iteratively walks the genealogy graph from child to parent in a loop from which it breaks when it reaches the beginning of the ancestry (when there are no more parents). The iteration is controlled by a state composed of two variables:

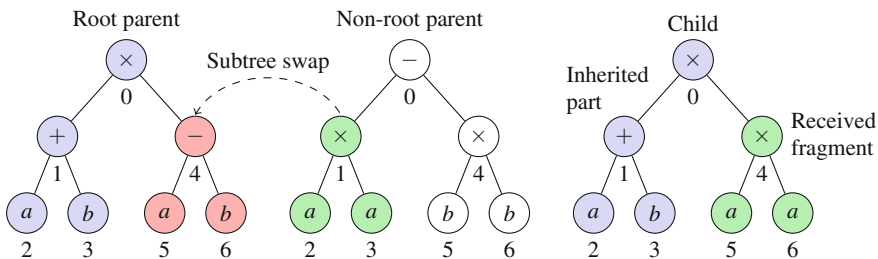


Fig. 5.5 Preorder arithmetics for subtree inclusion

Table 5.1 Sub-cases of the tracing problem

	Case	Description	Action
1.	$f_i = s_i$	Fragment and subtree coincide	Tracing is restarted from the non-root parent which originally contained the fragment
2.	$s_i + s_l < f_i$ or $f_i + f_l < s_i$	Fragment and subtree are distinct	Tracing is restarted from the root parent which contains the subtree
3.	$f_i < s_i < f_i + f_l$	Subtree contained by fragment	In this case, again, tracing is restarted from the non-root parent
4.	$s_i < f_i < s_i + s_l$	Subtree contains fragment	In this case, both the fragment and the original subtree (before the fragment was swapped) must be traced. This last case—the most general and useful case—represents a bifurcation point in the tracing procedure. The procedure recurses on each path (root and non-root parent)

We consider a subtree to be traced and a fragment inherited from the non-root parent. We use the notation s_i , s_l , f_i , f_l for the subtree and fragment preorder index and length, respectively

- The current subtree index: Marks the position of the subtree to be traced in its containing individual (according to the preorder node enumeration)
- The current individual in the genealogy. When the current individual has no parents, then we break the loop as we reached the beginning of the ancestry.

It is worth noting that for each of the cases 1–3 described above, the tracing method does not add vertices in the fragment graph, but instead—to produce a more compact representation—it skips individuals that remain unchanged relative to the subtree that is being traced. This explains why in Fig. 5.6, connected vertices are not required to be of consecutive ranks.

To save space, the pseudocode in Algorithm 4 only describes subtree tracing for crossover operations. A similar approach can be used for mutation, where the procedure is simpler as the individual only has one parent. It is obvious that in the case of mutation we cannot speak of proper inheritance but rather, we are trying to trace the original subtree before mutation, in an effort to find out if good mutations depend on favorable conditions which must have been created (evolved) in the tree.

In the code, the condition at line 10 corresponds to Case 1 in Table 5.1, the ones at lines 17 and 26 to Case 2, the one at line 14 to Case 3 and the one at line 22 to Case 4. Most of the time we are interested in tracing whole individuals, which means that the algorithm will often fall back to Case 4. Individuals can be successively decomposed

Algorithm 4: Method **Trace** for decomposing a subtree to its inherited parts**Input:** tree \leftarrow the starting point in the genealogy, subtree \leftarrow the subtree to be traced**Output:** A graph of all the inherited elements that formed the subtree

```

1 begin
2   currentTree  $\leftarrow$  tree
3    $s_i \leftarrow$  subtree index in currentTree
4   while node has parents do
5     parents  $\leftarrow$  the nodes parents
6     fragment  $\leftarrow$  inherited genetic fragment
7      $f_1, f_2 \leftarrow$  fragment indices in first and second parent
8      $s_i \leftarrow$  subtree index in tree
9      $f_l, s_l \leftarrow$  fragment and subtree lengths
10    if  $f_1 = s_i$  then
11      currentTree  $\leftarrow$  non-root parent
12       $s_i \leftarrow f_2$ 
13    else if  $f_1 < s_i$  then
14      if  $f_1 + f_l > s_i$  then
15        currentTree  $\leftarrow$  non-root parent
16         $s_i \leftarrow s_i - f_l + f_2$ 
17      else
18        currentTree  $\leftarrow$  root parent
19        len  $\leftarrow$  length of the subtree at index  $f_1$  in the root parent
20         $s_i \leftarrow s_i + \text{len} - f_l$ 
21    else if  $f_1 > s_i$  then
22      if  $f_1 < s_i + s_l$  then
23        add current node to fragment graph
24        Trace(root parent, subtree at index  $s_i$  in root parent)
25        Trace(non-root parent, subtree at index  $f_i$  in the non-root parent)
26      else
27        currentTree  $\leftarrow$  root parent

```

to their inherited parts (marked with black contours in each tree) until the beginning of the ancestry is reached. The characteristics of the resulting Fragment Graph can give an impression about how the solution was formed and how the initial genetic diversity was exploited by the algorithm.

Figures 5.6 and 5.7 show the best individual of the population being traced down to its inherited elements. For space reasons, only it and its parents are displayed. The color convention for the displayed individuals makes it easy to identify inherited fragments and crossover points:

- The subtree marked with gray represents the fragment received from the non-parent
- The node marked with red represents the subtree in the root parent that was replaced by the fragment
- The subtree marked with black contour nodes (with black text as opposed to gray) represents the subtree that is being traced

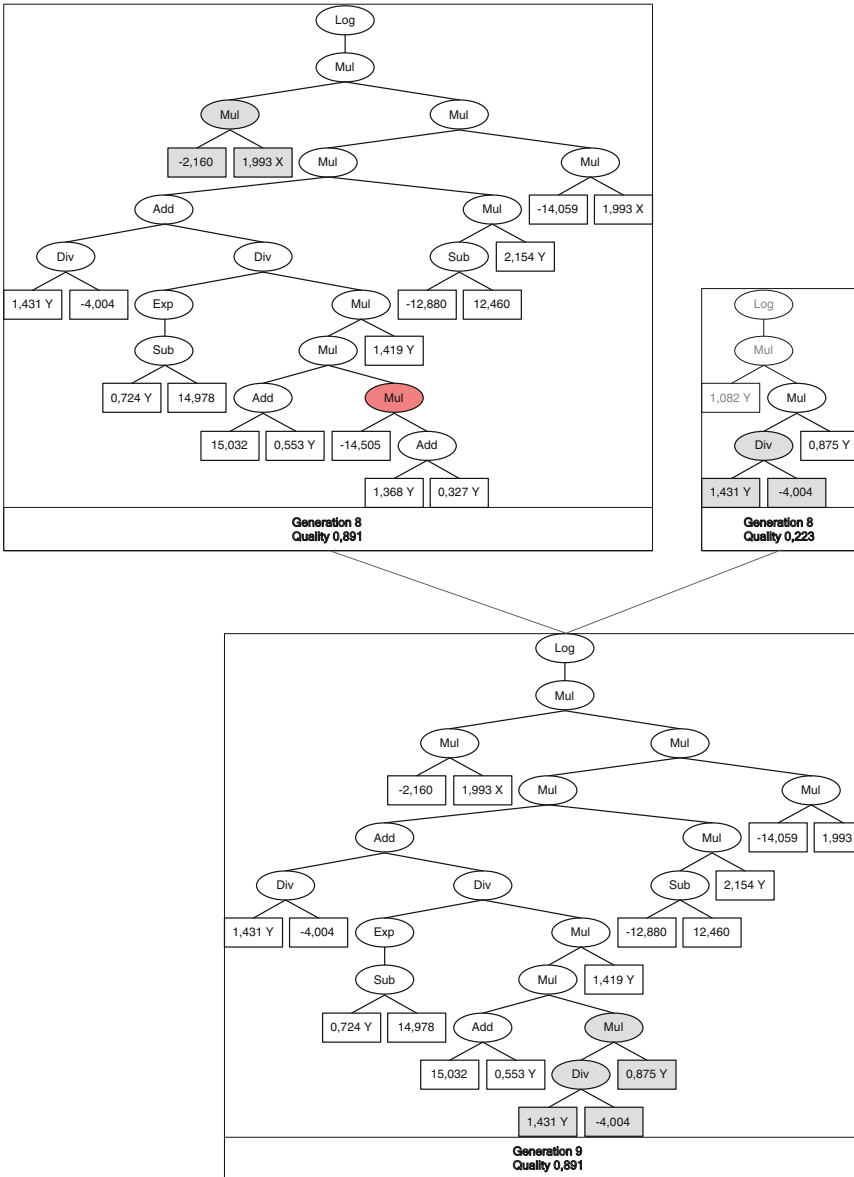


Fig. 5.6 Partial trace of the best individual of a GP run

Using this convention, we see in Fig. 5.6 an individual with two parents from which it inherited its genetic material. The smaller fragment from the non-root parent (marked with black contours) replaced the subtree marked with red in the other

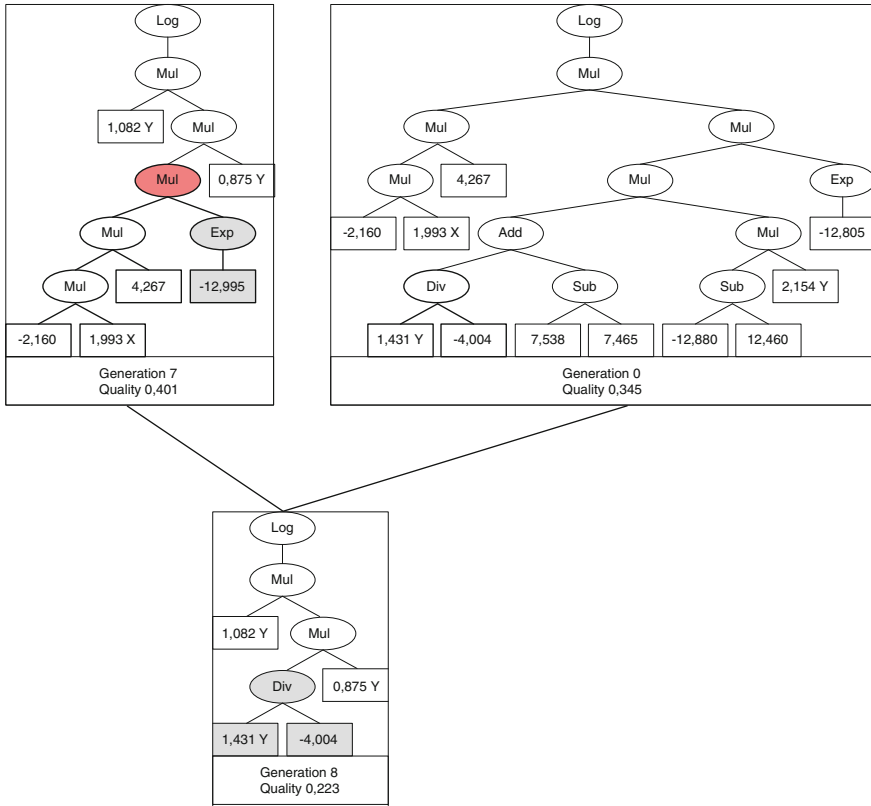


Fig. 5.7 Partial trace continuing from the non-root parent in Fig. 5.6

parent, producing the offspring (in which the fragment is marked with gray). Now, we continue the trace going up the genealogy to the non-root parent.

Figure 5.7 shows how the fragment from Fig. 5.6 was formed from elements belonging to an individual of Generation 7 and one of the initial population (Generation 0). This shows that in most cases the fragment graph turns out to be a simplified version of the full genealogy recorded in the Genealogy Graph.

Figure 5.8 shows how the fragment graph looks like when tracking the best solution of genetic programming runs. The three pictures displayed side-by-side show the implementation’s ability to zoom in and out of the fragment graph, showing the big picture or focusing on specific parts or individuals. Each red rectangle marks the part of the picture that was zoomed in and displayed in more detail in the next picture.

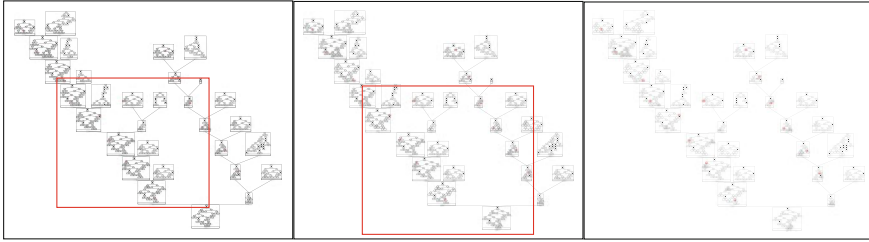


Fig. 5.8 HeuristicLab fragment graph View: birds-eye view of the whole graph (*left*), medium zoom level (*middle*), detailed zoom level (*right*). The *red rectangles* mark the areas where the fragment graph was zoomed in (from *left* to *right*)

5.3 Conclusions

5.3.1 Overview

In this contribution, we argued that GP is an intrinsically emergent process, in which genotypic representations and phenotypic properties are intercorrelated through the effects of the variation-selection loop. We claimed that emergent properties such as bottom-up and top-down causation and auto-organization are responsible for the occurrence of hierarchies in GP, of bloat, and ultimately, for the formation of solutions from smaller lower-level entities called “emergents”.

Next, we presented an approach for the investigation of the genotype-phenotype map, based on the recording and analysis of genealogical information during the run of the algorithm. If the transformations of the initial distribution of genetic materials in the initial population can be identified, then each single genotype–phenotype mapping can be explained in relation to what changed and where the change originated from.

We used the term “fragment” to describe an inherited piece of genetic information generally defined as the part of the genotype that is changed from parent to child. The problem we were interested in solving was to decompose an arbitrarily defined subtree into its constitutive parts. We demonstrated a new methodology for ‘tracing’ fragments, which allows us to follow an individual’s ancestry in the genealogy graph, producing a detailed image of how the genetic information was propagated across the ancestry.

All the analysis methods described in this paper were implemented and tested within HeuristicLab, an Evolutionary Computation framework developed at the Heuristic and Evolutionary Algorithms Laboratory in Hagenberg, Upper Austria. The genealogy analysis methods scale well with the number of generations and number of individuals in the population. They can be applied to optimization problems of realistic dimensions where the fragment graph can give a real clue about the algorithm’s emergent properties. Quantifiable indicators about fragment and population size distributions, genotypic diversity, effectiveness of genotypic operators (in terms of fitness changes) can be extracted from the produced graphs in order

to empirically determine which mathematical laws can accurately characterize the selection-variation loop.

5.3.2 Future Research

The next steps of this research are to extend the analysis methods for the newly developed fragment graph, and obtain empirical evidence concerning the behavioral dynamics of GP populations. The following topics present special interest for the further development of our methodology:

- Automated detection of building blocks based on their relevance (their quantifiable impact on the output of a given solution tree) and their frequency in the population
- Observe and analyze the dynamics of building block distribution in the population, using the tracing methodology
- Integrate a hyperschema representation into the methodology in order to investigate GP schema theories
- Develop improved analysis methods (concerning population diversity and genetic operator effects) exploiting the exact genotypic differences between populations from successive generations.

References

1. Altenberg, L.: Emergent phenomena in genetic programming. In: Sebald, A.V., Fogel, L. J. (eds.) *Evolutionary Programming—Proceedings of the Third Annual Conference*, pp. 233–241, pp. 24–26. World Scientific Publishing, San Diego (1994)
2. Altenberg, L.: The Schema Theorem and Price’s Theorem. *Foundations of Genetic Algorithms*, pp. 23–49. Morgan Kaufmann (1995)
3. Angeline, P.J.: Genetic programming and emergent intelligences. In: Kinnear Jr, K.E. (ed.) *Advances in Genetic Programming*, pp. 75–98. MIT Press, Cambridge (1994)
4. Banzhaf, W.: Genetic programming and emergence. *Genet. Program. Evol. Mach.* **15**(1), 63–73 (2014)
5. Burlacu, B., Affenzeller, M., Kommenda, M., Winkler, S., Kronberger, G.: Visualization of genetic lineages and inheritance information in genetic programming. In: *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation, GECCO’13 Companion*, pp. 1351–1358, ACM (2013)
6. Holland, J.H.: *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Michigan (1975)
7. Koza, J.R.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge (1992)
8. O’Reilly, U.-M., Oppacher F.: The troubling aspects of a building block hypothesis for genetic programming. In: Whitley, L.D., Vose, M.D. (eds.), *Foundations of Genetic Algorithms 3*, pp. 73–88, Estes Park, Colorado, 1994. Morgan Kaufmann. Published (1995)
9. Poli, R., Langdon, W.B., Dignum, S.: On the limiting distribution of program sizes in tree-based genetic programming. Technical Report CSM-464, Department of Computer Science, University of Essex (2006)
10. Poli, R., McPhee, N.F.: General schema theory for genetic programming with subtree-swapping crossover: Part I. *Evol. Comput.* **11**(1), 53–66 (2003)

11. Poli, R., McPhee, N.F.: General schema theory for genetic programming with subtree-swapping crossover: Part II. *Evol. Comput.* **11**(2), 169–206 (2003)
12. Spector, L., Langdon, W.B., O'Reilly, U.-M., Angeline, P.J. (eds.): *Advances in Genetic Programming*. MIT Press, Cambridge (1999)
13. Stadler, P.F.: Genotype-phenotype maps. *BIOLOGICAL THEORY*, 2, 2006
14. Wagner, G.P., Altenberg, L.: Complex adaptations and the evolution of evolvability. *Evolution* **50**(3), 967–976 (1996)
15. Wagner, S., Affenzeller, M.: *The heuristiclab optimization environment*. Technical report, Johannes Kepler University Linz, Austria, 2004

Chapter 6

Multi-Population Genetic Programming with Data Migration for Symbolic Regression

Michael Kommenda, Michael Affenzeller, Gabriel Kronberger, Bogdan Burlacu and Stephan Winkler

Abstract In this contribution we study the effects of multi-population genetic programming for symbolic regression problems. In addition to the parallel evolution of several subpopulations according to an island model with unidirectional ring migration, the data partitions, on which the individuals are evolved, differ for every island and are adapted during algorithm execution. These modifications are intended to increase the generalization capabilities of the solutions and to maintain the genetic diversity. The effects of multiple populations as well as the used data migration strategy are compared to standard genetic programming algorithms on several symbolic regression benchmark problems.

M. Kommenda (✉) · M. Affenzeller · G. Kronberger · B. Burlacu · S. Winkler
School of Informatics, Communications and Media,
University of Applied Sciences Upper Austria, Softwarepark 11,
4232 Hagenberg, Austria
e-mail: michael.kommenda@fh-hagenberg.at

M. Kommenda · M. Affenzeller · B. Burlacu
Institute for Formal Models and Verification,
Johannes Kepler University Linz, Altenbergerstr. 69,
4040 Linz, Austria

M. Affenzeller
e-mail: michael.affenzeller@fh-hagenberg.at

G. Kronberger
e-mail: gabriel.kronberger@fh-hagenberg.at

B. Burlacu
e-mail: bogdan.burlacu@fh-hagenberg.at

S. Winkler
e-mail: stephan.winkler@fh-hagenberg.at

6.1 Introduction

In the beginning of the 1990s coarse-grained parallel and distributed genetic algorithms [16, 22] were developed to speed up the evolutionary process by evolving subpopulations in parallel on different machines. If the individual subpopulations do not exchange any individuals at all, this approach is equivalent to running multiple genetic algorithms with varying initial conditions (e.g., the seed of the random number generator) sequentially. However, the more common approach is to exchange information between different subpopulations during the migration phase, when individuals move from one subpopulation to another at specific intervals. This information exchange is performed according to a predefined topology, which generally stays constant during the algorithm execution. Common examples for such topologies are uni- or bidirectional rings, grids, or hyper cubes [2].

The evolution of individuals on different subpopulation combined with an exchange of genotype information maintains the overall genetic diversity [17], which is considered a critical success factor of evolutionary algorithms. Another benefit of using multi-population evolutionary algorithms is that parts of the solution for separable problems can be evolved in different subpopulations and assembled to the final solution by the exchange of the solution parts through migration [21].

Although the behavior of island genetic algorithms has been extensively studied, the literature on multi-population genetic programming and especially symbolic regression is not as comprehensive. In *A Field Guide to Genetic Programming* [11] a detailed overview of parallelization techniques for genetic programming by using multiple populations, which focuses on faster algorithm execution, is given.

Symbolic regression is the task of finding a mathematical formula modeling the relation between a dependent variable and several independent ones. Symbolic regression is commonly solved by genetic programming [9], because of its ability to evolve models without any assumptions about the generated model structure and its strong, implicit feature selection [15]. Genetic programming is regarded as an evolutionary, population-based, heuristic algorithm that works similar to genetic algorithms, with the distinction that it uses a variable-length encoding for individuals. Therefore, the algorithmic advancements regarding multi-population evolutionary algorithms are easily transferred to genetic programming.

One of the first parallel genetic programming systems for solving symbolic regression problems is described in [12] and more detailed analysis regarding diversity, migration schemes, and population size were conducted in [3]. Although these studies reveal interesting theoretical aspects, a point of criticism is that the used symbolic regression benchmarks are nowadays solved by almost every genetic programming variant, because of the increased algorithm performance due to advanced operators and optimization techniques, such as linear scaling [7], offspring selection [1], semantic crossovers [18], or constant optimization [8]. For this reason, we tested the capabilities of multi-population genetic programming on recently published benchmark problems [20] and furthermore, included a novel sample selection strategy into multi-population genetic programming for symbolic regression problems.

6.2 Methods

Several techniques for sample selection in genetic programming were developed over the years, which all have in common that only a subset of the available data is used for evaluating the fitness of an individual. Hence, the algorithm runtime is reduced and due to regular changes to the selected subset the generalization capabilities of the evolved models should be increased. The techniques for sample selection range from random selection [4, 5], co-evolution of models and the samples used for fitness calculation [6, 13] or concepts derived from ordinal optimization [14]. Other methods are based on the difficulty of the samples or on the time since its last use in the evaluation [11].

We implemented a multi-population genetic programming algorithm for symbolic regression (illustrated in Fig. 6.1), where the samples for fitness calculation of each subpopulation differ and are exchanged at specific intervals. This mechanism should increase the generalization capabilities of the individuals and maintain the genetic diversity, which both should yield a better performance of the algorithm. Another benefit is that the execution time of the algorithm is reduced, because the individuals are only evaluated on parts of the training data and evaluation is in general the most time consuming part of the algorithm, which scales linearly with the size of the training data.

The assignment of the training data is performed during the initialization of the algorithm, where a different fixed subset of the training data, denoted as *FixedTrainingSubset* (FTS), is assigned to each island. If no migration is performed, the modified algorithm behaves exactly as concurrent standard genetic programming algorithm

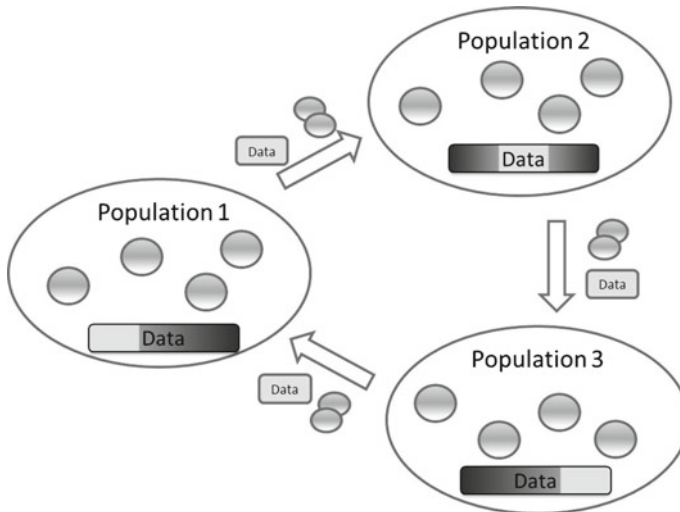


Fig. 6.1 Schematic representation of multi-population genetic programming with data migration

executions that are applied on different parts of the data. If migration is performed the immigrants from another subpopulation have to be re-evaluated on the FTS of the new subpopulations to adjust their quality accordingly, so that the selection is not misled by outdated quality information.

We use an unidirectional ring structure as migration topology, where the emigrants are moved to the clockwise neighboring subpopulation. As a result of choosing exactly the same number of emigrants and immigrants the population size of every island stays constant during the algorithm execution. A drawback of using varying training data partitions for fitness calculation per subpopulation is that individuals, which are never migrated, are learned only on their respective FTS and hence not all available training data is utilized.

Therefore, the FTS is joined with consecutively chosen samples, denoted as *VariableTrainingSubset* (VTS) that are not already included in the fixed partition. This additional data is migrated at regular intervals, which per default is equal to the individuals migration interval. As a result every individual is evaluated on the FTS of its island and on the changing VTS and over time all available data is presented to each individual.

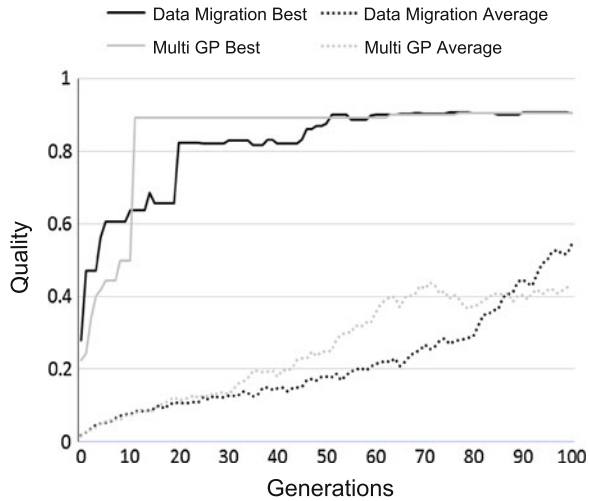
Algorithm 5: Multi-population genetic programming with data migration.

```

1 begin
2   for each Subpopulation do
3     Assign fixed training subset FTS
4     Assign variable training subset VTS
5     Generate initial population
6     Evaluate individuals on  $FTS \cup VTS$ 
7    $generations \leftarrow 0$ 
8   while  $generations < maxGenerations$  do
9     for each subpopulation do
10      Select parent individuals for reproduction
11      Generate offspring from selected parent individuals
12      Evaluate offspring on  $FTS \cup VTS$ 
13      Replace parents with newly generated offspring
14      if Migrate individuals? then
15        Comment: Data changes for migrated individuals
16        Migrate individuals to the neighboring subpopulation
17        Reevaluate migrated individuals on new data
18      if Migrate data? then
19        Comment: Data changes for all individuals
20        Migrate VTS to the neighboring subpopulation
21        Reevaluate all individuals
22       $generations \leftarrow generations + 1$ 
23    Reevaluate individuals on whole training data
24  return Best individual

```

Fig. 6.2 Best and average individual quality of multi-population GP with data migration (DataMigration) and standard multi-population GP (Multi GP) of exemplary algorithm runs on the Paige-1 problem



As a consequence of evaluating individuals only on a subset of the data, the individual with the highest fitness does not have to be the individual with the best fit on the training data. Thus, before selecting an individual of the final generation as the algorithm results, all individuals are reevaluated on the whole training data. The algorithm result is then the model with the best fit on whole training data and models that by chance explain their current training subset well, but could not explain the whole training data, are detected and ruled out. The whole multi-population genetic programming algorithm with data migration is stated as pseudo code in Algorithm 5.

The effects of the algorithm adaptations are shown in Fig. 6.2, which illustrates the quality of the best individual and the average quality of all individuals per generation for multi-population genetic programming (Multi GP) and multi-population GP with data migration (DataMigration). The shown results are the algorithm executions resulting in the best solution found for the Paige-1 problem (see Table 6.1). The multi-population GP algorithm uses all training data, whereas the algorithm with data migration uses 20% of the training data as FTS, as well as 20% as the VTS. In contrast to the standard algorithm, which finds the best model within the first 20 generations, the data migration algorithm takes longer to learn the characteristics of the data. The reason for this is that the data is presented to the individuals only gradually. In the end the best solutions are of approximately equal quality, but the average quality of the models is higher when using data migration, although that the algorithm uses only 40% of the training data.

6.3 Experiments

The effectiveness of the different algorithm variants and especially multi-population genetic programming with data migration were tested on a selection of artificial benchmark problems (Table 6.1), which were published as recommended symbolic

Table 6.1 Definition of synthetic benchmark problems and the respective number of training and test samples

<i>Nguyen-12</i>	$f(x, y) = x^4 - x^3 + 0.5y^2 - y$
Training	200 samples
Test	1,000 samples
<i>Keijzer-14</i>	$f(x, y) = \frac{8}{2+x^2+y^2}$
Training	200 samples
Test	798 samples
<i>Vladislavleva-5</i>	$f(x_1, \dots, x_3) = 30 \frac{(x_1-1)(x_3-1)}{x_2^2(x_1-10)}$
Training	300 samples
Test	2,700 samples
<i>Vladislavleva-7</i>	$f(x_1, x_2) = (x_1 - 3)(x_2 - 3) + 2\sin((x_1 - 4)(x_2 - 4))$
Training	200 samples
Test	1,000 samples
<i>Poly-10</i>	$f(x_1, \dots, x_{10}) = x_1x_2 + x_3x_4 + x_5x_6 + x_1x_7x_9 + x_3x_6x_{10}$
Training	200 samples
Test	1,000 samples
<i>Pagie-1</i>	$f(x, y) = \frac{1}{1+x^{-4}} + \frac{1}{1+y^{-4}}$
Training	200 samples
Test	1,000 samples

regression problems [20]. Two configurations of the benchmark problems were tested; one without any noise and another one where a normally distributed noise term was added to the generating data function, that contributes 20% of the total variance. Hence, the maximum coefficient of correlation R^2 that can be achieved on the noisy problems is 0.8.

6.3.1 Algorithm Configurations

In the first experiment three multi-population algorithm with different migration strategies were tested and compared to standard genetic programming with random sample selection (GP + RSS). One variant uses individual migration and data migration (IM + DM), one only data migration (DM), and another one only individual migration (IM). The algorithm with data migration use 20% of the training data as FTS and 20% as VTS, whereas the one with only individual migration uses 40% as FTS and no VTS. The standard genetic programming algorithm randomly selects 40% of the training data for fitness calculation. Hence, the number of used samples for fitness calculation is the same for all algorithm variants. The second experiment compares data migration algorithms (with 20% FTS and 20% VTS) with single and

Table 6.2 Algorithmic settings common for all algorithm variants

Tree initialization	PTC2 [10]
Maximum tree length	100 Nodes
Population size	500 Individuals/100 Individuals divided on 5 subpopulations
Elites	1 Individual
Selection	Tournament selection Group size 2/4
Crossover probability	100 %
Crossover operator	Subtree crossover
Mutation probability	25 %
Mutation operators	Single point mutation Remove branch Replace branch
Fitness function	Coefficient of determination R^2
Termination criterion	100 Generations
Terminal symbols	<i>Constant, weight * variable</i>
Function symbols	Binary functions(+, -, ×, /, exp, log)
<i>Further specific algorithm settings are described textually</i>	

multi-population genetic programming, that use all available samples used for fitness calculation of the individuals.

All multi-population algorithm configurations for both experiments use 5 subpopulations, where every subpopulation contains 100 individuals, and an tournament group size of 2 to adapt the selection pressure to the smaller population size. Single population algorithms use a population size of 500 and a tournament group size of 4 (see Table 6.2). Therefore, the number of evaluated individuals for all algorithm configurations is the same. The data migration interval was set to 10 generations and every fifth generation 10 % of the individuals were migrated according to an unidirectional ring topology. The fitness function used was the coefficient of determination R^2 [7] and as the R^2 is invariant to linear transformations, the individuals have to be scaled linearly before other quality measures can be calculated. All the algorithms were implemented and the experiments performed in HeuristicLab¹ [19], an open source framework for heuristic optimization.

6.4 Results

All results presented in this sections are reported as the coefficient of determination R^2 of the best model on the whole training data (even if sample selection strategies

¹ <http://dev.heuristiclab.com>.

Table 6.3 Median, average and standard deviation (first–third row) of the qualities of multi-population GP with individual and data migration (IM + DM), either one of them, and of GP with random sample selection (GP + RSS) on all problems without any noise

	IM + DM		DM		IM		GP + RSS	
	Training	Test	Training	Test	Training	Test	Training	Test
Nguyen-12	0.978	0.981	0.979	0.980	0.980	0.981	0.976	0.979
	0.979	0.981	0.980	0.967	0.980	0.981	0.975	0.955
	0.005	0.004	0.006	0.107	0.006	0.006	0.030	0.131
Keijzer-14	0.994	0.992	0.990	0.988	0.992	0.990	0.994	0.991
	0.992	0.990	0.990	0.988	0.991	0.989	0.991	0.986
	0.005	0.007	0.006	0.007	0.006	0.007	0.010	0.015
Vladislavleva-5	0.861	0.638	0.857	0.630	0.872	0.621	0.868	0.633
	0.852	0.627	0.865	0.655	0.854	0.610	0.777	0.574
	0.091	0.154	0.065	0.167	0.092	0.177	0.208	0.231
Vladislavleva-7	0.830	0.723	0.819	0.762	0.833	0.739	0.720	0.574
	0.794	0.669	0.803	0.740	0.807	0.683	0.721	0.602
	0.075	0.211	0.055	0.158	0.066	0.199	0.112	0.146
Poly-10	0.571	0.524	0.524	0.287	0.385	0.344	0.321	0.257
	0.549	0.510	0.515	0.318	0.454	0.323	0.391	0.308
	0.189	0.237	0.169	0.151	0.156	0.139	0.160	0.152
Pagie-1	0.903	0.839	0.900	0.837	0.903	0.836	0.903	0.834
	0.907	0.844	0.903	0.829	0.903	0.836	0.906	0.783
	0.017	0.033	0.013	0.075	0.006	0.015	0.018	0.169

are used) and are aggregated over 50 independent repetitions for every algorithm to take their stochasticity into account.

Tables 6.3 and 6.4 report the median, mean, and standard deviation of the best models for every benchmark dataset without any or with 20% noise. In the first two columns multi-population genetic programming with data migration and individual migration (DM + IM) and only data migration (DM) are reported, where 20% of the training data is used as FTS and another 20% as VTS is used. The third column shows the values obtained by multi-population genetic programming without data migration using 40% of the data as FTS and the last column the values for single-population GP with random sample selection.

All multi-population algorithm configurations achieve better results as the single population algorithm (GP+RSS) on every benchmark problem. The differences between the different migration strategies is not as clear, because no single variant outperforms all others. Still the best results were obtained when data migration is used, which gives a clear indication of the usefulness of the presented approach. The differences obtained on the Nguyen-12 and Keijzer-14 dataset are rather small and the absolute values quite high, which is an indication that these two problems are easily solved regardless of the used algorithm variant. The largest differences

Table 6.4 Median, average and standard deviation (first–third row) of the qualities of multi-population GP with individual and data migration (IM + DM), either one of them, and of GP with random sample selection (GP + RSS) on all problems with 20 % noise

	IM + DM		DM		IM		GP + RSS	
	Training	Test	Training	Test	Training	Test	Training	Test
Nguyen-12	0.782	0.784	0.780	0.784	0.782	0.785	0.775	0.779
	0.782	0.714	0.781	0.773	0.783	0.743	0.774	0.749
	0.003	0.182	0.004	0.047	0.004	0.148	0.006	0.143
Keijzer-14	0.792	0.783	0.789	0.784	0.791	0.783	0.781	0.773
	0.790	0.754	0.788	0.778	0.788	0.747	0.778	0.773
	0.010	0.101	0.011	0.018	0.012	0.114	0.016	0.019
Vladislavleva-5	0.586	0.505	0.588	0.524	0.563	0.467	0.558	0.459
	0.574	0.518	0.580	0.540	0.555	0.459	0.514	0.454
	0.069	0.127	0.076	0.140	0.065	0.125	0.129	0.177
Vladislavleva-7	0.629	0.589	0.630	0.611	0.634	0.590	0.477	0.380
	0.600	0.507	0.616	0.583	0.613	0.533	0.534	0.454
	0.067	0.207	0.046	0.120	0.055	0.189	0.096	0.182
Poly-10	0.486	0.430	0.357	0.287	0.385	0.344	0.328	0.257
	0.447	0.405	0.375	0.318	0.386	0.323	0.367	0.308
	0.128	0.160	0.122	0.151	0.113	0.139	0.135	0.152
Pagie-1	0.732	0.687	0.732	0.687	0.732	0.688	0.728	0.687
	0.734	0.685	0.732	0.665	0.734	0.675	0.724	0.661
	0.011	0.021	0.004	0.116	0.007	0.089	0.020	0.101

between the different algorithms occur on the Poly-10 dataset, where the algorithm with data and individual migration (DM + IM) clearly outperforms all other ones.

These observations hold for the results generated on the problems without any noise (Table 6.3) and with noise accounting for 20 % of the total variance of the target function (Table 6.4). The absolute values on the noisy problems are smaller, because it is harder to find the data generating function and due to the noise the maximum coefficient of determination, which can be achieved when finding exactly the data generating function, is 0.8.

After the comparison of the algorithm variants, which only use a proportion of the available training data, we wanted to compare the data migration approach to standard GP algorithms that use the whole training partition for fitness calculation. These results are reported in Tables 6.5 and 6.6, where either no noise or noise accounting for 20 % of the total variance was added to the target function. The first two columns report the values obtained by standard single population genetic programming (GP) and single population GP that uses the first 20 % of the data as FTS and alternates further 20 % (VTS) every tenth generation (GP + VTS). The last two columns report the results obtained by multi-population genetic programming without (Multi GP) and with data migration (Multi GP + DM). The data migration variant and GP + VTS

Table 6.5 Quality results of the best models obtained by single population (GP), single population genetic programming using a VTS (GP + VTS) and multi-population genetic programming (Multi GP) without and with data migration (Multi GP + DM)

	GP		GP + VTS		Multi GP		Multi GP + DM	
	Training	Test	Training	Test	Training	Test	Training	Test
Nguyen-12	0.991	0.991	0.991	0.991	0.981	0.982	0.978	0.981
	0.990	0.981	0.990	0.967	0.981	0.956	0.979	0.981
	0.006	0.043	0.005	0.141	0.005	0.134	0.005	0.004
Keijzer-14	0.997	0.995	0.996	0.994	0.993	0.991	0.994	0.992
	0.996	0.972	0.995	0.983	0.991	0.989	0.992	0.990
	0.004	0.108	0.003	0.043	0.005	0.007	0.005	0.007
Vladislavleva-5	0.948	0.738	0.887	0.670	0.872	0.654	0.861	0.638
	0.911	0.674	0.841	0.619	0.872	0.671	0.852	0.627
	0.100	0.235	0.149	0.217	0.083	0.121	0.091	0.154
Vladislavleva-7	0.791	0.635	0.810	0.589	0.825	0.712	0.830	0.723
	0.784	0.576	0.767	0.548	0.802	0.706	0.794	0.669
	0.085	0.215	0.105	0.250	0.066	0.150	0.075	0.211
Poly-10	0.549	0.341	0.364	0.202	0.571	0.520	0.571	0.524
	0.554	0.381	0.450	0.319	0.572	0.519	0.549	0.510
	0.142	0.245	0.186	0.255	0.150	0.193	0.189	0.237
Pagie-1	0.923	0.841	0.922	0.839	0.905	0.838	0.903	0.839
	0.920	0.817	0.923	0.834	0.907	0.840	0.907	0.844
	0.017	0.116	0.016	0.039	0.013	0.025	0.017	0.033

The shown values are the median, average, and standard deviation (first–third row) of 50 repetitions of every algorithm on the benchmark problems without any noise

use 40 % of the training data for fitness calculation evenly split between the FTS and VTS.

The multi-population algorithms perform better compared to the single population ones on all benchmark problems and especially on the Vladislavleva-7 and Poly-10 dataset the differences are considerable. This shows that using multiple smaller populations, instead of a larger, single one, helps the algorithm to produce better results. If the rationale for the better results, is that a higher diversity is preserved during the algorithm execution or the subpopulations explore a larger proportion of the solution space is a subject to further research.

A comparison between the algorithms with and without data migration does not show any significant differences. While for the single population algorithm the results obtained without data migration are slightly better on the training and on the test partition, they are almost the same for the multi-population algorithms. On the one hand this indicates, that models generated with data migration do not generalize better compared to models generated on the whole training data. On the other hand it shows that it is not necessary to use all available training samples to generate models, which describe the relationship between the dependent and independent

Table 6.6 Quality results of the best models obtained by single population (GP), single population genetic programming using a VTS (GP + VTS) and multi-population genetic programming (Multi GP) without and with data migration (Multi GP + DM)

	GP		GP + VTS		Multi GP		Multi GP + DM	
	Training	Test	Training	Test	Training	Test	Training	Test
Nguyen-12	0.796	0.771	0.788	0.781	0.786	0.784	0.782	0.784
	0.796	0.619	0.788	0.706	0.787	0.733	0.782	0.714
	0.005	0.248	0.004	0.185	0.004	0.163	0.003	0.182
Keijzer-14	0.812	0.758	0.806	0.764	0.803	0.781	0.792	0.783
	0.812	0.648	0.803	0.699	0.802	0.745	0.790	0.754
	0.006	0.213	0.009	0.150	0.005	0.115	0.010	0.101
Vladislavleva-5	0.657	0.555	0.612	0.518	0.591	0.511	0.586	0.505
	0.608	0.517	0.564	0.458	0.587	0.527	0.574	0.518
	0.108	0.174	0.118	0.207	0.065	0.137	0.069	0.127
Vladislavleva-7	0.621	0.495	0.646	0.462	0.631	0.568	0.629	0.589
	0.616	0.423	0.601	0.427	0.616	0.496	0.600	0.507
	0.063	0.218	0.076	0.236	0.049	0.228	0.067	0.207
Poly-10	0.368	0.180	0.313	0.131	0.492	0.428	0.486	0.430
	0.419	0.253	0.321	0.152	0.478	0.414	0.447	0.405
	0.114	0.170	0.104	0.129	0.100	0.154	0.128	0.160
Pagie-1	0.751	0.687	0.741	0.682	0.735	0.688	0.732	0.687
	0.751	0.619	0.742	0.641	0.737	0.676	0.734	0.685
	0.021	0.175	0.013	0.139	0.010	0.088	0.011	0.021

The shown values are the median, average, and standard deviation (first–third row) of 50 repetitions of every algorithm on the benchmark problems with 20% noise

variables sufficiently well. Therefore, a sophisticated sample selection strategy, such as data migration, can reduce the algorithm runtime, because the most execution time consuming part of genetic programming for symbolic regression is the fitness evaluation of the individuals.

6.5 Conclusion

In this chapter we presented a novel multi-population genetic programming algorithm for symbolic regression that incorporates a sophisticated sample selection strategy. In contrast to standard algorithms, every subpopulation is evolved on a different fixed training data subset and additionally a variable training subset, which is migrated between the subpopulations at specific intervals. Multiple benchmark datasets were used to test the performance of the new algorithm and the achieved results were compared with other algorithms using only a proportion of the training samples and standard genetic programming algorithms using all of the available training samples.

This new genetic programming algorithm with data migration performs better compared to other genetic algorithms that use other sample selection strategies. Furthermore, the usage of multiple, smaller populations produces better models compared to a single, larger population on all benchmark problems. An intention for developing the new algorithm was, that by changing the samples used for fitness calculation at regular intervals, the learned models do not over adapt to the presented data and hence, generalize better on the unseen test data. However, in the performed experiments this behavior could not be observed and the performance of algorithms using data migration or all the available training data was almost identical. A benefit of using data migration in combination with multiple populations in genetic programming is that while achieving as good results as algorithms that use all the training data, the execution of the algorithm is reduced, which is especially important when dealing with large datasets.

References

1. Affenzeller, M., Wagner, S.: Offspring selection: A new self-adaptive selection scheme for genetic algorithms. In: Ribeiro, B., Albrecht, R.F., Dobnikar, A., Pearson, D.W., Steele, N.C. (eds.) *Adaptive and Natural Computing Algorithms*. Springer Computer Series, pp. 218–221. Springer, New York (2005)
2. Cantú-Paz, E.: A survey of parallel genetic algorithms. *Calculateurs paralleles, reseaux et systems repartis* **10**(2), 141–171 (1998)
3. Fernández, F., Tomassini, M., Vanneschi, L.: An empirical study of multipopulation genetic programming. *Genet. Program. Evol. Mach.* **4**(1), 21–51 (2003)
4. Gathercole, C., Ross, P.: Dynamic training subset selection for supervised learning in genetic programming. *Parallel Problem Solving from Nature-PPSN III*, pp. 312–321. Springer, New York (1994)
5. Goncalves, I., Silva, S., Melo, J.B., Carreiras, J.M.B.: Random sampling technique for overfitting control in genetic programming. In: Moraglio, A., Silva, S., Krawiec, K., Machado, P., Cotta, C. (eds.) *Proceedings of the 15th European Conference on Genetic Programming, EuroGP 2012*, vol. 7244 of LNCS, pp. 218–229. Springer, Malaga, Spain, pp. 11–13 April 2012
6. Harper, R.: Spatial co-evolution in age layered planes (SCALP). *IEEE Congress on Evolutionary Computation (CEC 2010)*. IEEE Press, Barcelona, Spain, pp. 18–23 July 2010
7. Keijzer, M.: Scaled symbolic regression. *Genet. Program. Evol. Mach.* **5**(3), 259–269 (2004)
8. Kommenda, M., Kronberger, G., Winkler, S., Affenzeller, M., Wagner, S.: Effects of constant optimization by nonlinear least squares minimization in symbolic regression. In: *Proceeding of the Fifteenth Annual Conference Companion on Genetic and Evolutionary Computation Conference Companion*, pp. 1121–1128. ACM (2013)
9. Koza, J.R.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge (1992)
10. Luke, S.: Two fast tree-creation algorithms for genetic programming. *IEEE Trans. Evol. Comput.* **4**(3), 274–283 (2000)
11. Poli, R., Langdon, W.B., McPhee, N.F.: *A Field Guide to Genetic Programming*. Lulu.com (2008)
12. Salhi, A., Glaser, H., De Roure, D.: Parallel implementation of a genetic-programming based tool for symbolic regression. *Inf. Process. Lett.* **66**(6), 299–307 (1998)
13. Schmidt, M.D., Lipson, H.: Co-evolving fitness predictors for accelerating and reducing evaluations. In: Riolo, R.L., Soule, T., Worzel, B. (eds.) *Genetic Programming Theory and Practice*

- IV, vol. 5 of Genetic and Evolutionary Computation, chapter 17. Springer, Ann Arbor, pp. 11–13 May 2006
14. Smits, G., Vladislavleva, E., Yen, G.G., Lucas, S.M., Fogel, G., Kendall, G., Salomon, R., Zhang, B.-T., Coello, C.A.C.: Ordinal pareto genetic programming. In: Proceedings of the 2006 IEEE Congress on Evolutionary, pp. 3114–3120. IEEE Press (2006)
 15. Stijven, S., Minnebo, W., Vladislavleva, K.: Separating the wheat from the chaff: on feature selection and feature importance in regression random forests and symbolic regression. In: Gustafson, S., Vladislavleva, E. (eds.) 3rd symbolic Regression and Modeling Workshop for GECCO 2011, pp. 623–630. ACM, Dublin, Ireland, 12–16 July 2011
 16. Tanese, R.: Distributed genetic algorithms. In: Proceedings of the Third International Conference on Genetic Algorithms, pp. 434–439. Morgan Kaufmann Publishers Inc., San Francisco (1989)
 17. Tomassini, M., Vanneschi, L., Fernández, F., Galeano, G.: A study of diversity in multipopulation genetic programming. *Artificial Evolution*, pp. 243–255. Springer, New York (2004)
 18. Uy, N.Q., Hoai, N.X., O’Neill, M., Mckay, R.I., Galván-López, E.: Semantically-based crossover in genetic programming: Application to real-valued symbolic regression. *Genet. Program. Evol. Mach.* **12**(2), 91–119 (2011)
 19. Wagner, S., Kronberger, G., Beham, A., Kommenda, M., Scheibenpflug, A., Pitzer, E., Vonolfen, S., Kofler, M., Winkler, S., Dorfer, V., Affenzeller, M.: Architecture and design of the heuristiclab optimization environment. In: Klempous, R., Nikodem, J., Jacak, W., Chaczko, Z. (eds.) *Advanced Methods and Applications in Computational Intelligence. Topics in Intelligent Engineering and Informatics*, vol. 6, pp. 197–261. Springer, New York (2014)
 20. White, D.R., McDermott, J., Castelli, M., Manzoni, L., Goldman, B.W., Kronberger, G., Jaskowski, W., O’Reilly, U.-M., Luke, S.: Better GP benchmarks: community survey results and proposals. *Genet. Program Evol. Mach.* **14**(1), 3–29 (2013)
 21. Whitley, D., Rana, S., Heckendorn, R.B.: The island model genetic algorithm: on separability, population size and convergence. *J. Comput. Inf. Technol.* **7**, 33–48 (1999)
 22. Whitley, D., Starkweather, T.: Genitor ii: a distributed genetic algorithm. *J. Exp. Theor. Artif. Intell.* **2**(3), 189–214 (1990)

Chapter 7

Search Strategies for Grammatical Optimization Problems—Alternatives to Grammar-Guided Genetic Programming

Gabriel Kronberger and Michael Kommenda

Abstract In this chapter, we have a closer look at search strategies for optimization problems, where the structure of valid solutions is defined through a formal grammar. These problems frequently occur in the genetic programming (GP) literature, especially in the context of grammar-guided genetic programming [18]. Even though a lot of progress has been made to extend and improve GP in the last 25 years and many impressive solutions have been produced by GP, the initial goal of an automated programming machine for generating computer programs is still far away and GP is not yet established as a reliable and general method for solving grammatical optimization problems. Instead, many different GP variants have been described and used for solving specific problems. Today the term GP refers to a large set of related algorithms where the commonality mainly is that an evolutionary algorithm is used to produce solutions which often—but not always—represent code that can be executed by a problem specific virtual machine or an interpreter. This code is most frequently represented either as a tree or as a linear chain of instructions. The term genetic programming thus categorizes algorithms based on their approach to solution manipulation. However, the type of problems that is solved using these algorithms is more general. Especially for practitioners, it is often not relevant how a solution has been produced as only the solution itself is relevant. We argue that even though genetic programming is a powerful approach, it might not always be the optimal approach for solving “genetic programming problems” and instead other algorithms might work better for certain problems. Therefore, in this chapter we take a fresh look at those problems, that we in the following call grammatical optimization problems, and discuss various ways for solving such problems. A severely trimmed down extended abstract for this chapter appeared in [14].

G. Kronberger (✉) · M. Kommenda
University of Applied Sciences Upper Austria School of Informatics,
Communications and Media, Softwarepark 11, 4232 Hagenberg, Austria
e-mail: gabriel.kronberger@fh-hagenberg.at

M. Kommenda
e-mail: michael.kommenda@fh-hagenberg.at

© Springer International Publishing Switzerland 2015
G. Borowik et al. (eds.), *Computational Intelligence and Efficiency
in Engineering Systems*, Studies in Computational Intelligence 595,
DOI 10.1007/978-3-319-15720-7_7

7.1 Motivation

GP software systems are hard to use and largely incompatible [15]. One way to improve interoperability and to make it easier for practitioners to use GP could be a standardized solver-independent format for the definition of GP problems. Therefore, we defined a descriptive problem definition language (GPD L) for “grammatical optimization problems”, with the goal to separate the definition of the problem from solver semantics [15]. GPD L is solver- and framework-independent, and therefore, a problem defined using GPD L can be attacked with different algorithms or implementations with minimal effort.

We have made a first step into this direction by publishing a draft specification for the genetic programming definition language (GPD L) [15] and a reference implementation of a compiler for the tree-based GP implementation in HeuristicLab.¹ Even though the development of GPD L has been motivated mainly to simplify problem definition for GP, grammatical optimization problems occur in many applications (e.g. in machine learning [5, 13]), and completely different approaches for solving such problems are possible and might work better than GP. Separation of problem definition and solver semantics is the necessary first step to facilitate experimentation with different solver approaches and implementations. However, from a research perspective, it is especially interesting to get a better understanding of solution methods for grammatical optimization problems. For instance, the artificial ant benchmark problem, which is widely used in the GP research community, is rather hard for simple tree-based GP, and tree-based GP or hill-climbing do not outperform simple random search [21]² on this problem. In a recent contribution, it has been demonstrated, that even symbolic regression, which is considered as one of the problems where GP is especially useful, can be solved more efficiently in certain constrained settings by exhaustive search in combination with smart reduction of the search space and dynamic programming [25]. Another specialized solution for symbolic regression that can be mentioned in this context is FFX [17] which uses an iterative approach for learning a regularized sparse linear model of non-linear basis functions as frequently used in statistical learning [7].

For decomposable problems a simple greedy approach could be sufficient to find useful solutions using stepwise greedy improvement, as shown for instance in [5].

We are especially interested in innovative alternative approaches for solving grammatical optimization problems, which use the implicit syntactic and semantic structure of these problems to improve search efficiency.

¹ Available at <http://dev.heuristiclab.com/GPD L>.

² It should be noted that the ant problem is still a very simple toy-problem but GP is not much better than random search.

7.2 Grammatical Optimization Problems

In this section we define grammatical optimization problems which are mainly inspired by previous work in the area of grammar-guided genetic programming. A strongly related approach is described in [4] where the problem of choosing fast implementations of recursive algorithms such as the fast Fourier transform is formalized as “as an optimization problem over the language generated by a suitably defined grammar” [4]. In this work a bandit-based algorithm is used for the synthesis of algorithms. Other related work in the area of constraint programming describes so-called “grammar constraints” for restricting possible solutions to adhere to a given grammar [8, 23].

Grammar-guided genetic programming [18] is used to evolve variable-length solutions to an optimization problem, where the problem is defined through a formal grammar that specifies the syntactical structure of solutions and an objective function which is used to evaluate the quality of solutions that has to be optimized.

A grammatical optimization problem is a tuple $(G(S), f)$ consisting of an attributed grammar $G(S)$ with sentence symbol S which defines a language L , and an objective function $f : L \rightarrow R$ to be optimized. Informally, the attributed grammar is a set of rewrite rules, where each rule can be linked to a semantic action that is executed when a rewrite rule is applied (cf. [11]). The set of solution candidates of a grammatical optimization problem is equivalent to the language $L(G(S))$ that can be derived from repeated application of rewrite rules starting with the sentence symbol S . The goal is to find an optimal solution w.r.t. the objective function f , by repeatedly deriving different sentences using the grammar rules.

An example for a grammatical optimization problem might use the following grammar:

$$G(S) : S \rightarrow aS \mid bS \mid a \mid b. \quad (7.1)$$

and the objective function $f : L(G(S)) \rightarrow R$, where $f(l)$ is the number of occurrences of the symbol a in the sentence l , which has to be maximized.

Note, that the definition of a grammatical optimization problem is not unique because typically a given language can be expressed using multiple grammars. For instance, all grammars shown in Eq. 7.2 are equivalent to the original grammar in Eq. 7.1 as they all define the same language.

$$\begin{aligned} G_1(S) : S &\rightarrow Sa \mid Sb \mid a \mid b. \\ G_2(S) : S &\rightarrow SS \mid a \mid b. \\ G_3(S) : S &\rightarrow aaS \mid abS \mid baS \mid bbS \\ &aa \mid ab \mid ba \mid bb \mid a \mid b. \end{aligned} \quad (7.2)$$

This poses the interesting problem which grammar definition is suited best for a given solver. We elaborate on this in more detail in Sect. 7.4.

The solution space $L(G(S))$ of grammatical optimization problems is usually infinite, because sentences can become infinitely long. Therefore, for practical purposes the solution space is usually restricted to the sub-set containing sentences up to a maximal length k .

$$L' = \{l \mid l \in L, |l| < k\}. \quad (7.3)$$

The size of the restricted solution space L' for the grammar in Eq. 7.1 is

$$|L'| = \sum_{i=1}^k 2^i \quad (7.4)$$

However, the number of possible derivation trees is larger than that, as multiple trees might be possible for the same sentence. In this problem there is only a single optimal solution that has to be found.

7.3 Approaches for Solving Grammatical Optimization Problems

7.3.1 Random Search and Full Enumeration

Random sampling of sentences from the grammar as well as full enumeration of all sentences are straight forward to implement and are useful mainly for comparison purposes. Solvers for grammatical optimization problems must find solutions using on average less evaluations of the objective function than random search. We assume that problem instances of practical relevance cannot be solved by full enumeration. However, for benchmark problems or for the simple problems discussed in this paper it might be useful to use full enumeration to produce statistics of the fitness landscape of problems such as the frequency of optimal solutions.

Random search produces sentences from the grammar by repeated derivation of full sentences from the sentence symbol. This is accomplished using a recursive approach where a non-terminal in the chain is selected and replaced by a random alternative selected from the rule for that non-terminal. Notably, uniform selection might not be ideal for certain grammars because it could lead to very long sentences (rules with many non-terminal alternatives) or to very short sentences (rules with many terminal alternatives). An implementation of the sampler should allow to adjust the probabilities of sampling terminal and non-terminal alternatives.

Full enumeration can be implemented using a breath-first procedure or iterative deepening where all sentences of length k are produced before sentences of length $k + 1$.

7.3.2 Tree-Based Genetic Programming

Tree-based genetic programming evolves trees representing derivations using the rewrite rules of the grammar. Possible derivation trees for the example grammar G_2 shown in Eq. 7.2 are shown in Fig. 7.1. The tree on the right would be the optimal solution for the above mentioned objective function and a maximal sentence length $k = 8$.

The underlying idea in tree-based genetic programming is that trees with a high objective function value can be improved by subtree crossover.³ In subtree crossover two random subtrees of two parent individuals are swapped to produce two new solution candidates. Figure 7.2 shows the results of subtree crossover for two parent trees.

For this simple example problem the assumption that an improved tree can be produced by combining two above average trees holds. Informally, for a given tree of above average quality the expected change in quality when integrating any subtree into another tree using subtree crossover is positive. Therefore, this problem can be solved efficiently using genetic programming.

It is also instructive to consider a problem that is difficult to solve using tree-based genetic programming. We use the same grammar (Eq. 7.2) but consider the objective function $f(l)$ is the length of the longest palindrome in the sentence. When using this objective function, multiple optimal solutions for a given maximum length k are possible. Some examples for $k = 8$ are shown in Fig. 7.3. The number of optimal solutions is the number of possible sentences of length $k/2$ which is in this case $2^{k/2}$ and more generally $|T|^{k/2}$ where the set T is the set of terminal symbols. The total number of possible sentences for the grammar is the same as for the simple problem and shown in Eq. 7.4. So in comparison to the simple example there are many more optimal solutions and the number of optimal solutions grows asymptotically with the same order as the number of all sentences. Nevertheless, this problem is much harder for GP than the previous problem because the assumption that swapping subtrees of above average trees improves trees, does not hold anymore. In this problem the value of a subtree depends strongly on the context where the subtree is used in the overall tree.

7.3.3 Grammatical Evolution

Grammatical evolution [19] is a variant of genetic programming where an integer vector is used to encode solution candidates. The idea in grammatical evolution is to explicitly separate the genotype and the phenotype of solution candidates. The integer vector represents the genotype which is mapped to the phenotypic tree in a translation process. The evolutionary operators crossover and mutation are applied

³ Genetic programming usually also uses a mutation operator which changes parts of trees randomly to improve diversity of the genetic material. However, in this contribution we focus on the effects of the crossover operator.

Fig. 7.1 Derivation trees for the grammar G2 shown in Eq. 7.2

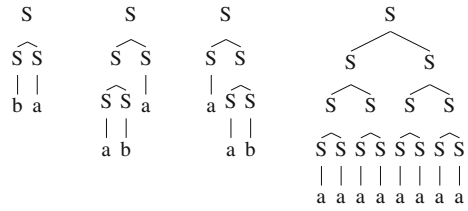
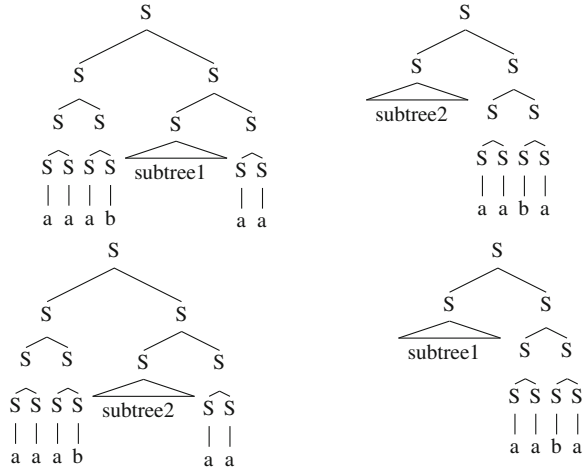


Fig. 7.2 Subtree crossover produces two new trees by swapping random subtrees from two parent trees



to the genotypes. Therefore, simple operators such as the single-point crossover or one-point mutation can be used and it is not necessary to implement more complex operators which work directly on the tree representations.

This separation is thus especially convenient for more complex grammars where the grammar constraints must be considered to produce valid solution candidates. One drawback is however the added complexity of the genotype to phenotype mapping. Especially, the so-called ripple effect is problematic because it hampers search effectiveness. The ripple effect occurs when mapping the genotype (integer vector) to the phenotype (tree). In this process the elements in the integer vector are used

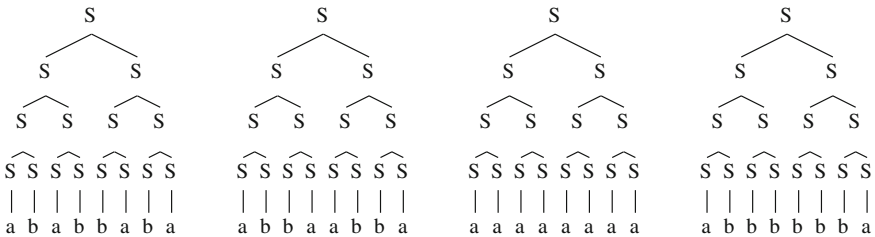
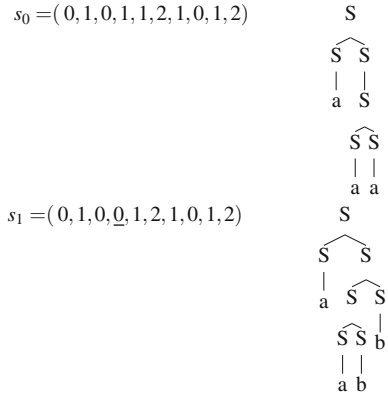


Fig. 7.3 Examples for derivation trees for the palindrome problem

Fig. 7.4 The ripple effect in grammatical evolution means that a small change in the genotype (*left side*) might lead to a large change in the phenotype (*right side*)



to select alternatives in grammar rules. Here, a small perturbation at the start of the integer vector might have a large effect in the mapping process and might lead to a completely different tree. In the mapping process it is also possible that a part of the genotypic information is not used because in the derivation of the grammar rules a full tree might be produced before the integer vector has been fully processed. Additionally, it could be possible that the tree is still incomplete even though the genotype has been fully processed.

Figure 7.4 demonstrates the ripple effect for the grammar $G_2(S)$ shown in Eq. 7.2.

Several variants of the genotype to phenotype mapping process for grammatical evolution have been described [20] to mitigate the ripple effect but a convincing solution is yet to be found. Grammatical evolution can therefore be recommended mainly because it is a simple way to implement grammar-guided GP as it is not necessary to implement complex operators for trees that adhere to grammar constraints.

7.3.4 Estimation of Distribution Algorithms

Estimation of distribution algorithms (EDA) use generative probabilistic models to first sample a set of solutions and then update the model parameters based on the empirical quality of these solutions [16]. Several variants of EDAs for trees and genetic programming that use different kinds of probabilistic models have been discussed in the literature (see for instance [9, 22, 24], or most recently [10]). One particularly interesting approach is PAGE (programming with annotated grammar estimation) which uses probabilistic context-free grammars with latent variables as the underlying model [6]. In this algorithm the grammar of the problem is first transformed to a probabilistic grammar where each alternative is derived with a certain probability. Through a probabilistic grammar it is possible to calculate a probability for each sentence of a language. The idea of the technique is to tune the parameters of the probabilistic model to increase the likelihood of good sentences.

Therefore, the algorithm repeatedly samples sentences from the grammar and updates the parameters based on the observed objective function values using a variational Bayes approximation. The authors observed that it is beneficial to add latent variables that determine the probabilities of alternatives to improve the search efficiency. Another extension of the idea uses a mixture of independent grammars with the same structure but different parameters.

It has been shown that PADE beats standard genetic programming on the deceptive max problem [6] with respect to the number of fitness evaluations. However, it has to be noted that the update of the parameters of the model is also computationally expensive.

The PADE approach is an attractive alternative to tree-based genetic programming for solving grammatical optimization problems. It only has a few parameters that need to be adjusted. Interestingly, PADE is currently not well known in the GP community and not used in studies for algorithm comparison.

7.3.5 Monte Carlo Tree Search

Monte Carlo tree search [3, 12] is a recent development in the area of reinforcement learning for optimization in Markov decision processes. A survey of MCTS algorithms and applications is given for instance in [1]. Monte Carlo tree search (MCTS) is also applicable to optimization problems [2] and grammatical optimization problems, as for instance described in [4].

MCTS for grammatical optimization problems is an extension to the simple random search algorithm as described above. The algorithm repeatedly samples sentences and chooses alternatives for non-terminal symbols randomly. However, the algorithm also keeps track of sampled sentences and their qualities in a tree. Using this information the algorithm tracks which parts of the solution space have already been visited frequently and which parts still need to be explored. In MCTS exploitation of alternatives, which lead to good solution candidates, and exploration of new alternatives is controlled by policies for bandit problems. An example for a search tree of MCTS is shown in Fig. 7.5. In this example the algorithm stores the number of tries for each alternative and the average quality of the sentences which have been generated via this alternative. The algorithm spends more effort sampling alternatives with a high average value but will also from time to time choose a relatively unexplored alternative.

In this figure the average quality of sentences for the 14 sampled sentences is $\frac{17}{14}$. The terminal alternatives “a” and “b” have been sampled only once and their qualities are one and zero, respectively. The non-terminal sentences have both been sampled six times and the quality of the first one starting with “a” is significantly larger than the quality of the alternative starting with “b”. Therefore, the algorithm would preferably sample sentences beginning with the chain “aa” in the next iterations.

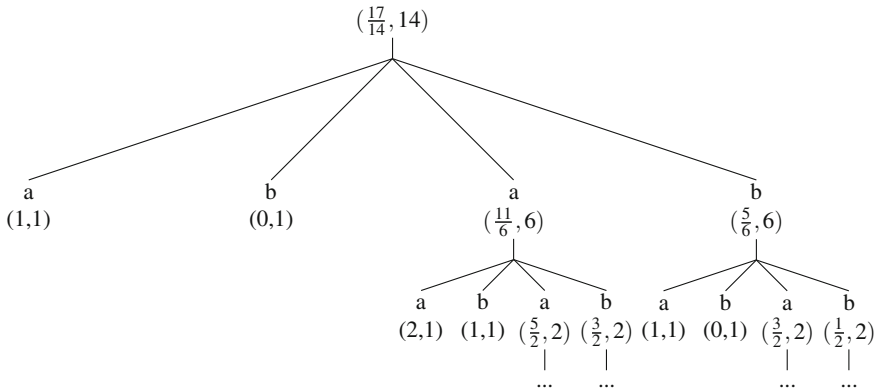


Fig. 7.5 Example of a search tree for Monte Carlo tree search. The algorithm stores the average quality and the number of visits of nodes (alternatives in the grammar)

7.4 Difficulty of Grammatical Optimization Problems

Consider the language $L(G(S))$ defined using the following regular grammar:

$$G(S) : S \rightarrow aS \mid bS \mid a \mid b. \tag{7.5}$$

where a and b are terminal symbols and S is a non-terminal symbol. Parsers for regular languages can be implemented efficiently as a finite state automaton which allow parsing of sentences in $O(n)$ time and $O(1)$ memory for sentences with n symbols using finite automata. Therefore, this grammar can be called less complex than context-free or even context-sensitive grammars. Nevertheless, it is possible to construct easy and hard grammatical optimization problems using this simple grammar as the difficulty can be controlled through the definition of the objective function.

In Sect. 7.2 we discussed the problem with the objective to maximize the number of “a” symbols and observed that this is a simple problem for genetic programming. We also discussed another problem using the same grammar with the goal to maximize the length of the longest palindrome, which is a harder problem for genetic programming.

For the first problem, if we sample sentences from the grammar we can determine the value of the objective function for each sentence. If we then go back and calculate an expected objective value over all sentences that can be generated through that alternative for each alternative in the rule for the non-terminal S , we observe that the alternatives that contain the symbol b have a lower expected objective value than the alternatives that contain an a . The expected value of each alternative can be calculated through a recurrence relation.

A simple and effective search algorithm would sample a sentence, evaluate the objective value, update an estimate for the expected objective value for the alternatives and in subsequent sampling iterations prefer those alternatives with a higher expected value. Notably, such a sampler can be implemented as a finite state machine with probabilistic state transitions.

For tree-based genetic programming this language would be represented differently to allow a balanced growth of trees. For example the same language could be defined via this context-free grammar:

$$G(S) : S \rightarrow a \mid b \mid SS.$$

In this grammar, the expected objective value of the alternative a is larger than the expected value of the alternative b and the expected value of the alternative SS would be the largest. Therefore, even if we use the above mentioned search method the optimal sentence would still be found efficiently because of the search bias for larger sentences and the symbol a .

Next, let's consider a different objective function leading to a harder problem for tree-based genetic programming. Suppose we want to find a sentence with a maximal number of pairs of different symbols (ab or ba) using the same grammar (Eq. 7.5).

When we try to apply the same approach as for the last problem, we quickly find out that both non-terminal alternatives " aS " and " bS " have the same expected objective value. Informally, we can argue that this is obvious as neither of them has a clear benefit w.r.t. the new objective function. Both alternatives would be sampled equally often. Thus, we might try to change the grammar so that alternatives can be selected greedily as before. The following reformulated grammar might be a possible idea:

$$G'(S) : S \rightarrow a \mid b \mid aaS \mid abS \mid baS \mid bbS. \quad (7.6)$$

Now, the alternatives " abS " and " baS " have the same expected value and are better than the the alternatives " aaS " and " bbS ". However, as the maximal sentence length becomes larger the difference between the two groups becomes smaller and smaller. The simple sampling approach would still sample many sentences starting with " aa " or " bb ".

For this objective function a context-sensitive grammar as shown in Eq. 7.7 correlates well with the fitness function

$$\begin{aligned} G_{context}(S) : \\ S &\rightarrow a \mid b \mid aS \mid bS . \\ aS &\rightarrow aa \mid ab \mid aaS \mid abS . \\ bS &\rightarrow ba \mid bb \mid baS \mid bbS . \end{aligned} \quad (7.7)$$

As before the two non-terminal alternatives in the first rule have the same expected value. However, for the following two rules there is only one alternative (" abS " and " baS ") that has a higher expected value than the other alternatives, and again,

a sampler that prefers those alternatives would efficiently find optimal solutions for the problem. The only added information that is necessary for the sampler is the context on the left hand side of the rule. Therefore, when updating the parameters of the sampler based on the quality of sampled solutions it is necessary to also learn weights of alternatives depending on the context.

In this example it is sufficient to consider only contexts of length one. However, generally it will be necessary to also consider longer contexts. One example, where it is necessary to learn long contexts is the palindrome problem discussed in the previous sections.

The palindrome problem is much easier to solve when the grammar only allows palindromic sentences. For instance the context-free grammar shown in Eq. 7.8 produces only palindromes. Therefore, the optimization task reduces to the trivial task of finding the longest sentence in this language.

$$G_{\text{palindrome}}(S) : S \rightarrow a \mid b \mid aa \mid bb \mid aSa \mid bSb. \quad (7.8)$$

However, in practice it is not possible to reduce the solution space in this manner and initially the full solution space of the original grammar has to be taken into consideration. An alternative context-sensitive grammar for the original language is shown in Eq. 7.9.

$$\begin{aligned} G'_{\text{palindrome}}(S) : \\ S &\rightarrow a \mid b \mid aS \mid bS. \\ aS &\rightarrow aa \mid ab \mid aSa \mid aSb. \\ bS &\rightarrow ba \mid bb \mid bSa \mid bSb. \end{aligned} \quad (7.9)$$

Again, when the alternative probabilities are updated depending on the context then in each rule the sampling algorithm would prefer the alternative that produces the best solution in expectation. Of course, the structure of the grammar plays a major role and it is not trivial to determine the matching grammar automatically.

Based on the discussion so far we propose the following MCTS algorithm that takes as input a grammatical optimization problem $(G(S), f)$ and the maximal length of sentences k and produces a sentence $s \in L(G(S))$ for which $f(s)$ is minimal/maximal. Start with the original grammar $G(S)$ and a uniform distribution for all alternatives in all rules and sample a set of sentences s_i where $|s_i| \leq k$. For each sampled sentence, store the full derivation tree so that it is possible to retrace each derivation step. Evaluate each sentence using f and update for each rule and alternative how often it has been tried and the sum of qualities of all sentences that have been derived using this alternative. In the next sampling iteration use weighted sampling for alternatives using an appropriate policy for bandit-problems. When the number of tries of one alternatives exceeds a predefined threshold t then create children for this node where each child stores statistics for this alternative dependent on the context of length one. The combination of the statistic values in all child nodes represents the statistic values of the parent node. Thus, when updating the statistics

for this alternative, the previous symbol in the actually sampled sentence determines which of the context-specific child nodes is updated. Subsequently, when sampling trees also take the previous symbol in the sampled sentence in consideration. Continue this procedure recursively for larger context lengths.

7.5 Summary

In this chapter we defined grammatical optimization problems as a general formulation for the problem types that are usually solved by grammar-guided genetic programming algorithms. Our motivation for the general formulation is that for some problem instances genetic programming might not be the ideal approach and we believe that it could be useful to attack those problems with other solver types. This argument is in line with our previous work [15] where we defined a framework- and solver-independent language for the definition of genetic programming problems (GDDL). A GDDL problem definition is a self-contained description of a grammatical optimization problem that contains the syntax and semantics of the language (solution space) and an objective function to be optimized. Furthermore, we discussed different approaches that can be used to solve grammatical optimization problems and discussed easy and hard synthetic problem instances.

It is especially instructive to consider which properties of problem instances make it easier or harder to solve a problem using genetic programming, as this information can be used to construct improved solvers. We observed that genetic programming is more efficient for problems where the syntactical structure of the search space correlates with the semantics. In particular, it is beneficial if subtrees from above average solution candidates can be integrated into other solution candidates at any position and still have a positive effect on the quality in expectation (position-independence). From this observation we concluded that in grammatical optimization problems the context of subtrees is especially relevant and problems where the average length of the relevant context is short are easier to solve. We described an approach based on policies for bandit-problems where we track the expected quality of alternatives based on their context.

The next steps are the implementation and refinement of this approach and the implementation of a GDDL compiler which generates a solver using this approach for any grammatical optimization problem.

Acknowledgments The work described in this chapter has been done within the project HOPL—Heuristic Optimization in Production and Logistics and supported by the Austrian Research Promotion Agency (FFG) within the COMET programme.

References

1. Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games* **4**(1), 1–43 (2012)
2. Chaslot, G., De Jong, S., Saito, J.-T., Uiterwijk, J.: Monte-Carlo tree search in production management problems. In: Proceedings of the 18th BeNeLux Conference on Artificial Intelligence. pp. 91–98 (2006)
3. Coulom, R.: Efficient selectivity and backup operators in monte-carlo tree search. In: *Computers and Games*, pp. 72–83. Springer (2007)
4. de Mesmay, F., Rimmel, A., Voronenko, Y., Püschel, M.: Bandit-based optimization on graphs with application to library performance tuning. In: Proceedings of the 26th Annual International Conference on Machine Learning, ICML'09, pp. 729–736. ACM, New York (2009)
5. Duvenaud, D., Lloyd, J. R., Grosse, R., Tenenbaum, J.B., Ghahramani, Z.: Structure discovery in nonparametric regression through compositional kernel search (2013). arXiv preprint [arXiv:1302.4922](https://arxiv.org/abs/1302.4922)
6. Hasegawa, Y., Iba, H.: Latent variable model for estimation of distribution algorithm based on a probabilistic context-free grammar. *IEEE Trans. Evol. Comput.* **13**(4), 858–878 (2009)
7. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning*, vol. 2. Springer, New York (2009)
8. Kadioglu, S., Sellmann, M.: Grammar constraints. *Constraints* **15**(1), 117–144 (2010)
9. Kim, K., McKay, B.R., Punithan, D.: Sampling bias in estimation of distribution algorithms for genetic programming using prototype trees. In: *PRICAI 2010: Trends in Artificial Intelligence*, pp. 100–111. Springer (2010)
10. Kim, K., Shan, Y., Nguyen, X., McKay, R.: Probabilistic model building in genetic programming: a critical review. *Genet. Progr. Evol. Mach.* **15**(2), 115–167 (2014)
11. Knuth, D.: Semantics of context-free languages. *Math. Syst. Theory* **2**(2), 127–145 (1968)
12. Kocsis, L., Szepesvári, C.: Bandit based Monte-Carlo planning. In: *Machine Learning: ECML 2006*, pp. 282–293. Springer (2006)
13. Kronberger G., Kommenda, M.: Evolution of covariance functions for Gaussian process regression using genetic programming (2013). arXiv preprint [arXiv:1305.3794](https://arxiv.org/abs/1305.3794)
14. Kronberger, G., Kommenda, M.: Search strategies for grammatical optimisation problems—alternatives to grammar-guided genetic programming. In: Proceedings of the 2nd Asia-Pacific Conference on Computer Aided System Engineering, APCASE 2014, 10th–12th February 2014, South Kuta, Indonesia, p. 101. APCASE Foundation (2014)
15. Kronberger, G., Kommenda, M., Wagner, S., Dobler, H.: GPDFL: a framework-independent problem definition language for grammar-guided genetic programming. In: *Proceeding of the Fifteenth Annual Conference Companion on Genetic and Evolutionary Computation Conference Companion*, pp. 1333–1340. ACM (2013)
16. Larrañaga, P., Lozano, J.A., *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*, vol. 2. Springer (2002)
17. McConaghy, T.: Ffx: fast, scalable, deterministic symbolic regression technology. In: Riolo, R., Vladislavleva, E., Moore, J.H. (eds.) *Genetic Programming Theory and Practice IX, Genetic and Evolutionary Computation*, pp. 235–260. Springer, New York (2011)
18. McKay, R.I., Hoai, N.X., Whigham, P.A., Shan, Y., O’Neill, M.: Grammar-based genetic programming: a survey. *Genet. program. Evol. Mach.* **11**(3/4), 365–396 (2010)
19. O’Neill, M., Ryan, C.: *Grammatical Evolution: Evolutionary Automatic Programming in an Arbitrary Language*, vol. 4. Springer (2003)
20. O’Neill, M., Ryan, C., Keijzer, M., Cattolico, M.: Crossover in grammatical evolution. *Genet. program. Evol. Mach.* **4**(1), 67–93 (2003)
21. Poli, R., Langdon, W.: *Foundations of Genetic Programming*, vol. 103, p. 107. Springer, New York (2002)

22. Sastry, K., Goldberg, D.: Probabilistic model building and competent genetic programming. In: Riolo, R., Worzel, B. (eds.) *Genetic Programming Theory and Practice*. Genetic Programming Series, vol. 6, pp. 205–220. Springer (2003)
23. Sellmann, M.: The theory of grammar constraints. In: Benhamou, F. (ed.) *Principles and Practice of Constraint Programming—CP 2006*. Lecture Notes in Computer Science, vol. 4204, pp. 530–544. Springer, Berlin (2006)
24. Shan, Y., McKay, R.I., Essam, D., Abbass, H.A.: A survey of probabilistic model building genetic programming. In: *Scalable Optimization via Probabilistic Modeling*, pp. 121–160. Springer (2006)
25. Worm, T., Chiu, K.: Prioritized grammar enumeration: symbolic regression by dynamic programming. In: *Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation Conference*, pp. 1021–1028. ACM (2013)

Chapter 8

Identification and Classification of Objects and Motions in Microscopy Images of Biological Samples Using Heuristic Algorithms

Stephan M. Winkler, Susanne Schaller , Daniela Borgmann, Lisa Obritzberger, Viktoria Dorfer, Christian Haider, Sandra Mayr, Peter Lanzerstorfer, Claudia Loimayr, Simone Hennerbichler-Lugscheider, Andrea Lindenmair, Heinz Redl, Michael Affenzeller, Julian Weghuber and Jaroslav Jacak

Abstract Heuristic algorithms are used for solving numerous modern research questions in biomedical informatics. We here summarize ongoing research done in this context and focus on approaches used in the analysis of microscopic images of biological samples. On the one hand we discuss the use of evolutionary algorithms for detecting and classifying structures in microscopy images, especially micro-patterns, cornea cells, and strands of myocardial muscles. On the other hand we show the use of data mining for characterizing the motions of molecules (for recognizing cells affected by paroxysmal nocturnal hemoglobinuria) and the progress of bone development.

S.M. Winkler (✉) · S. Schaller · D. Borgmann · L. Obritzberger · V. Dorfer · C. Haider · M. Affenzeller
Bioinformatics Research Group and Heuristic and Evolutionary Algorithms Laboratory,
University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria
e-mail: stephan.winkler@fh-hagenberg.at

S. Schaller
e-mail: susanne.schaller@fh-hagenberg.at

D. Borgmann
e-mail: daniela.borgmann@fh-hagenberg.at

L. Obritzberger
e-mail: lisa.obritzberger@fh-hagenberg.at

V. Dorfer
e-mail: viktoriam.dorfer@fh-hagenberg.at

C. Haider
e-mail: christian.haider@fh-hagenberg.at

M. Affenzeller
e-mail: michael.affenzeller@fh-hagenberg.at

8.1 Introduction

In several biomedical research approaches, fluorescence microscopy image analysis is an essential approach for researching and investigating the development as well as the state of various diseases. We here summarize ongoing research conducted at the Bioinformatics Research Group¹ and the Heuristic and Evolutionary Algorithms Laboratory (HEAL)² at the University of Applied Sciences Upper Austria in the application of heuristic algorithms for solving biomedical problems on the basis of microscopy images. The following approaches are pursued in this context:

- On the one hand we have successfully developed approaches using evolutionary algorithms, especially evolution strategies (ES), for identifying structures in micro- and nanoscale microscopy images.
- On the other hand we use machine learning approaches, especially regression and classification algorithms for identifying classification and prognosis models of dynamics seen in series of microscopy images. The approaches applied here

¹<http://bioinformatics.fh-hagenberg.at/>.

²<http://heal.heuristiclab.com/>.

S. Mayr · J. Jacak
Department of Medical Engineering, University of Applied Sciences Upper Austria,
Garnisonstraße 21,4020 Linz, Austria
e-mail: sandra.mayr@fh-linz.at

J. Jacak
e-mail: jaroslaw.jacak@fh-linz.at

C. Loimayr · S. Hennerbichler-Lugscheider
Red Cross Blood Transfusion Service for Upper Austria, Austrian Cluster for Tissue
Regeneration, Krankenhausstraße 7,4020 Linz, Austria
e-mail: claudia.loimayr@o.ropeskreuz.at

S. Hennerbichler-Lugscheider
e-mail: simone.hennerbichler@o.ropeskreuz.at

A. Lindenmair
Ludwig Boltzmann Institute for Experimental and Clinical Traumatology,
Donaueschingenstraße 13,1200 Wien, Austria
e-mail: andrea.lindenmair@trauma.lbg.ac.at

H. Redl
Trauma Care Consult, Gonzagagasse 11/25,1010 Wien, Austria
e-mail: office@traumacareconsult.com

P. Lanzerstorfer · J. Weghuber
Department of Food Technology and Nutrition, University of Applied Sciences
Upper Austria, Stelzhamerstraße 23, 4600 Wels, Austria
e-mail: peter.lanzerstorfer@fh-wels.at

J. Weghuber
e-mail: julian.weghuber@fh-wels.at

include the HeuristicLab implementations of support vector machines, random forests, neural networks, and genetic programming.

In detail, we here describe approaches for solving the following problems:

- Identification of patterns in microscopy images of biological samples (Sect. 8.2)
- Classification of endothelial cells in cornea (Sect. 8.3)
- Identification of PNH affected cells by classifying motion characteristics of single molecules (Sect. 8.4)
- Identification of strands in microscopy images of myocardial muscles (Sect. 8.5)
- Modeling and prognosis of bone development in amnion tissues (Sect. 8.6).

An extended abstract of the content of this chapter has been previously published in the Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014 [28].

8.2 Identification of Patterns in Microscopy Images of Biological Samples Using Evolution Strategies

In the analysis of images of fluorescently labeled biological samples using μ -patterning [21], one of the challenges is to efficiently identify symmetric grid structures in nanoscale images. In this context we apply an approach based on ES that is able to automatically identify such grids [4].

8.2.1 Biological Background

The cellular membrane of biological cells is an important integral part of cellular interaction processes. The membrane itself represents a physical barrier between intracellular and extracellular spaces and hereby generates a tiny reaction volume—the biological cell. In order to guarantee the proper function of a cell, interactions between the extracellular environment and the intracellular space are necessary.

A very sensitive method to detect protein-protein interactions in living cells is called μ -patterning assay [10, 21, 25]. A total internal reflexion fluorescence (TIRF) based microscopy imaging system was used to detect the interactions between a fluorophore-labeled protein and a membrane protein in a live cell context.

The goal of the research work summarized here in this paper is to develop an algorithm that is able to automatically identify grid structures in images of biological samples labeled by μ -patterning. Within these images the patterns have to be separated from areas that represent the lattices. Thus, what is needed is a method that is able to identify symmetric grids within the grid images and uses these grids to identify pattern areas in the corresponding pattern images.

8.2.2 Image Preprocessing

In order to decrease the runtime consumption of further image analysis steps (including the identification of grid structures), the analyzed images are down sampled [11] using a correlation threshold: Starting with down sampling rate (d_{sr}) 2, the d_{sr} is constantly increased until the correlation [18] of the down sampled image and the original image becomes less than θ . In order to distinguish pattern areas from lattice, the images are transformed to binary images using greyscale thresholding.

8.2.3 Optimization of Grids for Detecting Patterns in Biological Samples Using Evolution Strategies

8.2.3.1 Evolution Strategies

Evolution strategies (ES; [17, 22]), beside genetic algorithms (GAs, [1]) the second major representative of evolutionary computation, are population based, i.e., each optimization process works with a population of potential solution candidates that are initially created randomly and then iteratively optimized. In each generation new solution candidates are generated by randomly selecting parent individuals and forming new individuals applying mutation and optimally crossover operators; λ children are produced by μ parent individuals. Typically, parent selection in ES is performed uniformly randomly, with no regard to fitness; survival in ESs simply saves the μ best individuals, which is only based on the relative ordering of fitness values.

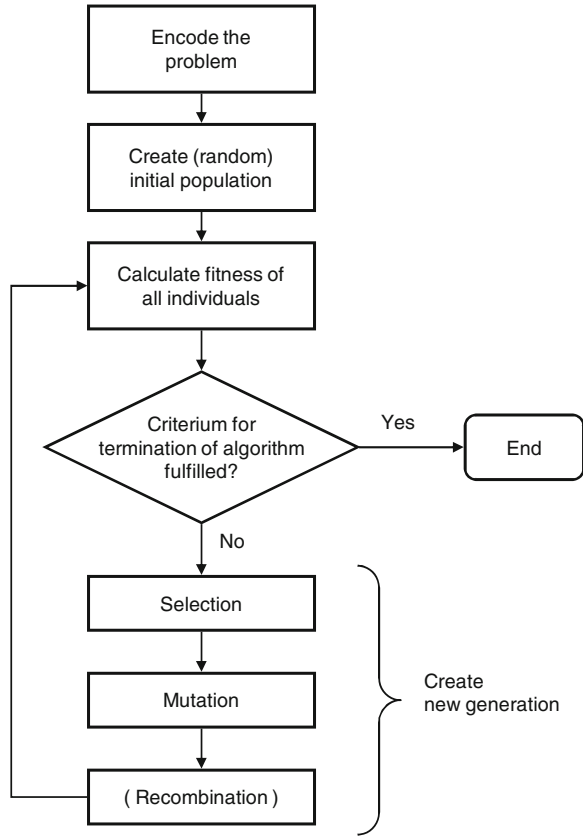
Thus, the main driving forces of optimization in ESs are offspring selection and mutation. For the optimization of vectors of real values, mutation is usually implemented as additive Gaussian perturbation with zero mean or multiplicative Gaussian perturbation with mean 1.0. Mutation width control is based on the quotient of the number of the mutants that are better than their parents: If this quotient is greater than $1/5$, then the mutation variance is to be increased; if the quotient is less than $1/5$, the mutation variance should be reduced. Additionally, this ES workflow is extended as proposed by Schwefel [22] so that there is an individual mutation strategy parameter for each parameter of the solution candidate. Thus, the mutation strength of each feature is also optimized during the evolutionary process.

Figure 8.1 shows the main workflow of an evolution strategy.

8.2.3.2 Solution Candidates Representing Grids

A solution candidate is represented as a composition of four parameters: the grid's deflection, width, horizontal offset and vertical offset of a reference grid line.

Fig. 8.1 The main workflow of an evolution strategy [27]



The actual *quality of a solution candidate (grid)* is computed by the comparison of each pixel of the expected grid image ($expected(grid)$) with the corresponding pixel of the binary representation of the original image ($binary$). The number of positive comparisons is summed up over all pixels p and divided by the total number of pixels N . The retrieved value is within the range $[0, 1]$.

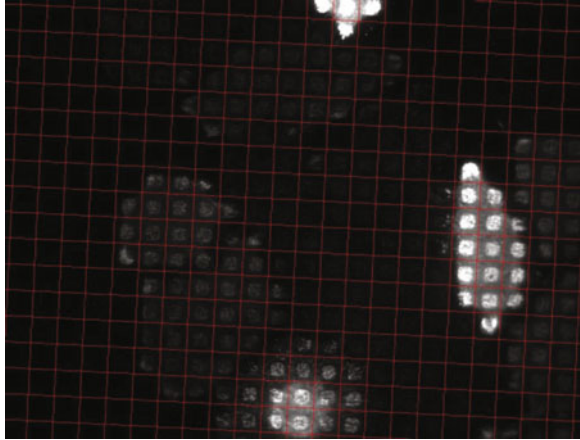
$$N(binary) = |\{p | p \in binary\}| \quad (8.1)$$

$$P(grid, binary) = |\{expected(grid)[p] = binary[p]\}| \quad (8.2)$$

$$quality(grid, binary) = P(grid, binary) / N(binary) \quad (8.3)$$

In order to start the identification of grid structures we determine *initial solution candidates* which are calculated from the binary image. The first step is to define three grid lines (two vertical and one horizontal) in order to be able to complete the grid definition. A grid line itself is composed of an offset-value, which determines the position within the image, and its deflection; it can be evaluated by calculating the

Fig. 8.2 Grid identified for an image taken using fluorescence microscopy. This result was produced by a 5 + 10 evolution strategy that was executed over 40 generations



percentage of white pixel-values compared to the binary image (Details regarding the genetic operators used here can be found in [4]).

8.2.4 Results

In order to demonstrate ability of this approach to identify optimal grids in biological microscopy images, series of exemplary samples have been selected. Analyzing numerous series of test runs we see that the here described approach was for all images able to find correct grids; the best results are achieved setting the down sampling threshold to 0.9 and using ES with plus selection and relatively small populations (e.g., $\mu = 5$ and $\lambda = 10$).

Figure 8.2 shows an exemplary grid identified for a fluorescence microscopy image.

8.3 Shape Classification of Endothelial Cells in Cornea

In the analysis of images of a cornea, the transparent front part of the eye, the two main challenges are to correctly identify endothelial cells and to classify the shape of the cells regarding the number of edges [6]. In ongoing research we develop an approach that uses ES for optimizing the corners of forms with four, five, six, or seven edges.

8.3.1 *Biological Background, Research Goal*

Cornea transplantation may be the only recovery chance for numerous diseases such as Fuchs' Dystrophy, keratoconus or similar disorders [8]. Therefore, allogeneic donor cornea is used for transplantation [6]. To minimize risks of cornea transplantations, the tissue donation has to pass strict quality checks. Amongst other requirements, the minimum of endothelial cell count should not fall below 2000 cells per mm^2 to be adequate for transplantation.

Usually, cornea endothelial cells have a hexagonal cell form. The typical form of endothelial cells gets lost during apoptosis. In this process cells lose their original hexagonal form. The higher the number of cells with hexagonal shape the better is the quality of a cornea and the higher is the possibility to be used for transplantation. Thus it is not only important to identify the endothelial cell count, but also to analyze the shape of the cells on the cornea that are to be used for transplantation.

8.3.2 *Pre-processing*

Images of corneas have to be preprocessed before cells can be identified and classified; the following preprocessing steps are performed: Identification of focused areas, dilation for detecting cell margins, transformation to binary images, and application of a flood fill algorithm for separating cells. Details of these methods are summarized in [14].

8.3.3 *Identification of Cell Forms Using Evolution Strategies*

A solution candidate for the ES used here to detect the correct cell morphologies is represented as a list of points which define the corners of a shape.

Initial solution candidates have a uniformly distributed possibility of being quadrangular, pentagonal, hexagonal or heptagonal. As the best fit can be created by using pixels that are part of the cell border, the cell border pixel with the minimum distance from the original corner of the initial solution candidate is calculated for each corner and assigned as the new corner. The so created solution candidates are the first individuals of the ES' population.

The *fitness function* has to return a value that reflects how well the shape fits into the original cell area. To get the fitness f of a solution candidate s , the area of the created shape is calculated and compared with the area of the cell. Overlapping areas lead to improvement of the fitness value, whereas an area that does not fit decreases the fitness value. Thereby the overlapping areas are weighted with a value α_1 and

points outside the shape are weighted with a value α_2 :

$$f(s) = \sum_{p \in ca(s)} I(p) \cdot \alpha_1 - \sum_{p \notin ca(s)} I(p) \cdot \alpha_2 \quad (8.4)$$

where $ca(s)$ is the area covered by the shape s . The aim of this approach is to maximize the values returned by the fitness function. The degree of reward/punishment depends on the grayscale intensity I of the pixels. Additionally, as it is easier for shapes with a higher number of corners to fit the original cell, fitness has to be weighted by the number of corners.

After calculating the fitness of each parent they are *mutated* in order to create children as new candidates from whom the best ones are promoted to the next generation. Mutation is accomplished by mutating points (i.e., randomly changing the position of the points defining the shape's corners) as well as by increasing or decreasing the number of points defining the structure.

8.3.4 Results

We have used evolution strategies that optimize cell form candidates in order to classify the cells given in cornea samples provided by the Red Cross Blood Transfusion Services; in [14] we give a detailed overview of these results.

The best results have been achieved by using the following weighting coefficients wc for the numbers of corners: $wc(4) = 1.6$, $wc(5) = 1.4$, $wc(6) = 1.2$, $wc(7) = 1.0$. Using (1+10)-ES and these weighting coefficients we were able to classify 70.03 % of the analyzed cells correctly³; for 22.77 % of the number of corners differed by not more than 1 from the correct classification.

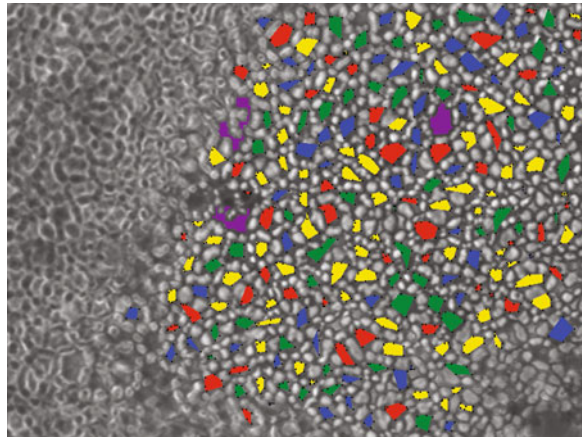
Figure 8.3 shows an exemplary image with cells that are (within the sharp region of the image) classified as quadrangular, pentagonal, hexagonal, heptagonal, or irregular.

8.4 Identification of PNH Affected Cells by Classifying Motion Characteristics of Single Molecules

In order to provide the evidence of paroxysmal nocturnal hemoglobinuria (PNH), a rare, life-threatening, mutational disease of the red blood cells, in nanoscale microscopy images, we pursue an approach that combines the identification of single molecules (SMs) and the classification of SM motions [20].

³ The correct cell shape classifications were in these tests defined by a tissue bank technician.

Fig. 8.3 Microscopy image of a cornea. Using an ES based approach, cells in the sharp region of the image are automatically classified. Cells identified as quadrangular are shown in *blue*, pentagonal ones in *green*, hexagonal ones in *yellow*, and heptagonal ones in *red*; irregular regions that are not identified as cells are shown in *purple*



8.4.1 Biological Background, Research Goal

Paroxysmal nocturnal hemoglobinuria (PNH) is a very rare disease that is characterized as a clonal disorder of hematopoietic stem cells [19] and affects blood cells; it is also referred to as the Marchiafava-Micheli syndrome. This disease disables some proteins from docking, especially CD59 (protectin), on the surface of the red blood cells; this can lead to an underproduction of those cells, hemolytic anemia, kidney failure, and a high incidence of life-threatening venous thrombosis.

The goal of the research work summarized here is to find an efficient approach for detecting and tracking single molecules to characterize and recognize PNH as well as its disease states based on microscopic images (Fig. 8.4). The following four major steps are necessary to accomplish this:

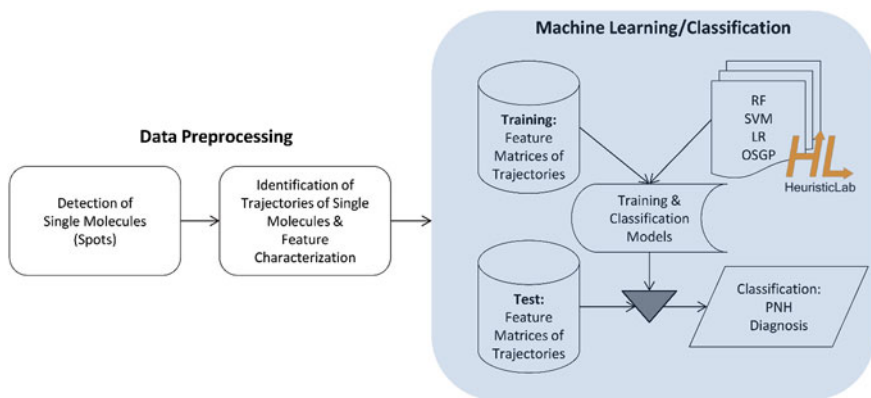


Fig. 8.4 An integrated image processing and machine learning workflow for detecting PNH affected cells [20]

- Single molecules are detected in a sequence of images including single molecule signals via Gaussian fitting methods.
- Trajectories of single molecules are detected over series of images.
- The so determined trajectories are characterized using a set of features as summarized in Sect. 8.4.2.
- Finally, the features of the trajectories are used as input for machine learning in order to learn classifiers that are able to distinguish between PNH affected and healthy patients.

8.4.2 Detection and Characterization of Trajectories of Single Molecules

An algorithm for the detection of single molecules has been developed based on the ideas described in [15] and [13]; details of this method are summarized in [20].

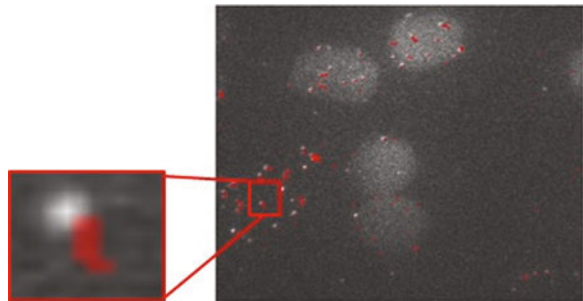
For identifying trajectories, the following approach described in [26] is used: Positions of a single molecule found in subsequent frames of images (i.e., series of images) are compared to each other. We check if a spot in frame i and in frame $i + 1$ can be identified with distance $d < D_{max}$; a spot in frame i is represented as the center of a circle with radius $d < D_{max}$. If exactly one spot in frame $i + 1$ is found with $d < D_{max}$, then a trajectory T is formed to the corresponding spot in frame i . If a circle in frame $i + 1$ contains more than one spot, the trajectory T is continued with that spot that best fits the trend of the trajectory. The accuracy of the trajectories as well as the algorithm can be tested by calculating the diffusion constant $Diff$ of the increase of the mean square displacement (msd) with different time-lags (formulae are used as described in [26]).

Figure 8.5 shows an exemplary trajectory of a single molecule over 12 frames.

The following features are defined for a trajectory T :

- The total number of single molecules in T , $length(T)$
- The average intensity of a trajectory, $intensity(T)$
- The average size of the hops in T , $spotDist(T)$

Fig. 8.5 Visualization of the trajectory of a single molecule tracked over 12 frames [20]



- The variance of the distances between the positions in T , $varDist(T)$
- The distance between the first and the last spot in T , $maxDist(T)$
- The ratio between the length of the trajectory and $maxDist(T)$, $distRatio(T)$
- The average motion changes within a trajectory, $motion(T)$.
- The convex hull of T , $polyArea(T)$.

This information is then in combination with class labels (healthy vs. PNH affected) used as input for machine learning algorithms.

8.4.3 Classification Test Results

The machine learning techniques applied here include support vector machines [23], random forests [5], neural networks [7], and genetic programming [9, 16, 27]; we have used the implementations of these algorithms in HeuristicLab [24], an open source framework for heuristic optimization developed and maintained by HEAL members.

In [20] we have documented results achieved using this approach that show that more than 80% of the analyzed test trajectories are correctly classified as belonging to healthy persons or patients suffering from PNH. Even after removing the feature $intensity(T)$, which is here the most significant feature for machine learning algorithms, 59.45% of the analyzed cells were correctly classified using genetic programming, 51.82% using k-nearest-neighbor classification, 59.89% using random forests, and 58.14% using support vector machines. The baseline for these classification tests was 50% as the number of trajectories of PNH-affected patients was equal to the number of trajectories of healthy persons.

8.5 Identification of Strands in Microscopy Images of Cardiomyocytes

In ongoing research we develop an approach based on ES for identifying sets of strands in microscopy images of cardiomyocytes, i.e., heart muscle cells. The regularity of such strands is an important factor as there might be connections to several heart diseases [2]. An example image is shown in the left part of Fig. 8.6: We see Z discs that are the boundaries of a sarcomere; a sarcomere is the smallest morphological subunit of a muscle that comprises the so-called actin and myosin fibers that slide into each other during muscle contraction leading to a constriction of the heart.

The goal here is to automatically identify such sets of strands. We are analyzing images recorded at the Research Center Linz of the University of Applied Sciences Upper Austria using atomic force microscopy (AFM, [3]).

The approach pursued here to identify sets of parallel strands is very similar to the one summarized in Sect. 8.2.

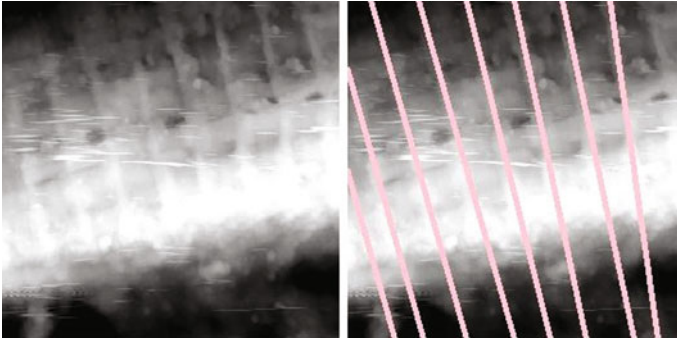


Fig. 8.6 *Left* Microscopy image of a cardiomyocyte. The goal here is to automatically identify sets of strands in the image. *Source* University of Applied Sciences Upper Austria, Research Center Linz. *Right* Parallel strands identified within the image

Initially, single strand candidates are created as lines with random position, declination, and thickness.

As such a line shall represent a strand found in the image and a strand seems to be a straight region that is lighter than its surrounding, the *fitness* of such a strand candidate sc is defined as the intensity of the pixels in sc compared to the intensity of the surrounding area of sc :

$$f(sc) = \frac{\sum_{p \in sc} I(p)}{|\{p \in sc\}|} - \frac{\sum_{p \in surr(sc)} I(p)}{|\{p \in surr(sc)\}|} \quad (8.5)$$

where $I(p)$ again is the intensity of a pixel p and $surr(sc)$ is the area surrounding sc , in this case all pixels with distance not more than 10 pixels to sc .

The best so found strand candidate is optimized, i.e., its thickness, declination, and exact position are optimized using an ES.

On the basis of the so found optimal initial strand, parallel strands are identified: Parallel copies of the strand are moved along the image and optimized using an ES-position, thickness, and inclination are optimized.

The result of this procedure is a set of parallel strands with optimized inclinations and varying distances from each other. An example is shown in the right part of Fig. 8.6.

8.6 Modeling and Prognosis of Bone Development

Humane amniotic membrane is the innermost of fetal membranes and a part of the placenta; for research purposes, human placenta is obtained naturally in the course of births. The viable (i.e., living) amniotic membrane releases a whole array of soluble factors which have a beneficial effect on wound healing after skin lesions [29].

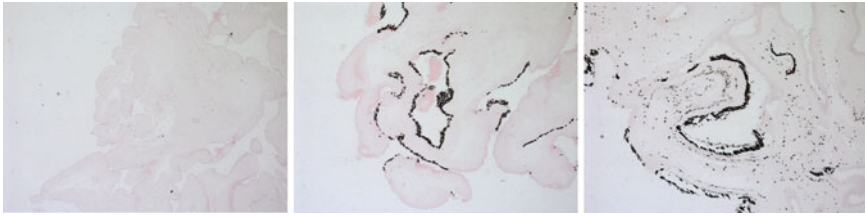


Fig. 8.7 Von Kossa staining of amniotic membrane under osteogenic conditions on day 0, day 14 and day 28; *black* regions show positive regions for bone development. *Source* Red Cross Blood Transfusion Service Linz

Being able to distinguish between viable and dead regions would be a very important screening opportunity for optimal wound treatment after big skin lesions. Furthermore, amniotic membranes are used in tissue engineering. It has been recently shown that it is possible to differentiate viable amniotic membrane towards the osteogenic tissue (i.e., bone cells) [12]. Mineralization and the expression of osteogenic markers in human amnion could be successfully induced under appropriate culture conditions. However, induced differentiation was not distributed homogeneously but rather showed patches of positive areas (as shown in Fig. 8.7).

Current research focuses on developing a machine learning approach that is able to automatically identify the mentioned cell development states. An automated prediction of the development of cells using advanced image analysis and machine learning is researched.

Acknowledgments The authors cordially thank their research partners at Red Cross Blood Transfusion Service of Upper Austria, Olympus Austria, Trauma Care Consult, and at the Research Centers Hagenberg, Wels, and Linz of the University of Applied Sciences Upper Austria for their ongoing support. The work described in this paper was done within the research projects *Micro-Prot* (sponsored by the University of Applied Sciences Upper Austria within its basic research programme) and *NanoDetect* (sponsored by the Austrian Research Promotion Agency within the FIT-IT programme).

References

1. Affenzeller, M., Winkler, S., Wagner, S., Beham, A.: Genetic Algorithms and Genetic Programming—Modern Concepts and Practical Applications. Chapman & Hall/CRC (2009)
2. Bernardo, B.C., Weeks, K.L., Pretorius, L., McMullen, J.R.: Molecular distinction between physiological and pathological cardiac hypertrophy: experimental findings and therapeutic strategies. *Pharmacol. Ther.* **128**(1), 191–227 (2010)
3. Binnig, G., Quate, C.F.: Atomic force microscope. *Phys. Rev. Lett.* **56**(9), 930–933 (1986)
4. Borgmann, D., Weghuber, J., Schaller, S., Jacak, J., Winkler, S.M.: Identification of patterns in microscopy images of biological samples using evolution strategies. In: Proceedings of the 24th European Modeling and Simulation Symposium EMSS 2012, pp. 271–276 (2012)

5. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
6. Dichtl, M., Gabriel, C., Hennerbichler, S., Seitz, B., Priglinger, S.: EU conformable eyebanking—a survey: *Eyebank linz. Spektrum der Augenheilkunde* **24**, 166–173 (2010)
7. Haykin, S.: *Neural Networks. A Comprehensive Foundation*, 2nd edn. Prentice-Hall, Upper Saddle River NJ (1999)
8. Kampik, A., Grehn, F.: *Augenärztliche Therapie*. Georg Thieme Verlag, Stuttgart (2002)
9. Langdon, W.B., Poli, R.: *Foundations of Genetic Programming*. Springer, Berlin (2002)
10. Lanzerstorfer, P., Borgmann, D., Schütz, G., Winkler, S. M., Höglinger, O., Weghuber, J.: Quantification and kinetic analysis of Grb2-EGFR interaction on micro-patterned surfaces for the characterization of EGFR-modulating substances. *PLoS One* **9**(3) (2014)
11. Lin, W., Dong, L.: Adaptive downsampling to improve image compression at low bit rates. *IEEE Trans. Image Process.* **15**, 2513–2521 (2006)
12. Lindenmair, A., Wolbank, S., Stadler, G., Meinl, A., Peterbauer-Scherb, A., Eibl, J., Polin, H., Gabriel, C., van Griensven, M., Redl, H.: Osteogenic differentiation of intact human amniotic membrane. *Biomaterials* **31**(33), 8659–8665 (2010)
13. Muresan, L., Jacak, J., Klement, E., Hesse, J., Schütz, G.J.: Microarray analysis at single molecule resolution. *IEEE Trans. Nanotechnol.* **9**, 51–58 (2010)
14. Obritzberger, L., Schaller, S., Dorfer, V., Loimayr, C., Hennerbichler, S., Winkler, S.: Identification of endothelial cell morphology in cornea using evolution strategies. In: *Proceedings of the European Modeling & Simulation Symposium* (2014)
15. Olivo-Marin, J.C.: Extraction of spots in biological images using multiscale products. *Pattern Recognit.* **35**, 1989–1996 (2002)
16. Poli, R., Langdon, W.B., McPhee, N.F.: *A Field Guide to Genetic Programming*. Lulu.com, (2008)
17. Rechenberg, I.: *Evolutionsstrategie*. Friedrich Frommann Verlag (1973)
18. Rodgers, J.L., Nicewander, W.A.: Thirteen ways to look at the correlation coefficient. *Am. Stat.* **42**, 59–66 (1988)
19. Rosse, W.F.: Paroxysmal nocturnal hemoglobinuria. *Curr. Top. Microbiol. Immunol.* **178**, 163–173 (1992)
20. Schaller, S., Jacak, J., Gschwandtner, D., Bettelheim, P., Winkler, S.M.: Identification of PNH affected cells by classifying motion characteristics of single molecules. *Proceedings of the International Workshop on Innovative Simulation for Health Care IWISH 2013*, pp. 52–57 (2013)
21. Schwarzenbacher, M., Kaltenbrunner, M., Hesch, M.B.C., Paster, W., Weghuber, J., Heise, B., Sonnleitner, A., Stockinger, H., Schütz, G.: Micropatterning for quantitative analyses of protein-protein interactions in living cells. *Nat. Methods* **5**, 1053–1060 (2008)
22. Schwefel, H.-P.: *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*. Birkhäuser, Basel, Switzerland (1994)
23. Vapnik, V.: *Statistical Learning Theory*. Wiley, New York (1998)
24. Wagner, S., Kronberger, G., Beham, A., Kommenda, M., Scheibenpflug, A., Pitzer, E., Vonolfen, S., Kofler, M., Winkler, S. M., Dorfer, V., Affenzeller, M.: *Advanced Methods and Applications in Computational Intelligence. Chapter Architecture and Design of the HeuristicLab Optimization Environment. Topics in Intelligent Engineering and Informatics*. pp. 197–261. Springer (2014)
25. Weghuber, J., Brameshuber, M., Sunzenauer, S., Lehner, M., Paar, C., Haselgrübler, T., Schwarzenbacher, M., Kaltenbrunner, M., Hesch, C., Paster, W., Heise, B., Sonnleitner, A., Stockinger, H., Schütz, G.J.: *Methods Enzymol.* **472**, 133–151 (2010)
26. Wieser, S., Schütz, G.J.: Tracking single molecules in the live cell plasma membrane—do’s and don’t’s. *Methods* **46**, 131–140 (2008)
27. Winkler, S. M.: *Evolutionary System Identification: Modern Concepts and Practical Applications*. Schriften der Johannes Kepler Universität Linz. Universitätsverlag Rudolf Trauner (2009)
28. Winkler, S. M., Schaller, S., Borgmann, D., Obritzberger, L., Dorfer, V., Affenzeller, M., Jacak, J., Weghuber, J.: Identification and classification of objects and motions in microscopy images

- of biological samples using heuristic algorithms. In: Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014, South Kuta, Indonesia, 10th–12th February, pp. 89–90 (2014). ISBN 978-0-9924518-0-6
29. Wolbank, S., Hildner, F., Redl, H., van Griensven, M., Gabriel, C., Hennerbichler, S.: Impact of human amniotic membrane preparation on release of angiogenic factors. *J. Tissue Eng. Regen. Med.* **3**(8), 651–654 (2009)

Part II
Communications, Networks,
and Cloud Computing

Chapter 9

Comparison of TCP/IP Routing Versus OpenFlow Table and Implementation of Intelligent Computational Model to Provide Autonomous Behavior

Ameen Banjar, Pakawat Papatwibul and Robin Braun

Abstract Software-Defined Networking (SDN) is an emerging programmable network architecture, where network control plane is decoupled from forwarding plane. The first standardize communication interface defined between the controls and forwarding layers of the SDN architecture is known as OpenFlow. OpenFlow is a key enabler for SDN that allows direct manipulation on the forwarding plane of network devices. SDN forwarding methods are based on flows, through a protocol like OpenFlow, which operates in contrast to conventional networking device methods, such as TCP/IP routing table and MAC learning table. In more details, OpenFlow protocol has the same forwarding methods to push L2-L4 functions which are simplified into a Flow-Table(s). This paper discusses the relationship between the processes of forwarding packets in conventional IP routing table versus OpenFlow-table. Then, the paper proposes the three phases of implementing a Distributed Active Information Model (DAIM) within OpenFlow to support an autonomic network management.

9.1 Introduction

The business and modern needs force network technologies to be increased in performance, complexity, and functionality. Moreover, current network paradigms are lack of adaptability, and limited to a single domain management, which is managed by network operators [17]. Also, networks have become massive and intractable

A. Banjar (✉) · P. Papatwibul · R. Braun
Centre for Real-Time Information Networks (CRIN), University of Technology
Sydney (UTS), 15 Broadway, Ultimo, Sydney, NSW 2007, Australia
e-mail: 11311103@uts.edu.au

P. Papatwibul
e-mail: 10926297@uts.edu.au

R. Braun
e-mail: robin.braun@uts.edu.au

due to complexity and that lead to challenges of scalability. Traditional operations struggle to cope, and thus a new management paradigm is required to fulfill the management of such dynamic infrastructures. Many attempts are introduced to cope with the higher demand on network management problems, but still not suitable for large scale networks and could suffer from scalability issues [19].

The Transmission Control Protocol (TCP) and the Internet Protocol (IP), commonly known as the TCP/IP standard, is widely used for network communications. TCP/IP attempts to make efficient use of the underlying network resources, by specifying how data should be transmitted, formatted, addressed, routed and received at the destination [6]. TCP/IP was developed to support maximum throughput over many kinds of different networks.

Although TCP/IP supports many current network services, it is not efficient for the requirements of business needs and end users. This has led to the development of alternative networking architectures, and the introduction of Software-Defined Networking (SDN).

SDN is a relatively advanced method for implementing communication networks. SDN separates decision maker, called control plane, which decides where packets are sent, from underlying infrastructure, called data plane, which forwards packets to the decided destination(s) [16]. The SDN approach uses intelligent abstractions to simplify the network including (1) Using a common API for programming network infrastructure; (2) Using a single state distribution algorithm for networking; and (3) global management where programs interact with the entire network instead of individual nodes [23]. A newly emerging standard for SDN is the OpenFlow protocol, which includes a standardized protocol for communications between the control plane and the data plane.

However, OpenFlow is still not widely standardised yet because of that in each year Open Network Foundation (ONF) introduce newly improved versions and extra functionalities. The latest version is known as OpenFlow specification 1.4 [15]. The evolution of OpenFlow protocol versions are kept experimenting within labs. This chapter compares the performance of TCP/IP with the newly emerging OpenFlow standard for software defined networking.

A variety of SDN approaches have been developed, but there is only limited information on the performance of each, and realistic performance comparisons are not widely available. Simulation tools such as OMNeT++ and INET Framework are suitable for tasks of designing, building, and testing network architectures, and provide practical feedback when developing real world systems [4, 5]. Such simulation tools allow system designers to determine the correctness and efficiency of a design before a system is deployed. Simulators also enable an evaluation of effects on various network metrics, and provide mechanisms to obtain results that are not experimentally measurable on larger geographically distributed architectures. However, very few performance evaluations of OpenFlow architectures using simulation tools such as OMNeT++ INET Framework have been published [11].

The performance comparison simulates OpenFlow and TCP/IP networks using the OMNeT++ INET Framework discrete events network simulator. By analyzing key network metrics including round-trip-time (RTT) and data transfer rate (DTR), the

results indicate that OpenFlow performed slightly better than TCP/IP in this analysis. The results also proved the correctness of OpenFlow implemented simulation model.

Thus, this chapter evaluates the performance of TCP and OpenFlow implementation in INET framework 2.0 using OMNeT++. In addition, the paper covers a number of multiple running from the simulation to preserve accuracy of the simulation results. The remainder of the chapter is organized as follows. In Sect. 9.2 we introduce the background and overview of TCP/IP and OpenFlow networks. In Sect. 9.3 we demonstrate the advantages and disadvantages from these two different network paradigms. We present more details of the fundamental features of IP routing and OpenFlow protocols in Sect. 9.4. Section 9.5, the performance of TCP/IP versus OpenFlow is evaluated in the simulations. Section 9.6 proposes the DAIM model within OpenFlow to support an autonomic network management. Finally, Sect. 9.7 concludes this work.

9.2 Background of TCP/IP and OpenFlow Network

This section introduces Transmission Control Protocol/Internet Protocol (TCP/IP) and provides a quick introduction to OpenFlow-Based SDN, discussing the architecture and how they operate in an overall context. We begin by defining both TCP/IP and OpenFlow network in the most general terms.

9.2.1 Overview of TCP/IP

The Transmission Control Protocol and Internet Protocol (TCP/IP) is a suite of communication protocols widely used to connect hosts on the Internet and on most other network communications as well [13]. TCP operates at the transport layer, which is the middle layer in the seven layers of OSI model [13]. This layer maintains reliable end-to-end communications between network devices. On the other hand, IP is a network layer protocol, which is responsible for packet forwarding including routing across intermediate routers (see Table 9.1). Because TCP/IP was developed earlier than the OSI 7-layer mode, it does not have 7 layers but only 4 layers [9].

Table 9.1 TCP/IP protocol layers

TCP/IP protocol suite	TCP/IP
FTP, SMTP, Telnet, HTTP,	Application
TCP, UDP	Transport
IP, ARP, ICMP	Internet
Network interface	Network access

One fundamental feature of the IP protocol is that it only deals with packets, addresses, and directing messages to where they are intended [21]. This is the most significant unit of TCP/IP data transmission. TCP allows two devices to complete a connection and exchange streams of data. TCP assures that the data and packets will be delivered to the destination in the same order in which they were sent [9].

Like the OSI model, functionalities of TCP/IP has been organised into four abstraction layers. (1) Network Access layer contains the network interface card which provides access to the physical devices. (2) Internet layer establishes network to network communications and therefore connects internet working. (3) Transport layer handles the end-to-end (host-to-host) communication. (4) Application layer offers the users with an interface to communicate and gives a way for applications to have access to network services [6, 16].

TCP/IP is a set of protocols developed to allow cooperating computers to share resources across a network and also to ensure networks robustness by recovering automatically from any failure of nodes on the network [6, 9, 21]. Furthermore, it can allow large scaled networks to be contracted with minimal requirements of central management.

9.2.2 Overview of OpenFlow-Based SDN

Software-defined networking (SDN) is a relatively advanced method for implementing communication networks [23]. SDN separates the decision maker, called the control plane, which decides where packets are sent, from the underlying infrastructure, called the data plane, which forwards packets to the decided destination [19]. This migration of control can be formerly and tightly bound in individual network devices enabling the underlying infrastructure to be abstracted for applications and network services, which can treat the network as a logical or virtual entity. In SDN, OpenFlow is considered as the first standardised protocol, which defines for communication between the control plane and the data plane [15, 22].

OpenFlow was initially introduced by Stanford University in 2008, as the first standardised communication interface defined between the control plane and the data plane of the SDN architecture [16]. Figure 9.1 explains OpenFlow as an open standard that enables researchers to run experimental protocols in the networks without having to expose vendors internal implementations of their network devices [14]. In classical switches and routers, the fast packet forwarding (data plane) and the high-level routing decision (control plane) happen in the same network element, whereas in OpenFlow they separate these two functions [24]. In OpenFlow, it separates these two functions. The data plane portion still resides in the switches, whereas the routing decisions are moved to a different device called the controller, typically a standard server. The controller communicates with an OpenFlow switch through a Transport Layer Security (TLS) and its predecessor, Secure Sockets Layer (SSL) channel using the OpenFlow protocol [22].

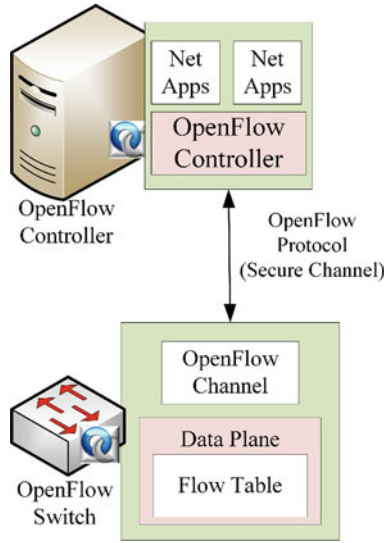


Fig. 9.1 OpenFlow-based SDN structure

9.3 Advantages and Disadvantages of TCP/IP and OpenFlow

In this section, some of TCP/IP and OpenFlow advantages and benefits, as well as some of their disadvantages and costs are described.

9.3.1 Advantages of TCP/IP

TCP/IP networks can offer many advantages in terms of performing fast packet forwarding. This section takes a look at the benefits of TCP/IP components. These two separate protocols, TCP and IP, work hand-in-hand to perform chores that control and direct the general mobility of data packets over the network. TCP/IP uses specific headers in order to define each packet’s header and payload. If there is more than one packet, it should know how many others should be expected. TCP is more concerned with making the connections to the remote hosts. IP, on the other hand, focuses on addressing so that messages are guaranteed that they will be delivered to the final destination. The development of TCP/IP has been a great success due to the result of the advantages they offer over other network protocols and protocol suites, including the following [6, 9, 21]:

- (1) TCP/IP is an opened and widely accessible protocol that is not owned by a single company as a secret protocol. This gives the opportunity for anyone with advance technical skills to improve and optimize it.

- (2) Most of the modern operating systems are virtually compatible with TCP/IP. As a result, almost any network systems can communicate with any other systems. Moreover, TCP/IP is also virtually compatible with every type of network configurations and computer hardware.
- (3) It provides efficient routing protocols, where routing decisions are determined by the best and most effective path for all packets as they traverse across the network. Therefore, this makes TCP/IP highly scalable with virtually unlimited size of the network.
- (4) TCP/IP can offer reliable data delivery. Reliability is by means of assuring that the data is delivered to its intended final destination. For example, through the process of error detection, and retransmission of missing and corrupted packets.
- (5) TCP/IP uses a relatively simple addressing principle known as IP addressing, which allows network operators to transfer their knowledge of TCP/IP to other TCP/IP networks that do not require the learning of a new addressing scheme.

9.3.2 Disadvantages of TCP/IP

9.3.2.1 Not an Operating System

Although TCP/IP supports many current network services, it is not an operating system (OS). An OS is the software that runs on a computer and creates the environment in which other applications can run. The various Windows platforms are examples of operating systems. When installed, TCP/IP becomes an important extension of your computer's operating system, but it does not replace it. This has led to the development of alternative networking architectures.

9.3.2.2 SYN Attacks

Another important principle of TCP/IP is using a three-way handshake to establish communication. By doing this the user connects to a server, and then sends back a reply to the user to confirm the connectivity. The host should then wait 75 s for the connection in order to receive a reply to the acknowledgment. By forwarding the first request but not the second request repeatedly is causing SYN attacks. This may result in controlling the computer's resources and affecting other legitimate connections to fail.

9.3.2.3 Sequence Manipulation

The Internet Protocol forwards packets where each includes a sequence number. In case of a user goes offline, any hosts that are currently connected to that user will close the communication if a reply is received without a proper sequential number.

If there are other users that can intercept these messages, they can recognize the ID number and try to respond as the actual user. This can cause significant security issues such as computer being hijacked or lost of important data.

9.3.2.4 Amorphous Identification

TCP/IP was developed with a function that only determines one connection at a time, and has this connection forward to a certain user. Moreover, the Internet has progressed worldwide making the IP address space being highly valuable. However, this IP address space is not dedicated anymore and is normally assigned within a group of users. The lack of unique identification can cause difficulties in verifying a user. The complexity of programs security also increases because trust in the identity of users is removed.

9.3.2.5 DNS Flaws

The Domain Name Service (DNS) provides a way to obtain a domain name by looking up IP addresses. IP protocol of the TCP/IP suite can only communicate to IP addresses, which needs Internet connection to take place through a DNS server maintained by the Internet Service Provider (ISP). One way to improve this flaw is to redirect the domain names to different IP addresses.

9.3.3 Advantages of OpenFlow

By using OpenFlow-based SDN technologies, it is possible for IT to address the high demand of business needs, dynamic nature of today's services and applications, adaptation to the ever changing network environments, and consequently reduce the complexity of network management and operations. OpenFlow-based SDN can bring several advantages to carriers and enterprises by modifying the network architecture in such way including [15, 22, 24]:

- (1) One of the greatest benefits of OpenFlow is creating flexibility in networks on how it will be use, operate and sold. The control program specifies the behaviour on the abstract model and configures the abstract network view, which can be written by network operators and administrators using common programming environments.
- (2) OpenFlow-based SDN can introduce rapid innovation through customization because network administrators can develop the features they need in the control program, instead of having to wait for software vendors to place in plan for their trademark products.
- (3) OpenFlow can also reduce the operating expenses because of centralized and automated management of network devices, uniform policy enforcement, and thus result in less network downtime and fewer configuration errors.

- (4) It enables network virtualization and network integration of computing and storage. This makes all the IT operations to be managed more easily by using a single viewpoint and toolset.
- (5) It is easy to integrate OpenFlow-based SDN with computing for resource management and maintenance.
- (6) OpenFlow increases the automation of management by using common APIs to abstract the underlying networking details from provisioning systems and applications to meet business objectives.
- (7) OpenFlow-based SDN provides a standard way to convey Flow-Table information to network devices, and therefore fosters open and multiple vendor markets.

9.3.4 Disadvantages of OpenFlow

9.3.4.1 Single Domain of Management

SDN is managed by a centralized domain and that leads to have a point of failure. Many solutions can be applied such as double the controller as a backup and that could prevent the failure of the network if one controller fails to functions. Also, other solution can be having a distributed controllers where that is possible if we have a logically centralized but physically distributed (can be distributed network applications, or just synchronise events between distributed controllers). However, many research projects have done that where they still in research phase where such a solution has more complexity on the current OpenFlow architecture.

9.3.4.2 Bottleneck of Request and Response Times

Conventional networks are able to handle events locally and do action according to those events. Whereas, in SDN architecture the forwarding plan of each device are not able to handle some events which need to be sent to the controller to manage that events. The controller received those events and calculates them and sends commands back to the corresponding switch. In this case, if sending many requests from massive switches to a single controller that could lead to bottleneck. Also, same thing can happen if many commands of setting many flow rules to one certain switch.

9.3.4.3 Scalability Issues with Centralized Controller

Single domain lead to scalability issues where the computing power is loaded to a single point also the distance lead to scalability issues where the Round Trip Time RTT for requesting unknown packets to be processed.

9.3.4.4 Enforcing QoS Is Difficult (OpenFlow Specific)

Dynamic adaptation based on OpenFlow protocol is not provide sufficient quality of service QoS, because the main focus of OpenFlow is the static configuration. As structure of OpenFlow the request queues processed within switches and that is outside of the protocol.

9.4 Functionalities of IP Routing and OpenFlow Protocol

SDN forwarding based on flow, through a protocol like OpenFlow, which embedded flow table(s) that simplified theory of methods to push L2-L4 of number of cases. OpenFlow network allows for more then L3 prefixes of traditional IP routed IPv4. OpenFlow able to mean matching exact values in fields across L2-Data (Ethernet), L3-Network (IP), L4-Transport (TCP/UDP) ports [7, 22].

Conventionally the method of forwarding uses vendors proprietary protocols based on Layer 3 prefixes routing. While, OpenFlow currently is a popular method which instantiates forwarding rules into network forwarding objects over TCP. OpenFlow offers low level network instruction set to the vendors then they can write controllers and applications that can control forwarding and extract valuable information about the state of the network [8, 24].

Statically programming flows by hand crafting in anywhere rather that a lab could be unmanageable include not only L3 forwarding but also L1-L4 forwarding. Operators setting up static flows for traffic forwarding can be unmanageable as throwing out IGP/EGP for dynamic exchange RIB state. Moreover, hand crafting static flows within OpenFlow is less scalable than operators how use static routs even less scalable than operators who not use IGP/EGP [8, 24].

Network had unique challenges in the past that warranted this use of dynamic change of a routing Information Base (RIB) state. The Operators used to use static routing in case of different networks to be communicated. Table 9.2 shows the routing

Table 9.2 Classification of routing protocols

Routing protocol type	Administrative distance	Example	Functionality
IGP (Interior gateway protocol)	Distance vector	Rip VI, Rip V2	Care about distance between routers and number of routers
	Link state	OSPF, IS-IS	Care about router state and speeds
	Advanced distance vector	EIGRP	Care about both of above functionalities
EGP (Exterior gateway protocol)	Path vector	BGP	Mostly used for Internet connections based on autonomous system (AS)

protocols in case of dynamic and different network to exchange network information. For example, a forwarding path from A to Z, operators should be able to push a policy from A to Z rather than set them statically A, B, C, D, and Z and operators should have the details of everything in between implemented without dealing with the irrelevant information to accomplish that policy [8, 23].

The main driven of the network are layers architectures, especially layer 2 up to layer 4. Layers have fundamentals information for routing and forwarding flows as an example Ethernet is layer 2, IP is layer 3 and Transport is layer 4, each has header. However, Protocols among these layers are processed and invented while the undoubtedly for networking are layers. Moreover, transferring data from end to end nodes use encapsulation, which is not encapsulating the frame, but it is inserting extra bits field on each layer. Most operators focused on IGP/EGP, framing and encapsulations for data interconnections, the following Fig. 9.2 illustrates an example of IP header and Wireshark captures [9, 23]. SDN approach change the focus of IGPs or other and came up with idea of flow based forwarding, then it is required to understand how layers are interacting with each other.

The following section below will introduce in more details connection setup and messages of OpenFlow.

9.4.1 Connection Setup

Figure 9.3 shows the connection establishment process of OpenFlow, where switches initiates a secure TCP channel to controller, that channel enable the controller to manipulate switches via OpenFlow protocol. Starting secure channel is required controllers IP address. The requirement has fulfilled through three ways hand check as you can see on the previous figure. After that the controller can detect all connected switches and initiate the connection. Firstly, upon the secure connection started both controller and switches corresponding hello messages to get OpenFlows version from each other. Secondly, if the OpenFlow versions are compatible in both nodes, then the controller send a message to know the capabilities of connected switch through features request message. Finally, after the switches received requested features message, they reply by message to inform the controller of supported functionalities [11, 15].

9.4.2 OpenFlow Messages

OpenFlow protocol is implementing some messages as shown in the following Table 9.3. The table shows few examples of OpenFlow messages and Wireshark captures of one message type which is Packet_Out message. The OpenFlow messages consists of the Features_Request and the Features_Reply, which are required for the set up, establish connection of secure TCP channel between OpenFlow switch and controller. The Packet_In message from switch to controller is used

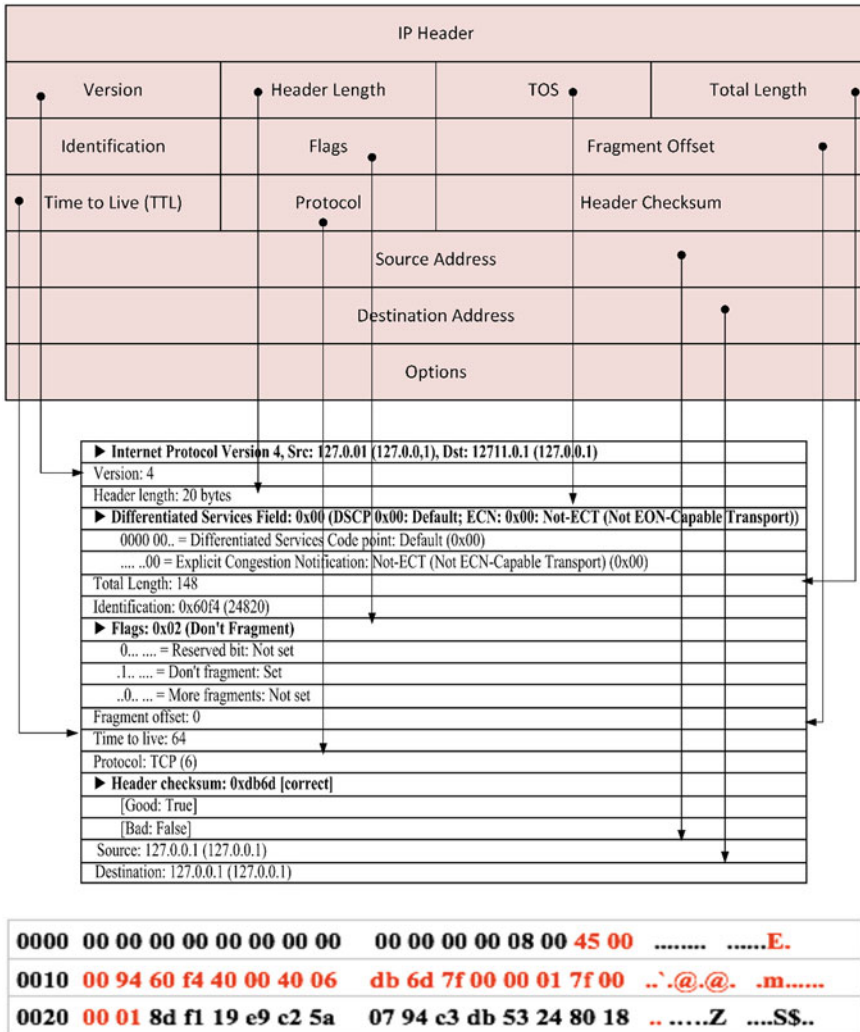


Fig. 9.2 Layer 3 IP packet header with Wireshark captures

when no matching rules are found to inform the controller about an unmatched packet or to send a packet to the controller if the associated action processed that.

Also the table has the Packet_Out message from controller to switch which is used to send a packet out from controller after calculate path then send it to the specific port at the switch. Lastly, Flow_Mod message from controller to switch, which enabled the controller to manage the forwarding table at the switch, embedded with packet matching rules fields and actions [11, 15].

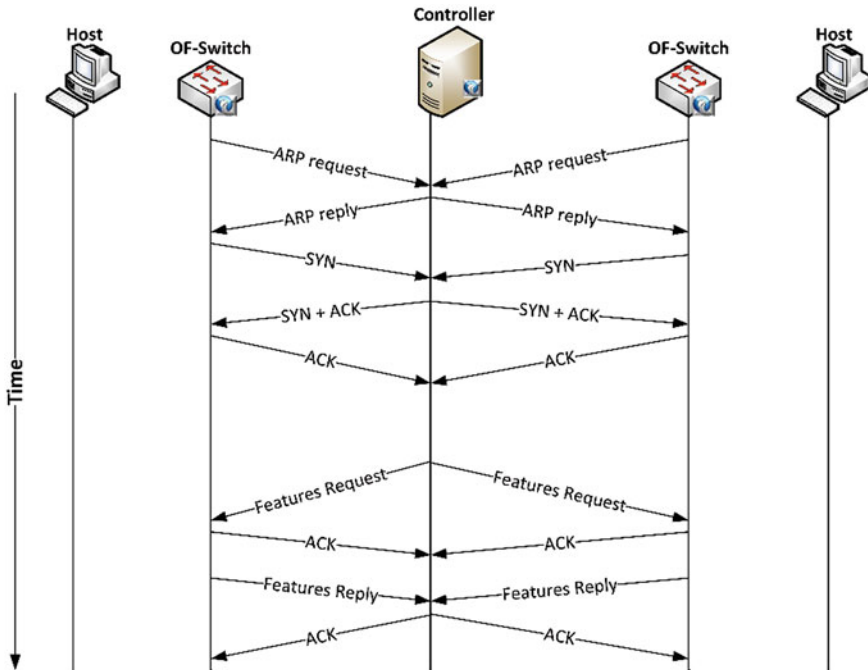


Fig. 9.3 Initialize communication set up

9.5 Comparison of Performance Results

This section provides a performance comparison of TCP/IP versus OpenFlow modules in INET Framework 2.0 using OMNeT++ network simulation. The network simulator makes it possible to evaluate how a network performs under circumstances with different versions of the protocol stack, and enables analysis of the effects of different variables including channel speed and delay [1, 12].

9.5.1 Measurement Methodology

The OMNeT++ INET Framework 2.0 network simulator is a C++ discrete event simulator [3, 4]. An advantage of this simulator is that it simplifies the integration of new modules, and allows existing modules to be customized. The INET Framework is a network simulation package that contains models for wired and wireless networking protocols, including UDP, TCP, SCTP, IP, IPv6, and Ethernet. The INET Framework has recently implemented an extension to enable OpenFlow to be modeled. The OpenFlow extension is still in early development, and is currently based on Switch Specification Version 1.2 [11].

Table 9.3 Wireshark capture and OpenFlow messages

Type	Description
<input type="checkbox"/> Packet-In [Switch > Controller]	When no matching rules found for a received packet on switch's flow table, then packet is forward to the controller
<input type="checkbox"/> Packet-Out [Controller > Switch]	Controller calculates rules and send a packet out to switch port(s)
<input type="checkbox"/> Flow-Mod (Controller > Switch)	Switch add a flow entry to its flow table, which is instructed by the controller
<input type="checkbox"/> Flow-Expired [Switch > Controller]	When the idle timeout comes up because of inactivity of flow entry.

▶ OpenFlow Protocol
▶ Header
Version: 0x01
Type: Packet Out (CSM)(13)
Length: 102
Transaction ID: 0
▶ Packet Out
Buffer ID: 257
Frame Recv Port: 2
Size of action array in bytes: 8
▶ Output Action(s)
▶ Action
Type: Output to switch port (0)
Len: 8
Output port Flood (all physical ports except input port and those disabled by STP)
Max Bytes to Send: 0
of Actions: 1

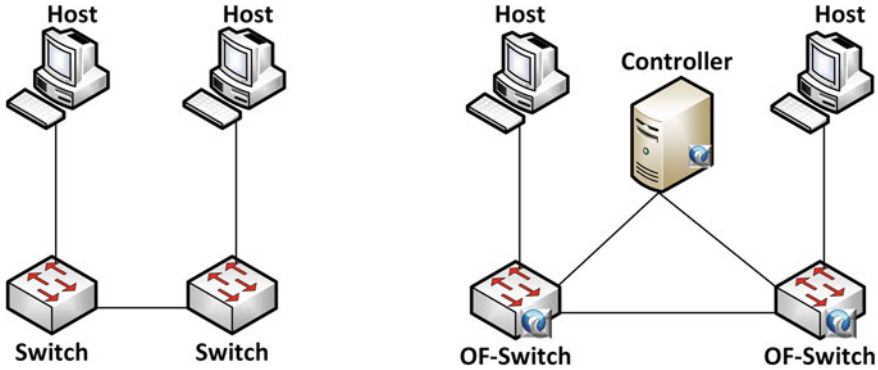


Fig. 9.4 Layout of the simulated networks

The OMNet++ network simulator is used here to simulate the operation of OpenFlow and TCP/IP while logging performance metrics including Data Transmission Rate (DTR) and the mean round-trip-time (RTT) for nodes in the simulated networks [3]. We analyse the DTR and RTT for the nodes within TCP/IP and OpenFlow networks, using a similar network topology for each. Each network includes a number of hosts, two switches and a destination server (see Fig. 9.4). The OpenFlow network also includes an additional device called the controller (standard server), which is directly connected via separate links to the OpenFlow switches. Hence the OpenFlow switches can perform Layer 2, 3, and 4 routing, as compared to the Layer 2 MAC learning table used by TCP/IP. The simulations are logged, and the logs are subsequently analysed to enable the performance of the OpenFlow and TCP/IP networks to be compared [1].

Performance is compared by varying the traffic and link delays by the same amounts in both the OpenFlow and the TCP/IP simulation channels. In this experiment, we measure the DTR and the mean values of measured RTT between the time-triggered nodes, both for the standard TCP/IP in INET, and OpenFlow. Sequential ping requests are generated, where each includes a sequence number, and replies are expected to arrive in the same order. After the simulation has run for 300 s, we measure the RTT for each ping and reply. This results in 48 measurements each for OpenFlow and TCP/IP, for each 300 s simulation period. Each simulation was run ten times (OpenFlow and TCP/IP five times each) to reduce simulation artifacts. A large number of samples are recorded, and the means and standard deviations of DTR and RTT are computed. RTT is approximated using following equations [10]:

$$Tr = \frac{\text{packet size}}{\text{link bandwidth}} = \frac{1,500 * 8 \text{ bit}}{100 * 10^6 \text{ bit/s}} = 0.12 \text{ ms} \quad (9.1)$$

$$RTT = \alpha * RTT + (1 - \alpha) * Tr \quad (9.2)$$

The packet size is known to be 1,500 bytes long. The link speed limited to 100 Mbit/s in both the TCP/IP and the OpenFlow networks is the transmission time between the segments sent and the acknowledgment arrival, and is a smoothing factor, which equals the value (7/8) 0.875.

9.5.2 Result Evaluation

We have tested the performance by sending video streaming TCP/IP and OpenFlow in both network environments. We have increase the size of the file from 1 MB to 512 MB and measure them against the time.

Figure 9.5 shows that as file-size increases, the performance of OpenFlow is slightly better than TCP/IP, in that it can forward video streaming UDP packets at higher rates (measured in megabits per second).

Figure 9.6 presents traces of the round trip times for both TCP and OpenFlow that was used to test the various algorithms, where each client has generated the pings and logged the results. It can be seen in Fig. 9.6 that TCP/IP has longer round trip times than OpenFlow, which indicates how well the client can ping to others.

Moreover, the performance comparison show that there is a slightly difference in the average value of RTT. In TCP/IP the average RTT is just below 0.8s, whereas in OpenFlow is around 0.6s. However, it is also evident that there are sudden spikes at the beginning and at the end caused by connection establishment and termination. The middle sudden spike was caused by the expired Idle-timeout.

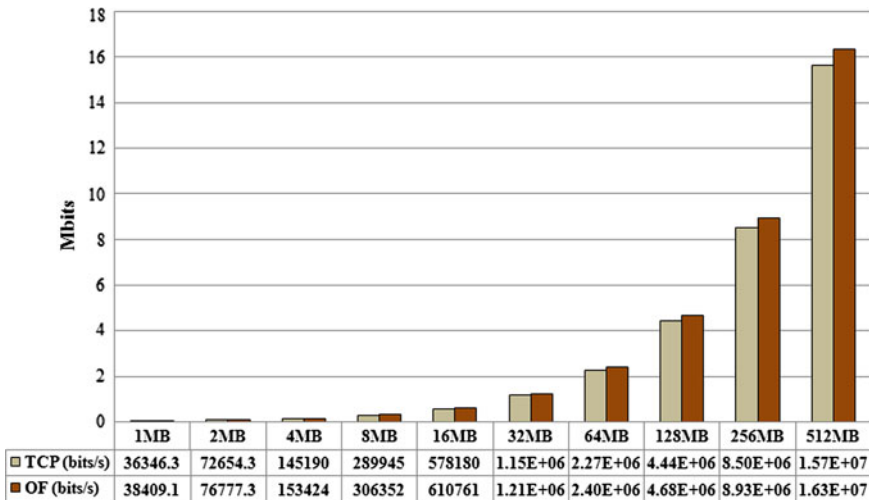


Fig. 9.5 DTR of TCP/IP versus OpenFlow

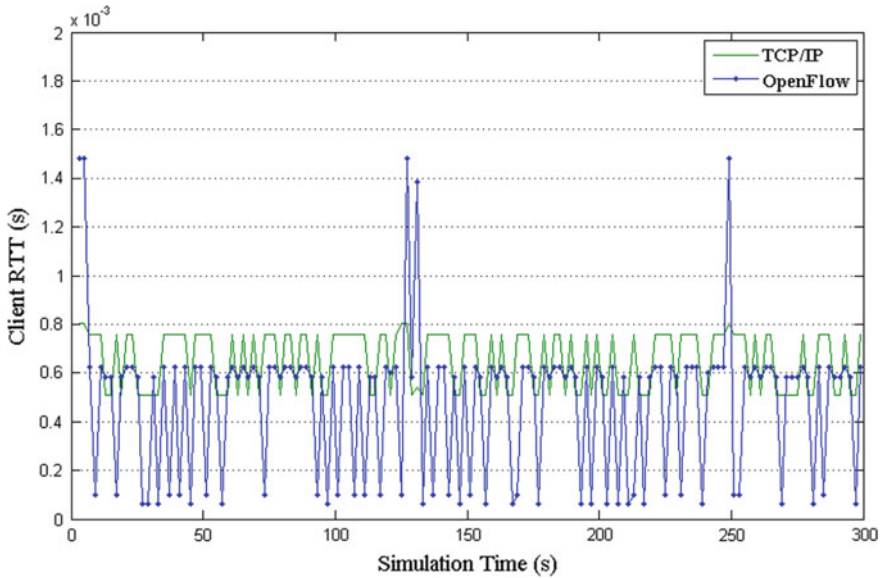


Fig. 9.6 Measured RTT of TCP/IP versus OpenFlow

The results shown in Figs. 9.5 and 9.6 indicate that OpenFlow outperforms TCP/IP in this analysis. This suggests that OpenFlow's performance benefits from the packet transmission methods run at Layers 2, 3, and 4, even though connection establishment and initialization has more steps than TCP/IP. This may confirm the assumption that OpenFlow performs slightly better than TCP/IP, with respect to RTT and DTR, and performs faster with the same circuitry and not incurring major performance losses.

9.6 Proposed DAIM Model and its Implementation

DAIM is a sustainable information model, which collects, maintains, updates and synchronizes all the related information. Moreover, the local decision making ability within each device, on the basis of collected information, allows DAIM to autonomously adapt according to the ever changing circumstances [20]. The DAIM model structure is proposed with the hope that it addresses the limitation of previous network protocols such as Simple Network Management Protocol (SNMP), Common Information Model (CIM) and Policy-Based Network Management [19].

Ultimately, the proposed DAIM model will address the limitations of current approaches and future distributed network systems, creating an autonomic computing management strategy. The DAIM model approach will also satisfy the requirements of autonomic functionality for distributed network components like self-learning, self-adaptation and self-CHOP (configuration, healing, optimization and

security). Each component can be adaptable according to any changed conditions of the dynamic environment without human intervention. This chapter introduces the three phases of implementing DAIM model as follows [19]:

9.6.1 Phase 1: Embedded Multi-Agent System in OpenFlow Switch (Managing Flow Tables)

In phase one, we have embedded the DAIM cloud by implementing a JVM environment in the OpenFlow switches, which include a multi-agent operating system like JadeX. Essentially it describes a distributed environment where the network information is the property of software agents residing in virtual machines distributed throughout the network elements. This JadeX environment can create, change and terminate DAIM agents. The architecture of OpenFlow will be the same and all the high-level routing decisions still be made from the controller, but will have the DAIM agents doing the work of carrying all OpenFlow messages from controller to switch and viscera. In this case, agents will own the values (variables) and properties of the flow entry. Where DAIM agent are able to operate get and set process driven by the controller as well as updating the forwarding table [19, 20].

9.6.2 Phase 2: Semi-Distributed Functionality

Phase two architecture is similar to the phase one, which include a multi-agent operating system like JadeX. Distributed agent resides on each network node and each network component (e.g., device components, service) operates with its own functions. The main difference of this phase is the controller can assign some functionality to DAIM agents (see Fig. 9.7). Now, the DAIM agents will not only be the carrier for OpenFlow messages between switches and controller, but also can gather information from controller attempting to perform some functions. For example, collecting RIB information from the controller, so it is possible to forward flows directly from switch according to the flow entries, which prefilled by DAIM agents. In more details, DAIM agents are able to update forwarding tables on the switches and can execute some functionality across the OpenFlow controller and switches. The controller uses a separate control channel to exchange information between network nodes by using DAIM agents. The network nodes listen to each other on a particular port for messages from the DAIM agents. Another example of DAIM agents functionality is performing a certain function for querying network statistics. Moreover, DAIM agents are able to inform the controller that a network problem is happening and also sends acknowledgment of the received messages. In all cases, DAIM cloud will trigger the JVM to instantiate a new unique agent in order to process those functions [17, 18].

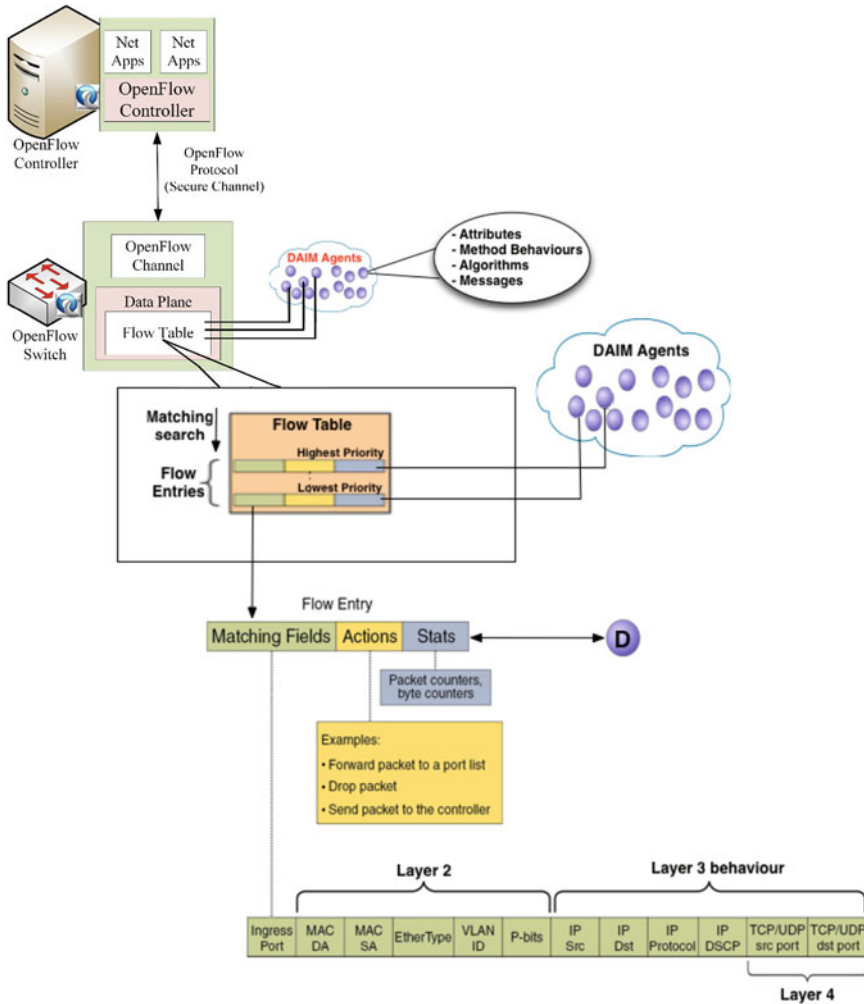


Fig. 9.7 DAIM approach include phase one and phase two

9.6.3 Phase 3: Fully Autonomous Functionality

Third phase is concerning on migrating all computational power to the DAIM agents which located within each switch to have autonomous behaviour of the network. System Requirement Database SRD and Control Database CD support these DAIM agents.

To understand this phase the chapter focused on migration of the computational power and sending packets by DAIM agents. The new architecture include DAIM cloud within each switch supported by databases so, each database actively synchronizes with others according to the events registered. DAIM cloud can publish

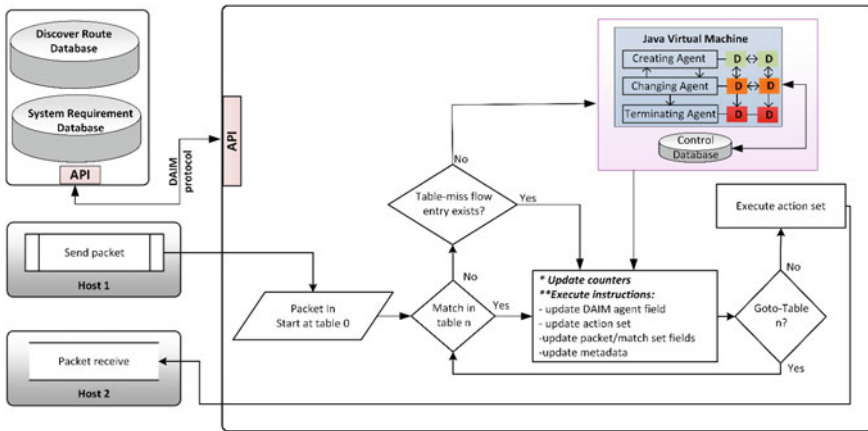


Fig. 9.8 DAIM ultimate approach for phase three

events and actively synchronize to databases, so that other switches can reconstruct the global network information. Moreover, individual switch can serve any coming packets locally or from other switches. While, self-configuration and self-adaptation will be applied if any local change happens as a benefit of database. Thus, DAIM cloud has the ability to synchronize information of the entire network and it has feasibility to be deployed in distributed system structures. Then, it could be possible to achieve that self-management if all autonomic functions are enabled [2].

When the packet hits the OpenFlow switch, it performs the operations shown in Fig. 9.8. Packet headers are used for table look-ups depending on the packet type, and typically include various packet match fields, such as Source IP, Destination IP, and MAC destination address. The switch begins with performing a table look-up in the first flow table, and may perform table look-ups in other flow tables [15, 16]. For example, the flow tables are sequentially numbered, so the packet is matched against flow entries of flow table 0. Other flow tables may be used depending on the outcome of the match in the table 0. If a flow entry is matched, the instruction set included in that flow entry is executed and the counters associated with the selected flow entry must be updated. Those instructions may direct the packet to another flow table, where the same process is repeated again. On the other hand, the instructions could forward the packet to Table-miss or the DAIM cloud in case of flow entry not matched.

In case of no matching within current flow entries for that unknown packet, then it be sent to Table-miss flow entry. The Table-miss entry holding old matching rules which could be a wild-card rules or specific table configuration for unknown packet. The Table-miss flow entry can process for unmatched packets by information from current flow entries. The third phase integrated DAIM cloud to support Table-miss flow entry, where it does not exist by default in a flow table, the DAIM cloud can add Table-miss flow entry or remove it. However, DAIM cloud is responsible of unmatched packets, if there is no Table-miss flow entry.

The autonomous DAIM cloud has a multi-agent operating system that can create, change, and terminate the intelligent DAIM agents named JADE/JADEx (Java Agent Development Framework). The software agents not only have responsibility for maintaining their own property (value) but also for adapting and modifying them as well. DAIM agents can make local decisions based on the system requirements and according to the collected information. Moreover, DAIM agents will also compute the forwarding destination for packets, exchanging data between agents, and update the flow table row using its built in methods [2, 17, 18].

Furthermore, business needs drive DAIM agents behaviours which being bounded to a particular variable such as flow entry variables and have some level of self-adaptation strategy to manage the variables for forwarding. The properties values are similarity notions of object-oriented programming. Thus, ingredients of DAIM agents are able to implement autonomic behaviours. For example, when the DAIM cloud receives the unmatched packet, then it creates agents that able to access and control network elements. Agents compute a path for the unmatched packet, and install flow entries on every switch along to the chosen path to forward that flow (see Fig. 9.8). For more details, DAIM agents are able to check this flow against system requirements and other policies to see whether it is allowed to be processed.

The DAIM agents provide a distributed environment where the network information is the property values of agents residing in virtual machines that are distributed throughout the network elements.

9.7 Conclusion and Future Works

In this chapter, we have introduced a thorough overview of TCP/IP and OpenFlow networks including the background, behaviour and architecture. Moreover, the advantages and disadvantages of TCP/IP and OpenFlow have been described to compare and contrast both the strengths and limitations of each network. We also presented more details of IP routing and its functionalities as well as how OpenFlow communicates with the switch via the OpenFlow protocol. The performance of TCP/IP protocol suite in contrast with OpenFlow is also evaluated using network metrics such as Data Transmission Rate and Round-Trip-Time in the topologies of our study. It has been evident according to the measurement outcomes that OpenFlow performs slightly more effective than TCP/IP with lower RTT values and can send streaming UDP packets at a higher rate (Mbps). Lastly, the proposed DAIM and its three implementation phases have been introduced with the hope to resolve the scalability issues and develop autonomic behaviours in OpenFlow.

For future works, we aim to implement the first phase of the DAIM cloud and extend OpenFlow structure based on intelligent agents to exchange information and install forwarding flow tables, which can be used in other distributed computing environment. Therefore, it could be applied to many different environments such as large data centers and road traffic systems.

Acknowledgments This work is sponsored by the Centre for Real-Time Information Networks (CRIN) in the Faculty of Engineering and Information Technology at the University of Technology, Sydney (UTS). This paper is an extended version of the ACASE14 conference paper [1].

References

1. Banjar, A., Papatwibul, P., Braun, R., Moulton, B.: Analysing the performance of the OpenFlow standard for software-defined networking using the OMNeT++ network simulator. In: 2nd Asia-Pacific Conference on Computer Aided System Engineering (APCASE), pp. 36–37 (2014)
2. Banjar, A., Papatwibul, P., Braun, R.: DAIM: A mechanism to distribute control functions within Openflow switches. *J. Netw.* **9**, 1–9 (2014)
3. Varga, A., Hornig, R.: An overview of the OMNeT++ simulation environment. In: Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops, p. 60 (2008)
4. Varga, A.: INET Framework for the OMNeT++ Discrete Event Simulator (2012). (Online) Available at: <http://github.com/inet-framework/inet>
5. Varga, A.: The OMNeT++ discrete event simulation system. In: Proceedings of the European Simulation Multiconference (ESM), p. 185 (2001)
6. Forouzan, B.A.: TCP/IP Protocol Suite. McGraw-Hill, New York (2002)
7. Salisbur, B.: OpenFlow: Coarse vs. Fine Flows (2013) (Online). Available at: <http://networkstatic.net/openflow-coarse-vs-fine-flows>
8. Salisbur, B.: SDN OpenFlow Policy Abstractions (2013) (Online). Available at: <http://networkstatic.net/sdn-openflow-policy-abstractions>
9. Comer, D.E.: Internetworking with TCP IP/1: Principles, Protocols and Architecture. Prentice Hall **1**, 77–88 (1999)
10. Snnen, D.: Performance Evaluation of OpenFlow Switches, Semester Thesis at the Department of Information Technology and Electrical Engineering (2011)
11. Klein, D., Jarschel, M.: An OpenFlow Extension for the OMNeT++ INET Framework. OMNeT++, Cannes France, (2013)
12. Jeroen, I., Heijenk, G., de Boer, P.T.: TCP/IP modelling in OMNeT++. University of Twente, The Netherlands, (2004)
13. Fall, K., Floyd, S.: Simulation-based comparisons of Tahoe, Reno and SACK TCP. In: ACM SIGCOMM Computer Communication Review, vol. 26, pp. 5–21 (1996)
14. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: Enabling Innovation in Campus Networks. In: SIGCOMM Computer Communication Review, vol. 38, pp. 69–74 (2008)
15. Open Network Foundation, OpenFlow switch specification version 1.3.0 (wire protocol 0x04), June (2012) (Online) Available at: <https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.3.0.pdf>
16. Open Network Foundation, Software-Defined Networking: The New Norm for Networks, ONF White Paper (2012)
17. Papatwibul, P., Banjar, A., Al Sabbagh, A., Braun, R.: An Intelligent Model for Distributed Systems in Next Generation Networks. In: Advanced Methods and Applications in Computational Intelligence, Springer, pp. 315–334 (2014)
18. Papatwibul, P., Banjar, A., Braun, R.: Using DAIM as a reactive interpreter for openflow networks to enable autonomic functionality, In: Proceedings of the ACM SIGCOMM Conference on SIGCOMM, pp. 523–524 (2013)
19. Papatwibul, P., Jozi, B., Braun, R.: Investigating O:Mib-Based Distributed Active Information Model (DAIM) for Autonomics, pp. 7–12 (2011)

20. Braun, R., Chiang, F.: A distributed active information model enabling distributed autonomies in complex electronic environments. *Third International Conference Broadband Communications, Information Technology & Biomedical Application*, pp. 473–479, IEEE (2008)
21. Stewart, R., Metz, C.: Sctp: New transport protocol for TCP/IP. *Internet Computing IEEE* **5**, 64–69 (2001)
22. Azodolmolky, S.: *Software Defined Networking with Openflow*, Packt Publishing Ltd. (2013)
23. Nadeau, T.D., Gray, K.: *SDN: Software Defined Networks*, O'Reilly Media (2013)
24. Shukla, V.: *Introduction to Software Defined Networking—Openflow & Vxlan*, North Charleston. SC: CreateSpace Independent Publishing Platform (2013)

Chapter 10

Designing Biomimetic-Inspired Middleware for Anticipative Sensor-Actor Networks

Christopher Chiu and Zenon Chaczko

Abstract Developing software environments for Sensor-Actor Networks (SANETS) is a promising research concern in systems engineering. Current concepts in software would adopt SANETS in a singular communications methodology, but the solution in this work is to take biological inspiration for the systems solution, thus the design of the system achieves a biomimetic construct as a result. SANETS are configurable for a variety of network structures and topologies, with the research aim in designing a network that is interactive and anticipatory to external and internal adaptations. Meanwhile, the event-based changes are composed of scenarios, and the interactivity between external and internal actors. From the requirements of the end-user, the system must be responsive and interactive from the user perspective in real-time, while in addition offering the contextual data to make useful interpretation of systemic conditions from an anticipative view Chiu and Chaczko Design of biomimetic middleware for anticipatory sensor-actor network systems. In: Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014, pp. 22–23. South Kuta, Indonesia, 10–12 February 2014. ISBN 978-0-9924518-0-6 [2].

Keywords Agent-centric simulation · Anticipative systems · Biomimetic-inspired methods · Sensor-actor based networks (SANETS) · Wireless Sensor Networks (WSN)

10.1 Introduction

The essential basis of co-operating Sensor Actor Networks (SANETS) is the interactivity amongst wireless networking hardware and human actors, both actuator and sensor, that consistently exchange information for mutual benefit. The interactivity

C. Chiu (✉) · Z. Chaczko
Faculty of Engineering and Information Technologies, University of Technology,
Sydney, Australia
e-mail: Christopher.Chiu@uts.edu.au

Z. Chaczko
e-mail: Zenon.Chaczko@uts.edu.au

of SANETS with the environmental surrounds demands a service model that is distributed and robust for the end-user requirements [6]. The delivery and provisioning of a distributed, harmonized service model is based on the coordination and collation of major tasks and requirements from heterogeneous data elements and resources [2].

The responsibilities of each SANET service can vary considerably, however the ultimate endpoint is to compliment the user's operational capability to predict and estimate developments that could occur amongst the SANET monitored environment. Developing the architecture for multi-agent concerns using SANET infrastructures ensures the elementary aspects of design is shaped during the engineering phase of development. The core aspects are highlighted below [9]:

- *Anticipative*: The design of the system must adapt to environments that evolve, because SANETS collectively sense and act upon the environment - with the exception that they have a limited lifespan until replacement or repair is required. Additionally, technological adaptations will lead towards new types of devices joining the final network. The architecture of the system that consists of the multiple devices, both external and internal to the SANET system, is suitably designed as software-based agents [1]. The software agents capture the thresholds, specifications and behavior of the SANET devices throughout the systemic network, and agency interaction between each other and the behaviors that are observed.
- *Interactive*: Captures the dimensional views of the SANET-based space of problems, with additional components adding to the complete actuation and sensory functionality, depending on the conditions of the environment. The information is collected in a procedural manner for eventual post-processes, and is executed at almost real-time for systemic efficiency. This is accomplished via the analysis of methodologies and the implementation of methods, by identifying heuristic-based approaches that is most suitable for the problem class that suits architectural designs for SANETS [7]. When the heuristic algorithm is sufficiently trained to detect local maxima and minima conditions, eventual computation of the data model is achieved.
- *Biomimetic in Conception*: Provisioning awareness of context to the architecture of the SANET, so it is capable for differences in the environment. The SANET network is configured for industrial, residential or commercial areas, with system complexity varying from heterogeneous to homogeneous domain concerns. The contextual variation is a core driver of the system architecture, while the system must be data focused—from the processing and aggregation of data as a complete structure [10]. This is achieved when one decouples the architecture with the heuristic software framework, to ensure the flexibility of the system accustoms to various methodologies of heuristics depending upon the space of the problem.

The results are expected to demonstrate the biomimetics-inspired facets of active heuristics combined with reinforcement-learned algorithms, optimally delivers a method for training systems in path navigation. Particularly, the construction of the SANET system as a multi-agent structure accomplished autonomic locality in perceiving senses, whilst ensuring global system objectives and targets. The robustness and adaptability of the system is done by constructing the SANET agency of service

with a devoted heuristic algorithm, while the space of the problem is deconstructed throughout the simplification of concerned subset information.

The research accomplished is composed of the following structures: The essential theories and concepts of SANET architectural solutions are provided, with the following sections describing about the contribution of engineering research and the experimental data in tabular form. The complete observations and conclusions are examined in detail, with the incorporation of learning-reinforcement algorithms in conjunction with active-based functions of heuristics.

10.2 SANET Environments for Agency-based Architectures

The main purpose of the experiment is to examine the Extended Spring Tensor Model (xSTEM) approach to analyze global perceptibility of the SANET domain. The main objective uses the open-based box experimental process, in order to contain the stages establishing the xSTEM-based model. The analytical approach of xSTEM assists in analyzing global interactivity, to view the SANET structure as motes driven by interactions throughout the externalized environment, because of the inherent distribution of interconnections with the SANET hardware motes.

The experimentation aims to evaluate the model of xSTEM to establish the magnitude and fluctuation of change within the SANET structure, so one can determine which model is suitable for distributed contexts in software-agent middleware systems. The ability of the xSTEM model to be used in contexts other than bio-computing constructs is an interesting concern, while reliability and performance are considered for future experimental work.

The experimental purpose is to determine the capability of xSTEM models for globalized SANET domains, resulting from the regionalized effects that influence the networking structure as a whole. The xSTEM model calculates the multitude of network structure sets, so the research objective determines the model's capability to estimate future SANET network configurations—resulting from environmental conditions occurring from an external standpoint. The experimental analysis will ultimately determine the final results, and reach an outcome to view the xSTEM model's ability to evaluate globalized concerns of SANETS, along with generalized experimental conclusions and research studies to consider in future.

The general procedure of the xSTEM-based experiment is elaborated as follows:

- (a) Determine the assumptions of the experiment for the Open-box Experiment approach;
- (b) Evaluate the approach of the algorithm of the xSTEM-based model for SANET structures;
- (c) Elaborate the method and procedure of the xSTEM-based model to observe globalized influences in the SANET structure;
- (d) Cater for the final results of the xSTEM-based Open-box model; and
- (e) Scrutinize the xSTEM-based model for SANET structures, along with future experimentation and examination.

10.3 Experimental Approach and Procedure

10.3.1 Approach of the Extended STEM Algorithm

The extension of the STEM approach is described in the algorithmic work of authors stated in [3, 10], by considering the SANET research work of [7] with the STEM algorithmic approach by [8]. The core steps of the XSTEM heuristic execution is focused on a reinforcement iteration in a singular fashion for feedback of sensor data from the externalized domain as summarized in Fig. 10.1.

The coordinates of the sensor data supports the first order derivative values as follows in Fig. 10.2:

- Values of bonds supported by length of radius between adjacent motes (r);
- Values of angles supported by angular bond amongst adjacent motes (θ);
- Values of dihedrals supported by angular dihedral amongst adjacent motes (ϕ);
and
- Contacts of non-localized motes based from the non-adjacent radius of motes (r_{ij}).

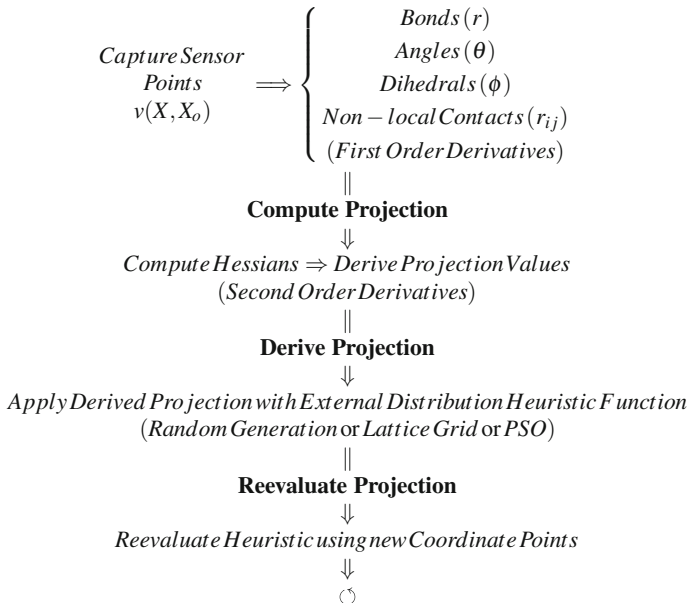
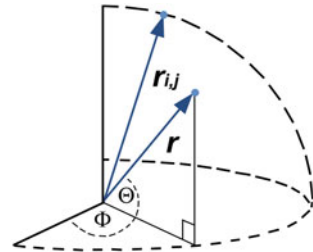


Fig. 10.1 Extended STEM Algorithm Process

Fig. 10.2 Geometric variables of the sensor coordinate space



10.3.2 Approach of the Mote Redistribution

Afterwards, the results of the second order derivative determines the projected values of direction and magnitude. The remodeled sensory information then undergoes post-processing for mote redistribution:

- Random arrangement with Mersenne-Twister random generation;
- Matrix arrangement with Equidistant Lattice Structure; and
- Agglomeration arrangement with Particle Swarm Optimization.

10.4 Procedure of Extended STEM Experiment

The research procedure is conducted in accordance to the experimental approach as shown in Fig. 10.3, with the main outcomes listed as follows:

- *Aims*: The aim of the xSTEM experiment is to ultimately determine the possibility of using a distributed SANET model for *perception via global means*, by evaluating against the heuristic of neural networks for optimization of the global optima. The method of Particle Swarm Optimization (PSO) is used to constitute the open-box experiments.
 - *Purpose*: The purpose of the xSTEM experiment is to find out the effectiveness of xSTEM in the distributed sensor field, as the research study executed in a defined framework generates results that can be interpreted and analyzed from a SANET structure perspective.
 - *Limitations*: The limits of the xSTEM experiment is that the model of the algorithm only considers the calculation of the projection of direction and magnitude from the SANET structure via global influences. The research study is operated on the MATLAB Toolkit by MathWorks Incorporated.
1. Whole amount of 'n' SANET motes is provided from the multi-agency input of data resources, in a 3-dimension network region of 50m^3 ($L \times W \times H \implies 5\text{m} \times 5\text{m} \times 2\text{m}$) in a 24h period. The general range from 25 to 200 motes an emulated rate of error for 5–10% of motes:

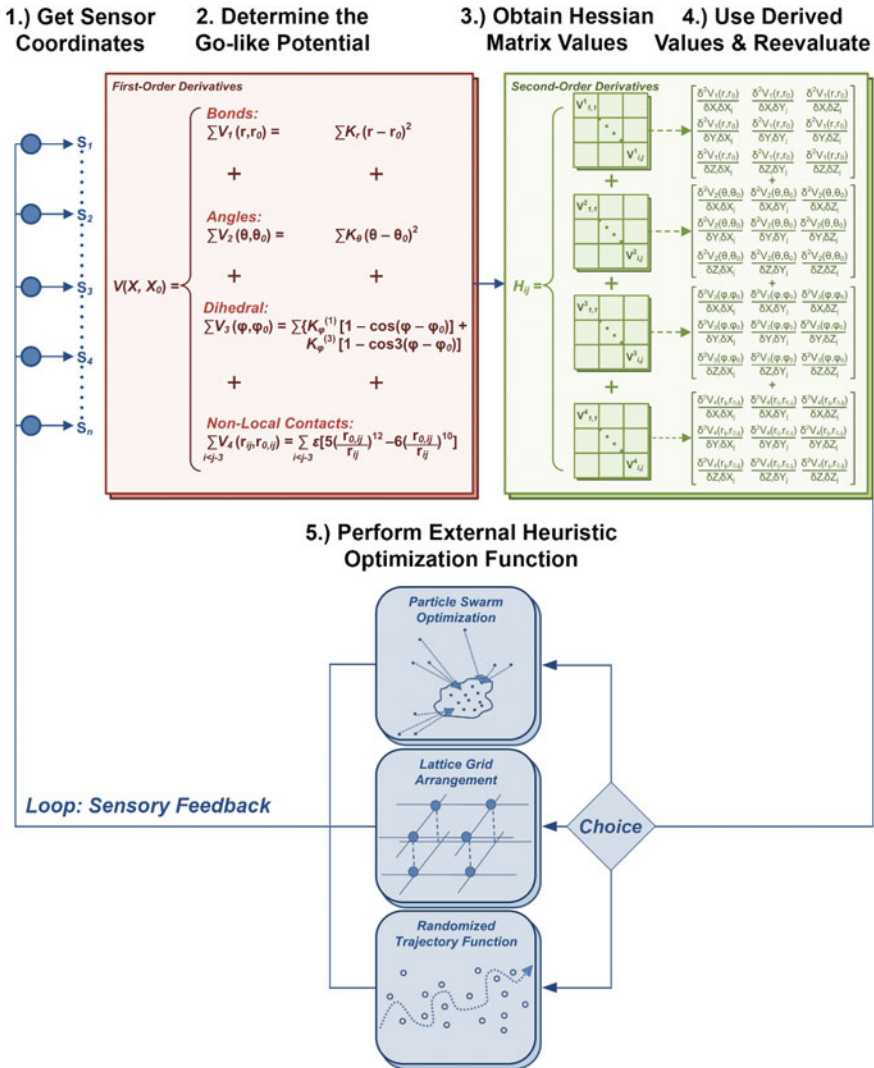


Fig. 10.3 Visual block diagram of extended STEM experiment

- (a) 25 mote elements; (e) 125 mote elements;
- (b) 50 mote elements; (f) 150 mote elements;
- (c) 75 mote elements; (g) 175 mote elements; and
- (d) 100 mote elements; (h) 200 mote elements

The controlling specifiers of the research experiment are described below:

(i) *Parameters of Data Input*

The sensory information is gathered from the ZigBee network stack via a virtual serial-port interface consisting of a 9,600 bits/s data rate. The program

connects to the serial interface, formatting the information into a text-file of format Comma Separated/Delimited Value (CSV). The MATLAB toolkit is executed to interpret the CSV text-file for the next process. The average sampling rate of the sensors is 2 reads/s.

(ii) *Laboratory Structure*

The laboratory structure is situated in the secure ICT Center located at the University of Technology, Sydney, Building 1, Floor 22. The laboratory is about 100 m above ground level, with air-conditioning set at 25 °C.

(iii) *Simulated Error Rate*

The Bit Error-Rate (BER) is generated using Sampling of Importance [11], with a probability of error that creates a randomized off-setting of the mote's position by $\pm 5\%$ of the real position of the mote:

- a. 5 % Generated BER: Creates in a random sample of 5 % of all SANET motes providing irrelevant data and coordinate points. By means of Sampling of Importance, henceforth 500 samples of data is necessary to evaluate the error rate of maximum 5 motes in the information set with a 1:100 sample ratio.
 - b. 10 % Generated BER: Creates in a random sample of 10 % of all SANET motes providing irrelevant data and coordinate points. By means of Sampling of Importance, henceforth 1,000 samples of data is necessary to evaluate the error rate of up to 10 motes in the information set with a 1:100 sample ratio.
2. The trajectory pathway is emulated in the SANET structure from a home geography of $x : 0 \text{ m}$, $y : 0 \text{ m}$, $z : 0 \text{ m}$ to the destination geography of $x : 5 \text{ m}$, $y : 5 \text{ m}$, $z : 2 \text{ m}$. The sensor events that motes will trigger proximity alerts only occur in a radius of 5 cm of the path. The distribution structure of the SANET network is described below:

a. *Distribution by Pseudo-Randomness*

The sensor motes are arranged randomly with the Mersenne-Twister approach. The seed is generated using the elapsed computer clock when the experiment takes place, with the value being the total seconds counted since January 1st, 1980.

b. *Distribution by Lattice-Grid Structure*

The sensor motes are equally distant amongst each other in a lattice-based structure. The length among motes is arranged according to the total volume of the experimental testbed as stated in 1(ii).

c. *Distribution by Particle Swarm Optimization*

The sensor motes aggregate according to the algorithm of global optimization, comparing between the distribution approaches of 2a and 2b. PSO is used for this case study, as the effectiveness of the heuristic is examined in the following literature [5]. PSO values are as described:

- (i) Total Heuristic Iterations 50
- (ii) PSO Factor of Inertia: 1.0
- (iii) PSO Corrective Factor: 2.0
- (iv) PSO Size of Swarm: SANET motes stated in 1 (a–h)

3. The xSTEM approach is evaluated to determine the globalized direction and magnitude of SANET motes, using the values as defined by the following literature [4]:
 - (i) K_r Length of Radius = 100ε
 - (ii) K_θ Angle of Bond = 20ε
 - (iii) $K_\phi^{(1)}$ Angle of Dihedral = ε
 - (iv) $K_\phi^{(3)}$ Angle of Dihedral = 0.5ε
 - (v) ε Value of Factor = 0.36

Evaluating the capability of the xSTEM heuristic model for SANET structures is accomplished by the manner as stated:

- a. Find the actual location of the SANET mote;
 - b. Find the predicted location of the SANET mote using the xSTEM heuristic approach;
 - c. Determine the new location of the SANET mote;
 - d. Determine the location differential for the predicted and new location of the SANET mote, using the mote's actual location as a reference point.
4. The procedure is repeated for 100 cycles to determine the average success of identifying the predicted location for each SANET mote. The success of identification occurs where the estimated location is within $\pm 5\%$ level of threshold of the predicted location:
 - a. The way to calculate the target of trajectory is achieved by calculating the spherical zone around the predicted location of the SANET mote. As the distance between destination and home is $\sqrt{(5-0)^2 + (5-0)^2 + (2-0)^2} = 7.34$ m, the predicted location is estimated to be within $7.34 \times 5\% = \pm 36$ cm of the mote's physical location.

10.5 Results of the Extended STEM Experimentation

10.5.1 Analysis of the Extended Spring Tensor Analysis Model

The experimental results of the xSTEM heuristic algorithm to estimate Random Data projection, shows that the increase of SANET motes from 25 to 200 devices results in a mean 16% betterment with 5% rate of error, and a mean 15% with 10% rate of error.

The trend in projection improves in a logarithmic fashion when the number of SANET mote agents increase, as the frequency of projection reaches 65% with 5% rate of error, and 60% with 10% rate of error. Increasing the overall rate of error from 5 to 10% results to a mean 3% decrease in estimating the projection.

10.5.2 Lattice Grid Structure Formation

The experimental results of the xSTEM heuristic algorithm to estimate Lattice Grid Structure projection, shows that the increase of SANET motes from 25 to 200 devices results in a mean 15 % betterment with 5 % rate of error, and a mean 14 % with 10 % rate of error.

The trend in projection improves in a logarithmic fashion when the number of SANET mote agents increase, as the frequency of projection reaches 66 % with 5 % rate of error, and 62 % with 10 % rate of error. Increasing the overall rate of error from 5 to 10 % results to a mean 4 % decrease in estimating the projection.

In comparison to Random Data arrangements, the Lattice Grid Structure operates in a more optimal fashion by a mean of 7 % with 5 % rate of error, and 6 % with 10 % rate of error. The differentiation in trend of projection for Random Data Structure and Lattice Grid Structure formation leads to an increase of 1 % with 5 % rate of error, and 2 % with 10 % rate of error.

10.5.3 Particle Swarm Optimization Structure Formation

The experimental results of the xSTEM heuristic algorithm to estimate Particle Swarm Optimization projection, shows that the increase of SANET motes from 25 to 200 devices results in a mean 16 % betterment with 5 % rate of error, and a mean 15 % with 10 % rate of error.

The trend in projection improves in a logarithmic fashion when the number of SANET mote agents increase, as the frequency of projection reaches 75 % with 5 % rate of error, and 70 % with 10 % rate of error. Increasing the overall rate of error from 5 to 10 % results to a mean 5 % decrease in estimating the projection.

In comparison to Lattice Grid arrangements, the Particle Swarm Optimization operates in a more optimal fashion by a mean of 8 % with 5 % rate of error, and 7 % with 10 % rate of error. The differentiation in trend of projection for Lattice Grid Structure and Particle Swarm Structure formation leads to an increase of 9 % with 5 % rate of error, and 8 % with 10 % rate of error.

10.5.4 Comparison of Results for Extended Spring Tensor Model Heuristic

The resulting improvement with the Extended Spring Tensor Model using Particle Swarm Optimization over the Random Data structure is elaborated below. By increasing the total population of SANET mote agents from 25 to 200 motes, a mean 10 % betterment with 5 % rate of error is recorded, and a mean 8.5 % betterment with 10 % rate of error is noted. Therefore, an overall mean betterment of 9 % is observed.

The resulting improvement with the Extended Spring Tensor Model using Particle Swarm Optimization over the Lattice Grid structure results in the increase in SANET mote agents from 25 to 200 motes leads to a mean 6.5 % betterment with 5 % rate of error, and a mean 5.5 % betterment with 10 % rate of error. Therefore, an overall mean betterment of 6 % is observed.

10.6 Conclusion of Research Study

The research study of the Extended Spring Tensor Model using the stated structure formations offers the analysis as follows, in order to establish the direction and magnitude of estimated change from a globalized Sensor-Actor Network perspective:

- **Random Mote Structure with Extended STEM Model**

xSTEM predicts the projection of randomly created sensor mote information by an average 40 % for sensor mote populations (25–200) with 14 % standard deviation. It is noted that the performance of projection is reduced by a mean of 4 % when error-injected input data increases from 5 to 10 %; as such it is noted that the reliability of sensor information is important for performance optimization of the xSTEM algorithmic model.

- **Lattice-Grid Structure with Extended STEM Model**

xSTEM predicts the projection of equidistant lattice structures by an average of 43 % for sensor mote populations (25–200) with 15 % standard deviation. The experimental results demonstrate that xSTEM heuristic optimization of lattice structures offers projection estimation improvement over random data, as lattice structures incorporates geographic structural stability instead of randomized structures.

- **PSO Distribution Structure with Extended STEM Model**

xSTEM predicts the project of Particle Swarm Optimization structures an average of 50 % for sensor mote populations (25–200) with 17 % standard deviation. The xSTEM heuristic model in combination with PSO distribution structure offers projection estimation improvement over lattice and random structures, as PSO has the ability to be adaptable to variability in the externalized SANET network.

- **Comparing xSTEM with PSO over Randomized & Lattice Grid Structures**

xSTEM projective capability of PSO distributions in comparison to random structures shows the predictive capability for PSO is increased by 9 % in comparison to random structures, and is increased by 6 % in comparison to lattice structures.

- **Effect of Rates of Error**

The estimation functionality of xSTEM is decreased when the rate of error being generated is increased from 5 to 10 % of SANET motes. The level of difference is a mean reduction of 4 % for random data structures, 5 % for grid data structures and 6 % for PSO structures.

The xSTEM heuristic results show for the distribution structure methodologies used in the research study, increasing SANET motes agents leads to an improvement in

the projection rate estimation capability. The improvement in prediction capability results from the increasing interactivity, particularly amongst SANET motes being analyzed by the xSTEM heuristic algorithm. It is noted that the xSTEM predictive frequency is not in proportion to the population of sensor motes, with the limiting factor being the methodology of distribution for the SANET system structure.

In addition, the rate of error for the sensory information has an effect on the projection capability of the xSTEM heuristic, as observed in Table 10.1. Throughout the distribution structures adopted, the xSTEM heuristic capability was lowered as the rate of error was raised from 5 to 10% of total SANET motes. Thus, the xSTEM heuristic must incorporate a filter for handling errors to reduce the sensory inputs of motes that malfunction, which ultimately improves the outcome of projection capability (Tables 10.2 and 10.3).

In conclusion, integrating xSTEM heuristics with PSO optimization approaches offers an optimum level of projection functionality over lattice-grid and randomized distribution structures, as PSO optimization behavior is more certain when compared

Table 10.1 xSTEM Open-box Experimental Results

No. of Agent Motes	Random Structure		Lattice-Grid Structure		PSOStructure	
	Rate of Error		Rate of Error		Rate of Error	
	5 (%)	10 (%)	5 (%)	10 (%)	5 (%)	10 (%)
25	16.891	16.223	18.353	17.562	20.980	19.920
50	25.532	24.682	27.622	26.501	31.601	29.751
75	35.881	34.581	39.021	37.312	44.862	42.562
100	42.911	41.332	46.730	44.543	53.493	50.382
125	45.632	44.210	49.423	46.991	56.592	53.581
150	51.371	49.581	55.882	53.131	63.812	60.183
175	54.041	52.172	58.792	56.433	67.201	63.502
200	58.870	56.893	63.811	60.722	72.940	68.670

Table 10.2 xSTEM with PSO Optimization versus Randomized Distribution

Number of Total Agent Motes	PSO versus Random Structure	
	Rate of Error	
	Randomized 5 (%)	Randomized 10 (%)
25	4.091	3.703
50	6.073	5.072
75	8.982	7.981
100	10.580	9.040
125	10.961	9.370
150	12.441	10.601
175	13.173	11.342
200	14.072	11.783

Table 10.3 xSTEM with PSO Optimization versus Lattice-Grid Distribution

Number of Total Agent Motes	PSO versus Lattice Grid Structure	
	Rate of Error	
	Randomized 5 (%)	Randomized 10 (%)
25	2.631	2.361
50	3.982	3.252
75	5.843	5.253
100	6.762	5.831
125	7.182	6.591
150	7.931	7.053
175	8.411	7.072
200	9.132	7.951

to randomized distributions. The capability of xSTEM heuristics to project the mote's direction and magnitude of change demonstrates potential, especially when applying global tensor functions in SANET-based infrastructures.

10.6.1 Future Task of Research Study

The magnitude and projection of changes for SANET infrastructures can be extended beyond Euclidean metrics. The projection of the physical realm with non-Euclidean metrics will result in overlapping global with local dimensional domains. Additionally, the parametric constants and thresholds adopted for the xSTEM heuristic requires more study, such that the input constants accurately model the perceptions of SANET-based infrastructures.

The further work suggested for the xSTEM heuristic algorithm is applicable for contexts where SANET network spaces is utilized. Essentially, the physical realm can be integrated further standardized approaches, both in terms of data structures and protocols adopted. The advantages of using the approach means the controller can harmonize with externalized components that add value for the SANET network structure.

References

1. Agrawal, D., Xie, B.: Encyclopedia on Ad-Hoc and Ubiquitous Computing Theory and Design of Ad-Hoc Networks. World Scientific Publishers, Singapore (2010)
2. Chiu, C., Chaczko, Z.: Design of biomimetic middleware for anticipatory sensor-actor network systems. In: Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014, pp. 22–23. South Kuta, Indonesia, 10–12 February 2014. ISBN 978-0-9924518-0-6

3. Chiu, C., Chaczko, Z.: Enhancement of surgical training practice with the spring tensor heuristic model. *Int. J. Electron. Telecommun.* **59**(3), 237–244 (2013)
4. Clementi, C., Nymeyer, H., Onuchic, J.N.: Topological and energetic factors: what determines the structural details of the transition state ensemble and En-route intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **298**(5), 937–953 (2000)
5. Clerc, M.: *Particle Swarm Optimization*. Wiley Publishers, Newport Beach (2006)
6. Dubois, D.: Introduction to computing anticipatory systems. *J. Comput. Anticip. Syst.* **314**(2), 3–14 (1998)
7. Karl, H., Willig, A.: *Protocols and Architectures for Wireless Sensor Networks*. Wiley Publishers, Hoboken (2007)
8. Lin, T.-L., Song, G.: Generalized spring tensor models for dynamics and conformation changes. *Struct. Biol. J.* **10**(1), 1–12 (2010)
9. Perlovsky, L., Kozma, R.: *Neurodynamics of Cognition and Consciousness*. Springer, Berlin (2007)
10. Rosen, R.: *Dynamical Systems Theory in Biology*. Wiley Publishers, New York (1970)
11. Song, W.T., Chiu, W., Goldsman, D.: Importance sampling techniques for estimating the bit error rate in digital communication systems. In: *Proceedings of the 37th Winter Simulation Conference*. United States of America (2005)

Chapter 11

Empirical Analysis of Terminal Reliability in Multistage Interconnection Networks

Nur Arzilawati Md Yunus and Mohamed Othman

Abstract Multistage interconnection network provide communication among processors, memory modules and other devices in parallel computer systems. These networks are designed to provide fast and efficient communication with a reasonable cost. The number of stages, interconnection topology, and network configuration differentiate the reliability of each network. Reliability is used to measure the system capability to transform information from input to output devices and its depends on the reliability of its component. Evaluation of reliability of interconnection network has been attempted by researchers in the past. This paper is an depth study of reliability evaluation in multistage interconnection network. The measurement includes terminal reliability of regular network known as shuffle exchange network and gamma network.

11.1 Introduction

Multistage interconnection network offer an attractive and cost effective solution to communication and interconnection between system components. The basic function of an interconnection is to transfer information from the input nodes of the network to the outputs nodes by setting up communication paths for routing through the network in an efficient manner [1]. The interconnection network enables communication between among the processor and memory modules themselves and also interconnects the processor and memory modules. Multistage Interconnection Network (MINs) is essentially a conciliation between crossbar and shared bus network [2]. MINs are used to design a network with a number of independent paths connections involving two modules [3]. Network topology can be defined as physical interconnection structure of the network graph. This topology are defined based on

N.A.M. Yunus (✉) · M. Othman
Department of Communication Network and Technology, Universiti Putra Malaysia,
UPM Serdang, 43400 Selangor, Selangor D.E., Malaysia
e-mail: arzilawati@gmail.com

M. Othman
e-mail: mothman@upm.edu.my

Fig. 11.1 Multistage interconnection network

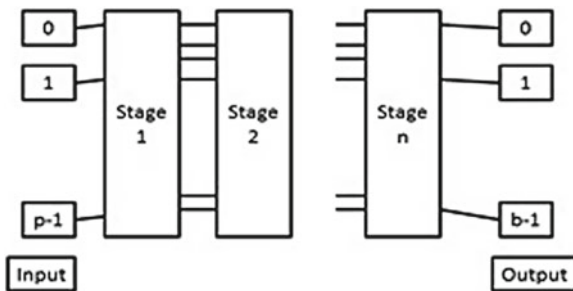
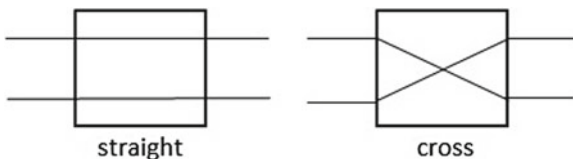


Fig. 11.2 Connection pattern



three parameters: (i) communication path, (ii) switching elements stages and (iii) links connectivity. MINs provide unique path between every pair of source and destination. Therefore, the failure of a single switching element destroys the several paths between input outputs (Fig. 11.1).

MINs normally connect N inputs and N outputs where N is referred to size of the network. The number of stage and the connection pattern between the stages determine the routing capability of the network [4]. The connection pattern for the interconnection network could be in straight or cross connection shows is Fig. 11.2. The main issue in MINs is to obtain reliable network topology to achieve the properly communication between processors and memories [4]. There are several different multistage interconnection network are studied in the literature. Figure 11.1 illustrates the structure of multistage interconnection network.

Reliability of an interconnection network defined as the possibility to facilitate a network works satisfactory for a predefined time scale and environment [5, 6]. Reliability of MINs is used to measure systems capability to transform information from input to output [1]. As the size and complexity increased the reliability become an important issue. When designing the communication network the reliability are important to consider especially in interconnection network.

For high reliability and performance, numerous methods have been recommended that offer fault tolerant to MINs. The network is considered failed once the connection between input and output is broken [4].When a switching network consists of a large number of switch nodes, the speed of the switching control over the whole network depends on how it is distributed. MIN is capable to assemble the reliability load if there is at least one substitute path to pact with collision or faults [7]. Evaluation of reliability interconnection network has been attempted by researchers in the past. This paper deals with the reliability problem and present a new network topology named Shuffle Exchange Network with Minus One Stage (SEN-). The designed pattern of SEN- is inspired by Shuffle Exchange Network (SEN). The reliability measurement

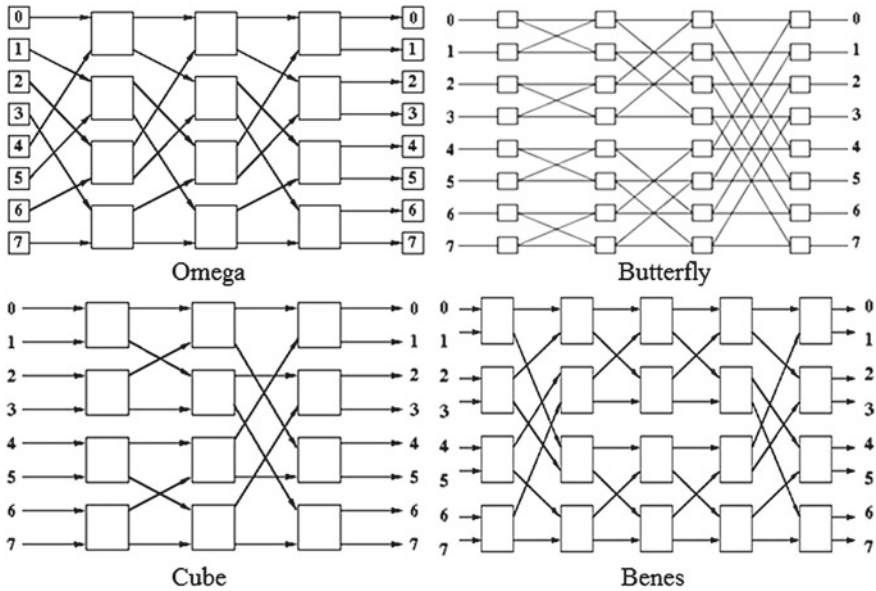


Fig. 11.3 Example of MINs

focusing on terminal reliability, including reliability of regular network known as shuffle exchange network and gamma network.

11.2 Literature Review

Multistage interconnection network is a large class network of topology that consist of the omega, gamma, cubes, baseline and binary networks [8]. MIN is a collection of switching elements and links distributed over different stages in a defined configuration. As the size and complexity system increases, reliability becomes an important issue [9]. Improves performance and increased reliability are two distinct advantages attributes to interconnection topologies [9]. It is not possible to send the entire source addresses at the same time to their corresponding destination because it will create the switch and link complete problem [10]. Figure 11.3 illustrate the example of MINs.

A fault tolerant MIN is one that provide service, in at least some cases, even when it contains a faulty component or components [11]. The unique path property also leads to a lack of fault tolerance. The failure of a single switch or link in the MIN can disconnect several inputs from several outputs [12]. Routing tags are particularly important for fault tolerant MINs since they should be able to specify a functioning path if one exists, tag limitation translate into fault tolerance limitation [11]. Routing tags are a way of describing a path through a network and providing for distributed network control [11]. When full access is the ultimate criteria of fault tolerance,

the extra switching stage provides a costly, though highly solution to failures [1]. Examples of such networks include Extra Stage Cube (ESC) [1], Extra Stage Omega, Augmented Shuffle Exchange Network (ASEN) [1], Irregular Augmented Shuffle Network (IASN) [13] and etc. For larger networks, the reliability in terms of mean time to failure (MTTF) is evaluated [14].

Reliable operation of interconnection networks is important to overall system performance. The reliability of interconnection network can be defined as the probability that the system has operated successfully over specific interval of time under static conditions. When designing communication network, reliability and cost are important to consider especially in interconnection network environment. The number of stages, interconnection topology, and the type of switching element used in the network configuration differentiate the interconnection reliability. As the switching element increase the terminal reliability of the network as function of the reliability of switching element (ranging from 0.990 to 0.999) increases [8]. There are generally three types of reliability which are important to MINs known as terminal, broadcast and network reliability [15, 16]. Reliability evaluation also typically assumes that each components in one of two modes completely working or completely failed [14, 17]. In order to define the reliability, behaviour of various interconnection network has been introduced. Omega network is an example of regular interconnection network. In extra stage Omega network, the additional stage is added and this network provides two paths communication between processor and memory modules [2, 16]. The reliability of extra stage Omega is higher than Omega network since the extra stage provide redundant path for communication process. [2, 16]. However the additional of two stages in the network decrease the reliability performance since the additional paths increase the links complexity in the network [2, 14, 16]. Enhanced MIN (E-MIN) referred as employment of two techniques to tolerate with fault tolerant which is extra staging and chaining. E-MIN is less expensive compared to chain network and also give a better performance in term of reliability and bandwidth compared to other fault tolerant MIN [18]. The modified SEN has been introduced to improve the terminal reliability. From the results shows that the modified SEN provide higher reliability than SEN for size $N > 4$ [8, 19].

The network architecture describes the network topology and its physical realization by determining the kinds and parameters of the network elements in detail. The network topology only gives the structure of the connection between the nodes related to graph theory. The local parts of such network are usually connected by switches. Different with Omega, Gamma has 3 input and 3 output crossbars. Gamma network are design in such way to provide redundant path between input and output terminal [6]. The previous researcher introduced an extra stage for Gamma network [6] to evaluate the reliability performance and compared with gamma network. The extra stage provide more redundant path in the network. However, the additional stage does not essentially improve the gamma network terminal reliability. The reliability for large size of Gamma network $N > 8$ is more complex due the complexity of Gamma configuration and availability of redundant paths [20]. The implementation of Crystal Ball in Gamma network, significantly reduce the network complexity and the results shows approximately of the exact reliability [4]. Due to the complexity of

Table 11.1 Properties of interconnection network

Properties	Bus	Crossbar	Multistage
Speed	Low	High	High
Cost	Low	High	Moderate
Reliability	Low	High	High
Configurability	High	Low	Moderate
Complexity	Low	High	Moderate

Table 11.2 Shuffle exchange network and gamma network comparison

Network	Shuffle exchange	Gamma
Input	N	2^n
Output	N	N
Stage	$(\log_2 N)$ stages	$(\log_2 N + 1)$ stages
Path	Single path	Redundant path
Switching element	2×2 SEs	3×3 SEs
Transmission	Cross/through connection	Routing tags
Intermediate stage	No	Yes
Reliability	Terminal, broadcast, network	Terminal

network configuration and availability of redundant paths, reliability bounds are used to estimate the exact reliability performance [20]. Table 11.1 shows the properties of interconnection network.

Reliability bound of Gamma network are introduced to estimate the exact reliability performance using the lower and upper bound method [20]. This method was reasonable and simple for reliability estimation of a complex large scale gamma network. As the size increases the reliability bound could be used to estimate the reliability of the networks [20]. In a gamma network, information from source nodes is transmitted through a specific set of routes to destination nodes [20]. The extra stage does not essentially increase the reliability of the network. However, by adding extra stages it can improve the fault tolerance of MINs. It also guarantees that there are two disjoint paths between each pair which are disjoint except for switches in the first and last stages [21]. Reliability evaluation of large Gamma network $N > 8$ size could be very complex due to the network complexity and availability of redundant path [20]. Table 11.2 shows the comparison between shuffle exchange network and Gamma network.

System performance could be analyzed based on its redundant paths, information transition time, and its reliability. Performance and reliability of interconnection network systems extensively depend on the interconnection of its components. As the number of interconnection paths and number of input and output of connecting switches increases, the performance measurement becomes more complicated. Reliability of the network is concerned with the ability of the network to carry out its desired network operation successfully. Lower bound reliability is useful for analysis larger networks [20, 22]. Cube network possess superior reliability comparison

with other extra stage. Existence of extra path increases the reliability of the system [22]. The generalize Cube topology is a unique path MINs with $N \times N$ input/output ports [22]. The Extra-Stage Cube (ESC) network is formed from generalize cube by adding an extra stage to the network. By using Stratified Monte Carlo Method in ESC the results shows that the reliability performance for ESC was improve [23]. For extra stage Cube network the lower bound network reliability closely approximates the exact network reliability for the extra stage cube network [20, 22]. There are mainly three types of reliability measures which are important to MINs, namely terminal, broadcast and network reliability. In this paper terminal reliability measure are considered for the analysis of the network. In general the network reliability $<$ broadcast reliability $<$ terminal reliability [15, 21]. The following assumptions are commonly made for the terminal reliability evaluation of MINs [24].

1. All failure are statistically independent.
2. All SEs are substantially less reliable than the links.
3. Each SE in the mIN has two states: failure, or operational with a known reliability.
4. All SEs identical and have constant exponential failure rates.
5. The SEs are not repairable. Source and terminal switches are always in operating condition.
6. A SE is in failure state when it cannot perform any the of the connection pattern.

11.3 Shuffle Exchange Network with Minus One Stage (SEN-)

Interconnection networks offer an attractive and cost effective solution to communication and interconnection between system components. The SEN- is proposed to improve the reliability performance in interconnection network and utilize the links in SEN. It has $(\log_2 N - 1)$ stage equal to two stages in the network. SEN- consist of stages that can route the switching through the path. The reason for reducing one stage is to allow paths communication with a lesser conflict between each source and destination. By reducing one stage in the network, the links and SEs also will be decreased. Since it has a lesser SEs compared to other SEN, it can decrease a link complexity in the network and be able to avoid a system from failure. It also helps to reduce the number of links failure in the network. SEN- is more reliable than SEN and Gamma network based on the switching element increase the results shows that the reliability performance increase compared to SEN and Gamma network. SEN- with eight inputs and eight outputs consists of 8 switches (SEs), is shown in Fig. 11.4.

11.4 Shuffle Exchange Network

The shuffle exchange network is self routing network. A message from any source to a given destination is routed through the network according to the binary representation of the destination address [25]. The input are transmitted the message

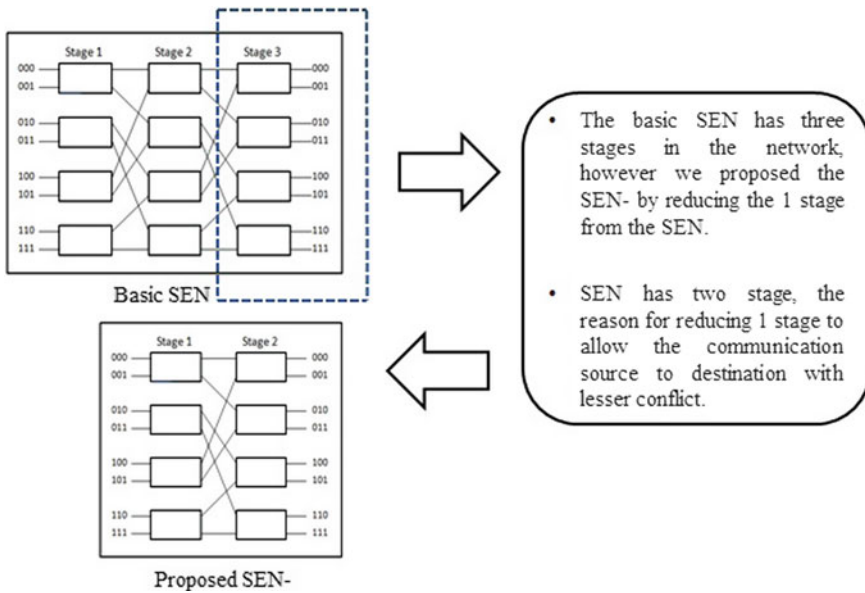


Fig. 11.4 Shuffle exchange network with minus one stage (SEN-)

to the output either straight or cross connection [8]. Usually this network has one path between particular input and output, therefore if any switching element fails the whole network will be affected. Shuffle exchange network is just one network in a large class of topologically [14]. A unique path between any source to a desired destination can be established by properly setting each switching element to a straight or cross connection [14]. The omega network is an example of a shuffle exchange network. It has $(\log_2 N)$ stages with $\frac{N}{2}$ SEs per stage. The message is routed based on the destination address from a given source to a given destination [25]. In this network, the address is shifted one bit to the left circularly in connection [10]. Shuffle exchange networks typically connect N inputs to N outputs and are referred to as $N \times N$. The parameter N is called the size of the network. Figure 11.5 shows the 8×8 size shuffle exchange network system.

11.5 Gamma Network

Advancement in processor technology with higher processing capability, makes interconnection networks highly demanded in computer networking areas. Interconnection networks provide the capability of connecting multiple processors, allowing sharing of resources such as memory and processing time. Much research has been performed in investigating the communication process, particularly in the area of interconnection networks. The properties of MINs can be broadly divided into three categories:

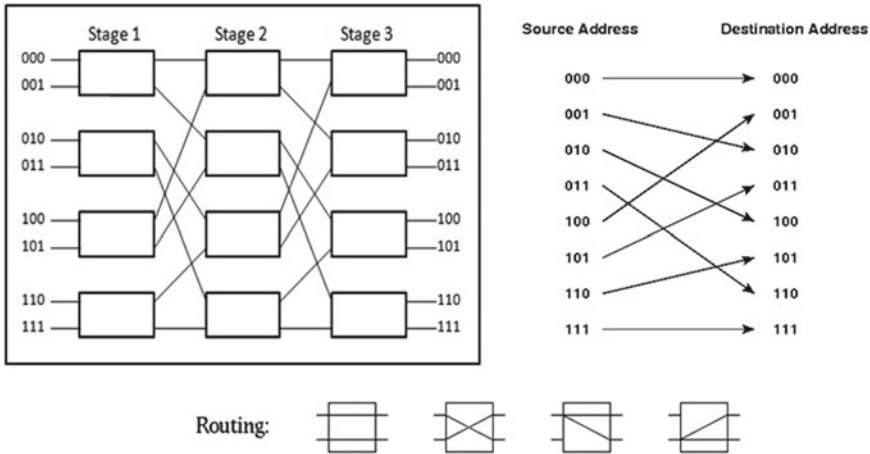


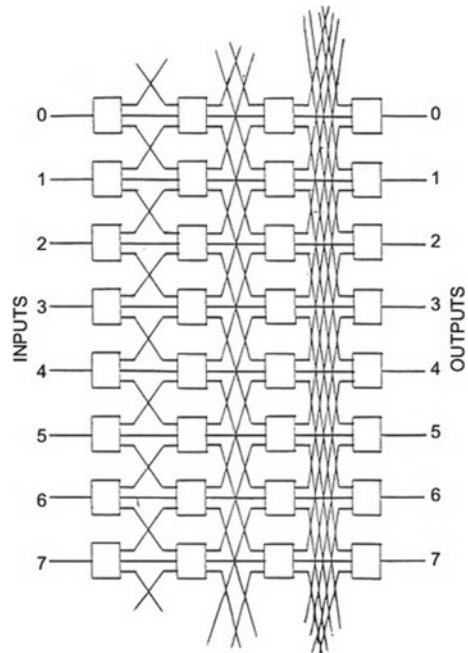
Fig. 11.5 Shuffle exchange network (SEN)

(i) one path, (ii) two path and (iii) multipath [24]. Gamma network belong to multipath network. This interconnection network connecting $N = 2^n$ input switches to N output switches [26]. It consists of $(\log_2 N + 1)$ stage with N switches per stage [20]. The stages are linked via power of two to ensure redundant paths exists between the input and output terminal [26]. Each switching element is represented by a block. In gamma networks, any switching element at the intermediate stage consists of three incoming and three outgoing links. The stages are linked in such way to ensure the redundant paths exist between the input and output terminals [6]. The network are considered failed once the connection between an input output are broken. The communication was established by using a routing tag procedure. Tag value has a positive and negative value. Both value are essential to generate the routing path [4]. Figure 11.6 illustrate a Gamma interconnection network with 8×8 sizes.

11.6 Terminal Reliability

Reliability of the network is concerned with the ability of a network to carry out a desired network operation successfully. To evaluate the performance of networks is a challenging task. Many parameters influence the performance of particular network architecture. In this paper the reliability measures focus on terminal reliability performance. Terminal reliability of a MIN is defined as the probability of existence at least one fault free path between a designated pair of input and output [16]. The terminal reliability for SEN, SEN- and gamma network can be calculated as shown below.

Fig. 11.6 Gamma network



11.6.1 Shuffle Exchange Network

By adding Shuffle-Exchange Networks (SENs) in MIN it has been commonly considered as convenient interconnection systems due to their size of its switching elements (SEs) and uncomplicated configuration. All the switches in a MIN are assumed to be of size 2×2 , and well known that 2×2 switch has only two possible states, straight and cross connection [10]. Let r be the time dependent reliability of the basic switching element. Terminal reliability for SEN can be formulated as follows:

$$R_t(SEN) = r^{\log_2 N} \tag{11.1}$$

11.6.2 Shuffle Exchange Network with Minus One Stage (SEN-)

The proposed SEN- consists of two stages for transmitting input from the source to the destination. It has N inputs and N outputs. Let r be the time dependent reliability of the basic switching element. There are $(\log_2 N - 1)$ defined as the number of stage in SEN- architecture. By reducing one stage from SEN it will decrease the number of links in the network. Terminal reliability for SEN- [27, 28] can be formulated as

Table 11.3 Terminal reliability comparison

R(t)	0.990	0.992	0.994	0.995	0.996	0.998
SEN-	0.9801	0.9840	0.9880	0.9900	0.9920	0.9960
SEN	0.9702	0.9761	0.9821	0.9850	0.9880	0.9940
Gamma	0.7857	0.8247	0.8655	0.8867	0.9083	0.9531

follows:

$$Rt(SEN-) = r^{\log_2 N - 1} \quad (11.2)$$

11.6.3 Gamma Network

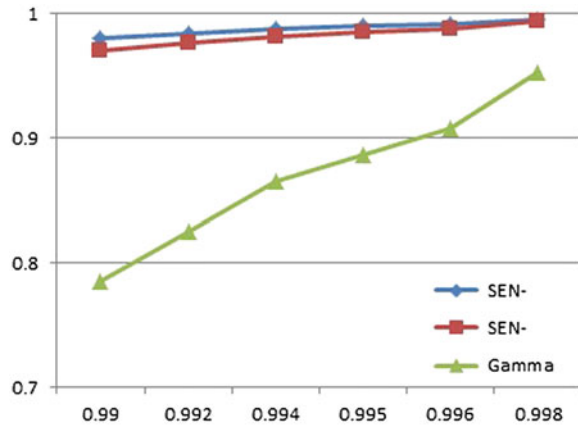
For the Gamma network, the lower bound reliability closely approximates the exact reliability of terminal therefore the terminal reliability of Gamma network are formulated as follows:

$$Rt(Gamma) = r^{2N} [1 - (1 - r^{\frac{N}{3}((\log_2 N) - 2)})^3][1 - (1 - r^{\frac{N}{2}})^2] \quad (11.3)$$

11.7 Results and Discussion

The interconnection network facilitates communication not only among the processors and the memory modules but also between the processors and memory modules. In this section the terminal reliability analysis is demonstrated. As explained in equation (1), (2) and (3) we will calculate the reliability performance of SEN-, SEN and Gamma network based on this equation. Figure 11.7 show the comparison for the three types of network namely SEN, SEN- and Gamma with the reliability performance for each network. From the figure the results explain that the increasing of switching element in the network will significantly increase the terminal reliability. However for the interconnection network comparison, the results shown that the SEN- has the highest terminal reliability, approximately 0.5 %, and 13.5 % compared to SEN and Gamma network respectively. From the graph it can be concluded by reducing one stage from SEN will decrease the number of links in the network, consequent from that the results for terminal reliability for SEN- will higher than SEN and Gamma network. Since gamma network provide redundant paths in the system, therefore the complexity occur in the configuration process and resulting the lower reliability compared to shuffle exchange network system (Table 11.3).

Fig. 11.7 Terminal reliability comparison



11.8 Conclusion

Interconnection networks play a major role in the performance of modern parallel computer. With advancement research in reliability field, an increased focus needs to be placed on the interconnection networks. In this paper, a shuffle exchange with minus one stage is introduced and the terminal reliability for this network is derived. In shuffle exchange network environment the increasing of network size significantly decrease the reliability performance in the network. This situation occurs because of links complexity in the network to passing the message from source to destination. By reducing one stage in SENs the switches is able to route with a lesser links complexity. Many systems are too complex to be determining by simulation method if detailed and accurate results are required. Then, measurement often turns out to be the only feasible solution. The performance evaluation of interconnection network is broadly based on the reliability performance in the networks. In this paper the terminal reliability of shuffle exchange network and gamma network are compared. The terminal reliability reveals that the Gamma network is less reliable than shuffle exchange network. The experimental results show that the SEN- provide higher terminal reliability compared to the SEN and gamma network. A Gamma network provide redundant paths in the system, therefore the complexity occur in the configuration process and resulting the lower reliability compared to shuffle exchange network system. It is shown that a specific modification in the structure of a standard shuffle exchange network will increase the reliability performance in the network. It is also shows that as the reliability of switching element increase the terminal reliability of this three network also increase.

Acknowledgments This work was supported by the Fundamental Research Grant Scheme (FRGS), Universiti Putra Malaysia (FRGS Number: OS/01/14/14S0FR).

References

1. Saini, N., Gupta, S., Thakral, B.: Multistage interconnection networks a review. *Int. J. Electron. Commun. Technol.* **4**, 1–4 (2013)
2. Lata, S., Dahiya, N., Chaudhary, S.: Reliability analysis of omega networks and its variants. *Int. J. Comput. Sci. Manag. Stud.* **12**, 138–141 (2012)
3. Rastogi, R., Verma, N., Nitin Chauhan, D.S.: 3-Disjoint paths fault-tolerant multi-stage interconnection networks. *Advances in Computing and Communications. Communication in computer and information science.* **190**, 21–33 (2011)
4. Gunawan, I., Gan, M.L.: Reliability analysis of gamma interconnection network systems. *Int. J. Perform. Eng.* **5**, 485–492 (2009)
5. Dash, R.K., Barpanda, N.K., Tripathy, P.K., Tripathy, C.R.: Network reliability optimization problem of interconnection network under node-edge failure model. *Appl. Soft Comput.* **12**, 2322–2328 (2012)
6. Gunawan, I., Fard, N.S.: Terminal reliability assessment of gamma and extra-stage gamma networks. *Int. J. Qual. Reliab. Manag.* **29**, 820–831 (2012)
7. Nitin Garhwal, S., Srivastava, N.: Designing a fault-tolerant fully-chained combining switches multi-stage interconnection network with disjoint paths. *J. Supercomput.* **55**, 400–431 (2011)
8. Fard, N.S., Gunawan, I.: Terminal reliability improvement of shuffle exchange network systems. *Int. J. Reliab. Qual. Saf. Eng.* **12**, 51–60 (2005)
9. Jena, S., Sowmya, G.S., Radhika, P., Reddy, P.V.: Reliability analysis of multipath multistage interconnection network. *Int. J. Comput. Sci. Inf. Technol.* **4**, 63–74 (2012)
10. Bhardwaj, V.P., Nitin Tyagi, V.: Minimizing the switch and link conflict in an optical multistage interconnection network. *Int. J. Comput. Sci.* **8**, 206–214 (2011)
11. Adam, G.B., Agrawal, D.P., Siegel, H.J.: A survey and comparison of fault tolerant multistage interconnection networks. *Computers* **20**, 14–27 (1987)
12. Kumar, V.P., Wang, S.J.: Reliability enhancement by time and space redundancy in multistage interconnection networks. *IEEE Trans. Reliab.* **40**, 461–472 (1991)
13. Sharma, K., Kahlon, K.S., Bansal, P.K.: Reliability and path length analysis of irregular fault tolerant multistage interconnection network. *ACM SIGARCH Comput. Archit. News* **37**, 16–23 (2009)
14. Chaudhary, S., Madaan, D., Kumar, S.: Empirical comparison and analysis of shuffle exchange network and its variant. *Int. J. Comput. Sci. Manag. Stud.* **12**, 147–150 (2012)
15. Fard, N.S., Gunawan, I.: Multistage interconnection networks reliability. In: CiteseerX, pp. 1–4 (2007)
16. Gunawan, I.: Reliability analysis of shuffle-exchange network systems. *Reliab. Eng. Syst. Saf.* **93**, 271–276 (2008)
17. Trahan, J.L., Wang, D.X., Rai, S.: Dependent and multimode failures in reliability evaluation of extra stage shuffle exchange MINs. *IEEE Trans. Reliab.* **44**, 73–86 (1995)
18. Arya, K.V., Ghosh, R.K.: Designing a new class of fault tolerant multistage interconnection networks. *J. Interconnect. Netw.* **6**, 361–382 (2005)
19. Yunus, N.A.M., Othman, M.: Shuffle exchange network in multistage interconnection network : a review and challenges. *Int. J. Comput. Electr. Eng.* **2**, 724–728 (2011)
20. Gunawan, I.: Redundant paths and reliability bounds in gamma networks. *Appl. Math. Model.* **32**, 588–594 (2008)
21. Menezes, B.L., Bakhru, U.: New bounds on the reliability of augmented shuffle exchange networks. *IEEE Trans. Comput.* **44**, 123–129 (1995)
22. Fard, N.S., Gunawan, I.: Reliability bounds for large multistage interconnection networks. *Appl. Parallel Comput.* **2367**, 507–514 (2002)
23. Gunawan, I.: Multistage interconnection networks reliability evaluation based on stratified sampling Monte Carlo method. *Int. J. Model. Simul.* **28**, 209–214 (2008)
24. Blake, J.T., Trivedi, K.S.: Multistage interconnection networks reliability. *IEEE Trans. Comput.* **38**, 1600–1604 (1989)

25. Feng, T.S.: A survey of interconnection networks. *Computer* **14**, 12–27 (1981)
26. Parker, D.S., Raghavendra, C.S.: The gamma network. *IEEE Trans. Comput.* **33**, 367–373 (1984)
27. Yunus, N.A.M., Othman, M.: Reliability performance of shuffle exchange omega network. In: *Proceedings of 1st IEEE International Symposium on Telecommunication Technologies*, pp. 26–30 (2012)
28. Yunus, N.A.M., Othman, M.: Empirical analysis of terminal reliability in multistage interconnection networks. In: *Proceedings of 2nd Asia-Pacific Conference on Computer-Aided System Engineering*, pp. 42–43 (2014)

Chapter 12

Authorial, Adaptive Method of Users' Authentication and Authorization

Robert Sekulski and Marek Woda

Abstract The main goal of this paper was to introduce an adaptive method of users' authentication and authorization. By *adaptive* one can understand having an ability to suit different conditions (namely different users devices, different user's profiles and different users behaviors). Over a time working conditions and habits have changed and employees may often use own devices from different locations to connect company's resources. This situation poses a severe threat to security, and tightening security rules is not always an option. This brought a need of an adaptive system, which would choose methods adequate to the current threat level. Presented *authorial* solution not only minimizes the risk of unauthorized access to company's data, but also simplifies users' authentication process. Example implementation and performed test scenarios showed that taken approach works and all theoretical assumptions are valid and possible to implement in the real world scenario.

12.1 Introduction

It should not be questionable to say that Internet have become an integral part of people's lives nowadays. In survey [1], more than half respondents answered that they could not live without the Internet. Different people list different reasons of its popularity. Among the most common responses, fast access to information on almost every topic, easiness of data share, facilitation of communication between people regardless of the distance between them and multiplicity of on-line services available mostly for free, are placed. An ease of data share poses a serious threat to privacy and possibility of confidential data leakage. This applies especially to a various type of business for which any leak of corporate information may be pernicious. It became

R. Sekulski · M. Woda (✉)

Department of Computer Engineering, Control and Robotics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: robert.sekulski@gmail.com

M. Woda

e-mail: marek.woda@pwr.edu.pl

© Springer International Publishing Switzerland 2015

G. Borowik et al. (eds.), *Computational Intelligence and Efficiency in Engineering Systems*, Studies in Computational Intelligence 595, DOI 10.1007/978-3-319-15720-7_12

obvious that there is a need of protecting access to at least some. D. Todorov [2] describes three security processes that should be used jointly in order to provide protected access to resources. Over the last years many various implementations of them were created. As R. Smith indicates: *Today's authentication systems evolved from decades of attacks, many of them successful*—[3]. Blissful ignorance of users combined with new threats released every day, cause that the number of infected computers rises every year [4]. K. Miller brought in his publication [5] concerns about security of BYOD trend.

The biggest gripe with BYOD is a small or even zero level of control which company has over devices, which are not its assets. BYOD postulate along with people's tendency to not respecting basic password security policies, like not to use the same password for various systems [6] entails risks that adversary will be able to obtain a remote access to the victim's account and later will get access to company's confidential data. As it is stated [1] companies will have to change their current security systems to ones based on a user, role and device type. In order to do so, understanding who is using a device, where it is being used, and what information is accessed, is vital. Currently there is a growing demand for identity management systems that are usually neither universal nor simple in use. They work in following manner the more suspicious (non-typical) is user's activity then the more detailed authentication process is conducted [7].

12.1.1 Existing Solutions

There are a number of approaches to adaptive security control unfortunately none is the best to fulfill a desired level of granularity in terms of usefulness and easiness to implement. Threat-Adaptive Security Policy [8] mimics the human approach of placing trust on others based on their actions, and despite its reasonable good results is not the best solution for all systems. It creates a problem of choosing a set of positive and negative behaviors used for calculating the trust. Also a creation of users profile may be a complicated task in systems which allows multiple users' devices as users' behaviors, may vary based on them. Self-Adaptive Authorization Framework [9] integrates with the existing RBAC/ABAC authorization infrastructure to manage its configuration in an automate way. The preventive action is being chosen from the set of actions based on a result that comes from evaluation function (e.g. function minimizing costs). It may happen that best solution will not be selected due to its high cost. This approach is fairly complex to implement. Dynamic Authorization Model Based on Security Label and Role [10] minimizes complexity of roles assignments when users need access to resources on different security levels it is based on Web Services and introduces fairly intricate ontology. Proposed Policy Resolver manages user rights, obligations, may be too difficult to implement for the most organizations. OpenAM is all-in-one access management platform for protecting any type of resource across enterprise [10] however it cannot exists solely on its own.

12.2 Proposed Solution

The main goal of this study was to propose an adaptive method of users' authentication and authorization. Adaptive means *having an ability to change to suit different conditions* [11]. In this case different conditions were: different users devices, different users profiles and different user's behaviors. As it was said in introduction working environment and habits have recently changed and employees may often use different devices from different locations to connect company's resources. This situation is a severe threat to security, however as was later said in the same section, tightening security rules is not always an option. This brings a need of an adaptive system, which would choose methods adequate to the current threat level. Adaptive solution would not only minimize the risk of unauthorized access to company's data, but also simplified users' authentication process. Within many types of currently available devices which user may use to work, mobile devices are the less secure. It is due to their size, what makes them easier to be stolen or lost. Following this line of thinking, desktops are the most secure, next are laptops, then tablet mobile devices and the least secure are mobile phones. This assumption is not entirely correct. Mobile phones may be the easiest to stole or lose, but if user was logged off of the system, even when someone unauthorized obtained physical access to the device, he will not be able to pass even the first step of authentication process. In the same time desktops and laptops may be targets of malicious software which steal users credentials what would allow adversary to log into the system. Of course such software exist for mobile phones as well as for desktops, but their number is still very low comparing to threats for desktops and laptops. This shows that device type should not be the only factor deciding which authentication methods should be used. Just as important should be match with users behavior profile. If someone works 5 days a week between 8 am and 4 pm, when he suddenly tries to log into the system at 6 pm authentication process should be more complex for him. Such a profile should be created for each of users devices. That would increase profiles' accuracy as it should be suspicious when user normally uses only his mobile phone after 5 pm and suddenly he tries to log in from his desktop. The next case when system should decide which authentication method should be used is during user's work. If user starts to behave suspicious, not as always, for example he starts to copy all data, what he never did before and what is described as intrusive action for his role, then system should perform additional authentication. The authentication complexity in that case should be dependent of the user's current threat level. Of course the method in that case should be also different than methods used to authenticate the user for this session. Adaptive authentication may be also used to provide better users' experience. Some methods may be cumbersome to pass on some devices, while other providing the same security level are easier to pass. Taking that into account, different set of methods may be used depending on the users device type. That practice would speed up users authentication process and provide him more comfort.

12.2.1 Threat Calculating Algorithm

In order to provide adaptive authentication, a solution similar to OpenAM's authentication chain was used. During first request in a new session user's device is being recognized using device fingerprinting and system matches it to an appropriate user's profile. If no profile for such device exists new one is being created. Based on that match user has to pass selected authentication methods from the chain to obtain required security level for the current request. After that he can perform actions to which he has permission granted within his role and security level. If he tries to perform action, which is defined in his role, but requires higher security level, then additional authentication has to be performed.

The security level can be lowered during user's work in response to his abnormal behavior or after performing suspicious actions. The abnormal behavior means that user uses the system in a different way than usual. The suspicious actions set is defined for each role. Higher security level can be obtained by passing more methods from the authentication chain. Lowering it is done by system as a response to user's suspicious actions and abnormal behaviors.

12.3 Main System Components

The main goal of this work was to create a solution which can be easily configured and extended to best fit different users' requirements. To empirically test base assumptions and validate the core design an proof-of-concept implementation has been done. This chapter describes in details that implementation and shows how it can be configured and extended.

Figure 12.1 shows the high-level diagram of system components. Separation of different parts of the system allows users to replace only small part of the code in order to modify or enhance chosen functionality.

12.3.1 Request Filter

Request Filter's role is to intercept all requests to the resources to which access is restricted. It then passes the request to Decision Point (described in Sect. 12.3.1) which handles it. In sample application only one filter is being used, however in other situations many filters can be used if filter's role is more complex than just passing requests. The only requirement is that each of used filters must call Decision Point's "allow" method with intercepted *HttpServletRequest* and *HttpServletResponse*.

12.3.1.1 Decision Point

Decision Point is the core component which handles whole processing of each request. It dispatches requests to adequate components depending on the current state

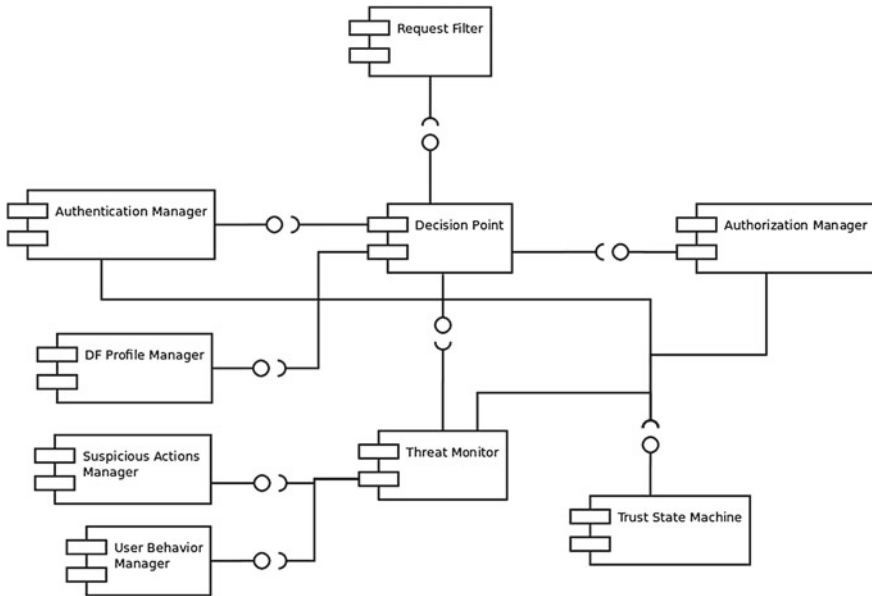


Fig. 12.1 Main components of the system

of the process and processes returned information. Figure 12.2 presents the process of handling requests by Decision Point. To work correctly, system has to load information like device type, role, behavior profile and a list of suspicious actions for current user. The first attempt to recognize the user is using device fingerprinting. When first request in new session is being processed user is being redirected to the site which tries to collect his device fingerprint. This process is described in Sect. 12.3.2.

When Decision Point gets the request with collected data it passes it to DF Profile Manager which returns the best matching profile or information that there is no suitable profile. This process is described in Sect. 12.3.5. No matter what was returned from the DF Profile Manager in the next step Decision Point calls Authentication Manager to perform authentication what is described in Sect. 12.3.4.

Decision Point and Authentication Manager use Authentication Request and Authentication Response to communicate. The first one is being sent when authentication of the user is required and the second one when Decision Point passes authentication data obtained from the authentication method. In current implementation they both contain the same data which is users Profile, *HttpServletRequest* and *HttpServletResponse*. As they do not introduce any logic and are only used to carry data, they do not implement any interface, but could be easily extended when additional information would have to be sent in future implementations. Because profiles returned by the DF Profile Manager can be mismatched and some of the authentication methods do not require entering username just password, while profile has not been confirmed (before first successful login) user has a possibility to inform the system that his profile was wrongly selected. In that case Decision Point

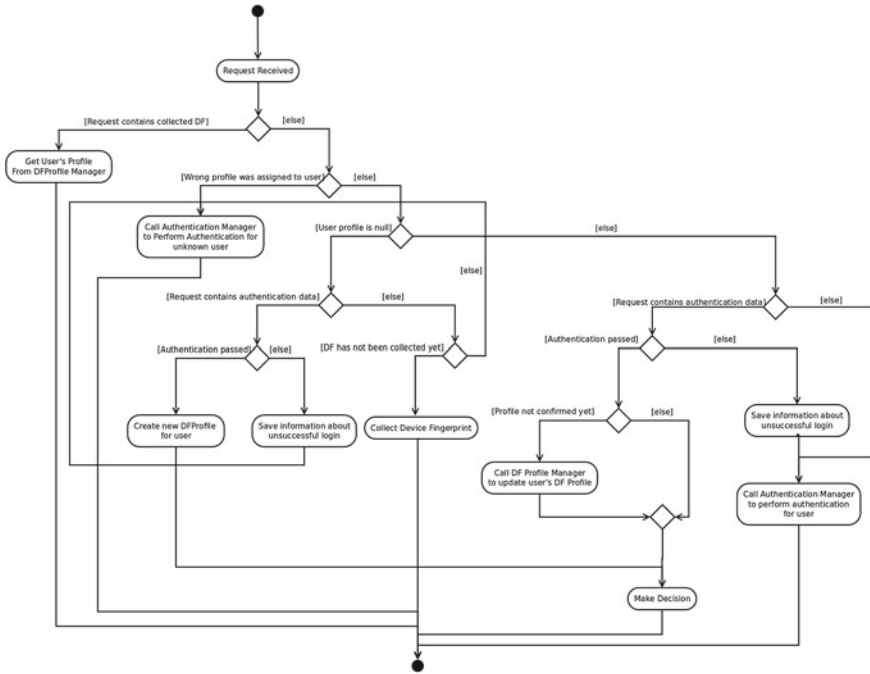


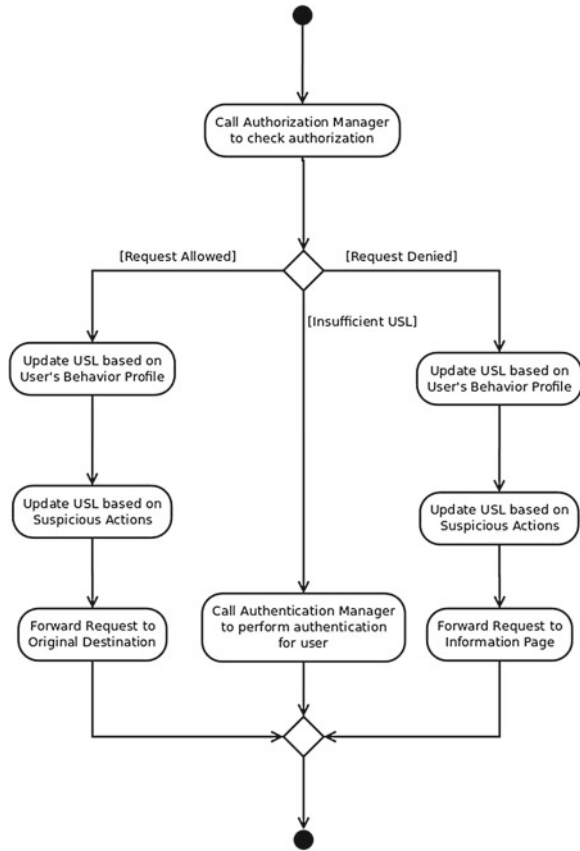
Fig. 12.2 Steps of processing a request in Decision Point

will order Authentication Manager to authenticate user using default authentication method which is obligated to require providing username. When user authenticates himself for the first time in new session if his DF profile was matched correctly it is being updated. Otherwise new DF profile is being created for the user. After this Decision Point makes decision how to respond to users request. If user did not pass the authentication then information about unsuccessful attempt is being saved and user has to try again. In order to make decision Decision Point consults Authorization Manager. To communicate they use Authorization Request and Authorization Response they carry only data and do not implement any interface. Authorization Request contains *HttpServletRequest* and user's Profile.

It is being sent to Authorization Manager which responds with Authorization Response containing Authorization Decision and a text message. Authorization Decision can be "ALLOW", "DENY" or "INSUFFICIENT_USL" and the text message is being used to explain the reason of denying a request. "INSUFFICIENT_USL" means that user has to pass additional authentication steps to be able to finalize his request. Figure 12.3 presents how Decision Point handles different Authorization Managers decisions.

When request is allowed Decision Point calls Threat Monitor to check user's action against his behavior profile and the list of suspicious actions. This process is described in Sect. 12.3.3. It may result in decreasing user's security level what will

Fig. 12.3 Process of handling Authorization Response in Decision Point



have impact on users next requests. After this user's request is being forwarded to its original destination. In case when request was denied the flow is almost the same but instead of forwarding to the original destination user is being forwarded to the page with the information that his request could not be finalized. When Authorization Manager decides that user has insufficient security level then Decision Point sends Authentication Request to Authentication Manager which forwards a user to next authentication step. No interaction with Threat Monitor happens in this case as it is considered as an internal action of the system.

12.3.2 Device Fingerprinting

In the first draw of the system, Device Fingerprinting method was planned to be used as a one of the authentication methods. Just by recognizing his machine user would have been granted some security level. However in the final solution Device

Fingerprinting is used only as a recognition to choose correct authentication method for a user. This change was made because as Device Fingerprinting method allows to generate unique enough identifiers to be used in limited group of users, the system cannot operate only on raw calculated fingerprint. Due to the fact that data used to generate fingerprint can be changed (e.g. by upgrading web browser or one of the plugins) system has to allow some errors in matching a profile. How it is implemented in the example implementation is described in Sect. 12.3.3. This section describes what data and how is collected.

After analyzing available papers describing different Device Fingerprinting methods, a decision was made to use only basic methods in the example implementation. To collect required data, a user is being redirected to the site which uses javascript and flash technologies providing required functionalities. Following information are being obtained using javascript:

- available plugins,
- is DOM local storage available,
- is DOM session storage available,
- are cookies enabled,
- current timezone,
- screen resolution.

For browsers which do not support javascript's "navigator.plugins" functionality an fallback method has been created. The availability of following plugins is being checked using a library "PluginDetect" version: 0.8.1 created by Eric Gerds, available at: <http://www.pinlady.net/PluginDetect>.

- Java, QuickTime, DevalVR, Shockwave, Flash, WindowsMediaPlayer, SilverLight, VLC, AdobeReader.

The information about available fonts was collected using flash script written by Henrik Gemal, available at: <http://browserspy.dk>. Data like preferred locales, IP address and User Agent were collected from the *HttpServletRequest*.

The Fingerprint is stored as a hash of the concatenated string values of all collected data. What is worth of noticing is fact that information about available plugins and fonts are also stored as hashes. This approach significantly reduces time needed to compare and find a fingerprint in the database and introduces another unique factor which is an order of fonts and plugins in which they were collected.

12.3.3 DF Profile Manager

Device Fingerprint Profile Manager's role is to save, retrieve and update information about users devices. One of the main system's assumptions is that a user can use many different devices in his daily work. This component manages DF Profiles which are bond with the user's profile Fig. 12.4. Decision Point communicates with DF Profile Manager using an interface which provides methods to create new DF Profile, update


```

<abnormalBehaviors xmlns="http://www.w3schools.com"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3schools.com abnormalBehaviorsConfig.xsd">
  <profileClass name="WORK">
    <behavior name="timeOfFirstLogin" points="5" variance="10"/>
    <behavior name="numberOfRequests" points="6" variance="10"/>
    <behavior name="proportionsOfRequestsTypes" points="4" variance="10" />
  </profileClass>
  <profileClass name="PC">
    <behavior name="timeOfFirstLogin" points="3" variance="10"/>
    <behavior name="numberOfRequests" points="6" variance="10"/>
    <behavior name="proportionsOfRequestsTypes" points="4" variance="10" />
  </profileClass>
  <profileClass name="MOBILE">
    <behavior name="numberOfRequests" points="6" variance="10"/>
    <behavior name="proportionsOfRequestsTypes" points="6" variance="5" />
  </profileClass>
</abnormalBehaviors>

```

Fig. 12.4 Example configuration of the user's behavior profile

existing one and get profile which the best matches collected data. The process of collecting data is described in Sect. 12.3.2.

The mostly often used functionality—matching profile based on the data—does not guarantee correct results. The reason for that was explained in Sect. 12.3.2. In the example implementation system first calculates a fingerprint from all collected data and checks if it exists in the database. If yes, then it is a sure match, because fingerprints are marked as a unique in the database. In other case, DF Profile Manager loads from the database a set of candidate profiles that are the profiles which have the same hashes of fonts and plugins as the ones generated from the collected data. For each of the candidate profiles a number of minor points is being calculated. Minor points are being granted for the equal values of the other fields like timezone, agent string, etc. From the set a profile with the highest number of minor points is being returned. If two or more profiles have the same amount of minor points then the first one of them is being returned. In case when the set of candidate profiles was empty, that means there were no profiles with the same hashes of the plugins and fonts strings, a new set is being created with the profiles having only one of those strings equals. The procedure of choosing one of the candidate profiles remains the same, however the selected profile has to have at least a specified amount of minor points to be considered as a match. In other case no profile will be returned.

DF Profiles have classes. A device class is required because it is being used to choose correct settings for user's abnormal behaviors and also to choose correct authentication chain. These functionalities are described in Sects. 12.3.5 and 12.3.4. Device class is being signed to the profile during its creation. In the sample implementation classification of the profile is done by the Profile Class Manager. The requirements for all classes are stored in the XML file. Each class has to have specified name and maximal USL (User Security Level) which can be obtained by the user using the device from this class. In case when device can be classified to more than one class, the class with the highest possible USL is being chosen. Each class

```

<roles xmlns="http://www.w3schools.com"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3schools.com rolesConfig.xsd">

  <role name="DEVELOPER">
    <permission path="/classified/classified" method="GET" usl="8" />
    <permission path="/data/*" method="GET" usl="5" />
    <permission path="/data/*" method="POST" usl="5" />
  </role>
  <role name="ADMINISTRATOR">
    <permission path="/data/" method="POST" usl="5" />
    <permission path="/admin/" method="POST" usl="8" />
    <permission path="/data/" method="GET" usl="5" />
    <permission path="/admin/" method="GET" usl="8" />
  </role>
  <role name="HR">
    <permission path="/users/" method="GET" usl="3" />
    <permission path="/users/" method="POST" usl="6" />
    <permission path="/data/*" method="GET" usl="8" />
  </role>
</roles>

```

Fig. 12.5 Example configuration of the roles

can have specified constraints which should be fulfilled by the device to be classified to the class. Every property can have an accepted value or range of accepted values. To be classified to the class device has to fit in all specified constraints. The configuration should contain one default class which has no restrictions for those devices which were not classified to any of the other classes. The last functionality provided by the DF Profile Manager—update existing profile—updates all data stored for the profile in the database, but those changes cannot result in changing devices class.

12.3.4 Authorization Manager

Authorization Manager's responsibility is to decide if user has a right to make an action on specified resources. It makes its decisions based on the roles' configuration and the user's current security level. Each user of the system has assigned a role. Each role has configured permissions to specified resources. It is possible to configure required USL for specified action on selected resources for each role. Figure 12.5 shows the example configuration of the roles. Each role has to have a name. It can contain unlimited number of permissions which have to have defined path to the resources, HTTP method and required USL. The path with the "*" at the end means that the permission covers all subpaths of the path.

When Authorization Manager makes a decision if user's request should be allowed, it first gets user's current security level from Trust State Machine. If this level is equals to 0 than it means that users account has been blocked and his request is denied with the message informing about the blocked account. Otherwise

```

<USLconfig xmlns="http://www.w3schools.com"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3schools.com USLConfig.xsd">

  <level number="1" minPoints="1" initialPoints="5"/>
  <level number="2" minPoints="6" initialPoints="10"/>
  <level number="3" minPoints="11" initialPoints="15"/>
  <level number="4" minPoints="16" initialPoints="19"/>
  <level number="5" minPoints="20" initialPoints="23"/>
  <level number="6" minPoints="24" initialPoints="27"/>
  <level number="7" minPoints="28" initialPoints="29"/>
  <level number="8" minPoints="30" initialPoints="30"/>
</USLconfig>

```

Fig. 12.6 Example configuration of the USLs

Authorization Manager tries to match requested path and method with the permissions provided for the user's role. The permissions are listed in the same order as they occur in the XML file. If it finds a match then it compares user's security level with the required by the permission and if it is greater, than the decision is "ALLOW" else it is "INSUFFICIENT_USL". If Authorization Manager did not find any match then request is denied with the appropriate information. Roles configuration is being provided to the Authorization Manager by the Roles Parser which loads whole configuration from the XML file at the start of the system. This approach reduces time which would be needed to parse an XML file each time Authorization Manager makes a decision.

12.3.4.1 Trust State Machine

Trust State Machine handles User Security Levels which determine permissions defined within a role, a user has been granted. USL is being increased after successful completion of the authentication step and it is decreased by Threat Monitor (described in Sect. 12.3.5) in some situations. Figure 12.6 shows the configuration of the USLs used in the sample implementation. Each of the levels has a number, minimal amount of points which are required to stay on this level and an initial amount of points which are granted to the user when he enters the level. Trust State Machine holds the information about users points and his current Security Level. It provides methods to set a new USL (used by Authentication Manager), get the current USL (used by Authorization Manager and Authentication Manager) and to change amount of users points (used by Threat Monitor) which can result in the change of the current USL. The configuration of USLs is loaded at the system start and provided by USL Configuration Parser similarly as it is done in Authorization Manager.

```

<suspiciousActions xmlns="http://www.w3schools.com"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3schools.com suspiciousActionsConfig.xsd">
  <role name="DEVELOPER">
    <action name="unsuccessfulLogins" value="1" points="3" />
    <action name="timeWithoutAction" value="20" points="5" />
    <action name="notPermittedRequest" value="1" points="3" />
  </role>
  <role name="ADMINISTRATOR">
    <action name="unsuccessfulLogins" value="1" points="4" />
    <action name="timeWithoutAction" value="30" points="5" />
    <action name="notPermittedRequest" value="1" points="5" />
  </role>
  <role name="HR">
    <action name="unsuccessfulLogins" value="1" points="3" />
    <action name="timeWithoutAction" value="20" points="2" />
    <action name="notPermittedRequest" value="1" points="8" />
  </role>
</suspiciousActions>

```

Fig. 12.7 Example configuration of the suspicious actions

12.3.5 Threat Monitor

Threat Monitor realizes the main concept of this work—adaptive authorization. Its role is to monitor all user’s requests and react on them according to the provided configuration. As was described in Sect. 12.2.1 each of the user’s requests is being checked against his behavior profile and the list of suspicious actions. User has different behavior profiles depending on the class of the used device. After each session user’s behavior profile is being updated to better adapt to user’s recent habits. Those updates however take into account previous profile, to prevent of rapid changing of a profile what could be a result of a masquerade attack. The other technique used to prevent such attacks are the suspicious actions lists. Those lists are defined for roles and differently than behavior profiles are unmodifiable.

Suspicious actions list and user’s behavior profile are managed by Suspicious Actions Manager and User Behavior Manager respectively. Threat Monitor consults them to obtain required information for specified user. Because suspicious actions are defined for roles and are unmodifiable, they are stored in an XML file which is presented in Fig. 12.7.

Each role has actions which are considered as suspicious. Each action has to have provided a name, a value and a number of points which will be deducted from the users account for this action. The “value” attribute can have different meanings depending on the name of the action. It is Suspicious Actions Manager’s role to recognize and handle appropriately all suspicious actions configured in the XML file. The second component used by Threat Monitor—user behavior profiles—are being stored in the database however they are created according to the configuration which is similarly to the suspicious actions defined in the XML file.

Each DF profile class has specified behaviors which should be monitored by Threat Monitor. Each of the behaviors in the XML file has to have provided name, points and variance attributes. Similarly to the suspicious actions, name of the behavior has to

be recognized by User Behavior Manager. Points have the same meaning but instead of the value, which is different for each user, configuration of the user's profile allow to define tolerance to the difference between current value and the value stored in the user's profile.

Due to the fact, that most of the suspicious actions and user's behaviors are of the different type and should be checked in different situations, Threat Monitor, Suspicious Actions Manager and User Behavior Manager are strongly coupled. That means that in current implementation design of Threat monitor is dependent of the suspicious actions and user behaviors which it should monitor.

The main difference between them is that suspicious actions are defined and rigid while user's profile may change during the time what changes a definition of abnormal behaviors. The motivation to use this combined approach was described in [8]. Each action/behavior defined on those lists can have different severity. That means for each action different number of points can be subtracted from user's account. Each security level has threshold of points. When user's current amount of points is lower than required by his current security level then the level is being lowered. The differences in points between specified security levels do not have to be constant to create more sensitive security levels (especially the highest ones).

The abnormal behaviors and suspicious actions may vary in different applications depending on their structure and purpose. In the example application developed for the purpose of this study they contain following entries: different working hours, too many actions in specified time interval, different proportions in access to different resources types, number of unsuccessful attempts during last authentication process, exceeded time without any action, attempt to access resources which are forbidden for user's role.

12.4 Preliminary Tests

Preliminary tests shows that system reacts correctly and blocks the user who became a threat. By flooding system with requests to different resources user started to behave differently than it was defined in his behavior profile and by trying to access unavailable resources he performed suspicious actions. Depending on a role and a device class, the user was blocked sooner or later what is correct and results from the configuration. Another scenario tested systems resistance to the unauthorized access attempts when villain tries to pass authentication process using a brute force method. In this case it is assumed that attacker obtained an access to the users device but has problems with passing the first authentication step. For all roles and classes the test runs looked as follows: requested specified resources, multiple unsuccessful login attempts, successful login, get blocked. The system works in a way that it decreases user's points related to the unsuccessful login attempts after user finally passes this step. Depending on the user's role more or less unsuccessful attempts are needed to block user's account. All the preliminary test runs show that system reacts correctly and blocks the user before he is able to access any resources.

12.5 Conclusions

Having regard to the new trends and changing working conditions that require more traveling and mobility and in the same time minimizing a risk related with these requirements, the solution described in this paper was designed. Introduction of adaptive authentication and authorization let users to omit a complex authentication process when it is not required but in the same time protects the classified data for being misused even by authorized employees.

This approach not only increases users' performance and facilitate their work but also improves company's overall security. The designed system complements currently available solutions, joins them and extends to respond to the new requirements like use of multiple devices by one user.

As the example implementation and performed test scenarios showed, taken approach works and all theoretical assumptions are valid and possible to implement. This of course does not mean that a universal solution has been created, the current implementation still has many aspects that should be improved, nevertheless this paper was meant to show the direction in which new security solutions should head to adapt to new demands.

References

1. Cisco annual security report (2011) [Online]. Available at: www.cisco.com/en/US/prod/collateral/vpndevc/security_annual_report_2011.pdf
2. Todorov, D.: *Mechanics of User Identification and Authentication: Fundamentals of Identity Management*, 1st edn. AUERBACH, Boca Raton (2007)
3. Smith, R.: *Authentication—From Passwords to Public Keys*. Addison-Wesley (2002)
4. Mendyk-Krajewska, T., Mazur, Z.: Problem of network security threats. In: 3rd Conference on Human System Interactions, May 2010, pp. 436–443 (2010)
5. Miller, K., Voas, J., Hurlburt, G.: BYOD: security and privacy considerations. *IT Prof.* **14**(5), pp. 53–55 (2012) [Online]. Available at: <http://dx.doi.org/10.1109/MITP.2012.93>
6. Gkaraffi, S., Economides, A.A.: Comparing the proof by knowledge authentication techniques. *Int. J. Comput. Sci. Secur. (IJCSS)* **4**(2) (2011)
7. Sekulski, R., Woda, M.: Adaptive method of users authentication and authorisation. In: *Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014*, 10–12 February 2014, South Kuta, pp. 174–176. ISBN: 978-0-9924518-0-6
8. Venkatesan, R., Bhattacharya, S.: Threat-adaptive security policy. In: *IEEE International Performance, Computing, and Communications Conference, IPCCC 1997*, February 1997, pp. 525–531 (1997)
9. Bailey, C., Chadwick, D.W., Lemos, R.D.: Self-adaptive authorization framework for policy based RBAC/ABAC models. In: *Proceedings of the IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, SER*. pp. 37–44. IEEE Computer Society, Washington (2011) [Online]. Available at: <http://dx.doi.org/10.1109/DASC.2011.31&apos>
10. Gao, J., Zhang, B., Ren, Z.: A dynamic authorization model based on security label and role. In: *IEEE International Conference on Information Theory and Information Security (ICITIS)*, pp. 650–653 (2010)
11. *Cambridge Advanced Learner's Dictionary*, Cambridge University Press, Cambridge (2005)

Chapter 13

Estimation of Quality of Service in Stochastic Workflow Schedules

Michal Wosko and Jan Nikodem

Abstract This paper investigates the problem of estimating the quality of a given solution to a workflow scheduling problem. The underlying workflow model is one where tasks and inter-task communication links have stochastic QoS attributes. It has been proved that the exact determination even of the schedule length distribution alone is #P-complete in the general case. This is true even if the problems of processor-to-task allocation and inter-task communication are abstracted away, as in program evaluation and review technique (PERT) approaches. Yet aside from the makespan, there are many more parameters that are important for service providers and customers alike, e.g., reliability, overall quality, cost, etc. The assumption is, as in the distributed makespan problem, that all of these parameters are defined in terms of random variables with distributions known *a priori* for each possible task-to-processor assignment. This research provides an answer to the open question of the complexity of the so formulated problem. We also propose other than naive Monte Carlo methods to estimate the schedule quality for the purpose of, e.g., benchmarking different scheduling algorithms in a multi-attribute stochastic setting. The key idea is to apply to a schedule a novel procedure of transformation into a Bayesian network (BN). Once such a transformation is done, it is possible to prove that, for a known schedule, the problem of determining the overall QoS is still #P-complete, i.e., not more complex than the distributed PERT makespan problem. Moreover, it is possible to use the familiar Bayesian posterior probability estimation methods, given appropriately chosen evidence, instead of a blind Monte Carlo approach. As the schedules are usually required to satisfy well-defined QoS constraints, it is possible to map these to appropriately chosen conditioning variables in the generated BN.

M. Wosko (✉)

FICT, Swinburne University of Technology, Melbourne, Australia
e-mail: mwoosko@swin.edu.au

J. Nikodem

Faculty of Electronics, Wroclaw University of Technology, Wroclaw, Poland
e-mail: jan.nikodem@pwr.edu.pl

13.1 Introduction

The service industry must nowadays come to terms with an increasing complexity of workflow scheduling problems, with the arbitrary structure of the involved workflows, the multiplicity of the workflow instances co-existing and competing on a scheduling platform, the heterogeneity of the platform and volatility of its elements being some of the challenges. Another significant challenge is the fact that the execution of any single instance is unavoidably affected by random events and so must be also the performance of a given schedule with respect to any objective for which such randomness applies. In general, there exists a complex relationship between the performance for a global objective, such as the length or the quality of a schedule, and the variables (decisions) that describe the (atomic) task-to-processor allocations. The “classical” scheduling approaches, both offline parallel processing heuristics (see, e.g., Drozdowski [4], Kwok and Ahmad [10]) and online stochastic scheduling policies (see, e.g., Niño-Mora [14] or Pinedo [16]), usually limit their area of interest and optimization capability to a single execution time-related objective, such as makespan, flowtime of the (single) application or the weighted completion time of its tasks. Multi-objective scheduling problems may have their performance targets defined, again, with respect to a single resource, e.g., when makespan and flowtime are to be optimized simultaneously, or with respect to a set of conceptually disjoint QoS attributes, e.g., time, quality, cost, etc. A non-extensive list of solutions for both cases includes metaheuristics such as stochastic algorithms and linear (LP) or mixed non-linear programming (MNLP) approximations and tailored heuristic solutions. In general, stochastic algorithms explore a small part of the actual decision state space and of the corresponding performance space defined with respect to the global objective functions. The LP/MNLP-approximation metaheuristics explore an approximate, reduced state space and the performance space corresponding to the approximation. The tailored heuristics (sometimes called “ad-hoc” heuristics) explore a decision state space that is locally defined at each step and optimize a likewise locally defined set of objectives which, with various degrees of belief, is likely to contribute to optimizing the set of global objectives. An obvious example is the most common list scheduling heuristic approach (see, e.g., Graham [6], Munier et al. [13]). In this approach, the global makespan minimization problem, which is *NP*-complete, is transformed into a series of locally defined, polynomial problems of finding the processor allocation for the currently considered task that guarantees its earliest finish time.

Historically, the first results on the complexity and estimation of distributed QoS values were published for the problem of the aggregated execution time distribution of projects with uncertain task durations, modeled by activity-on-arc (AoA) graphs in the area of program evaluation and review technique (PERT). A (stochastic) PERT problem is essentially a $P \infty | v_i \sim \text{stoch}, c_{ij} = 0, \text{prec} | *$ problem in the Graham-Allahverdi notation scheme [1, 2, 7], and a PERT “schedule length” is computed under the assumption of greediness. In other words, there are, by definition, no communication delays and an activity that is ready (all of its predecessors are completed) can, and always is, executed immediately, without waiting for a processor to

become free or for any algorithm logic-related reason. Therefore the PERT problem is simplified with respect to, e.g., a $P|v_i, c_{ij} \sim \text{stoch}, \text{prec}|*$ problem, as it actually does not require any non-trivial scheduling decision, such as the choice of the task-to-processor allocations. Yet the evaluation of the resulting schedule length alone is known to be #P-hard [8, 9]. Some estimation techniques to get around this problem were proposed, among others, by Schmidt and Grossmann [19], Ludwig et al. [12]. For the problem of estimating the length of a schedule with data dependencies and actual allocation of tasks onto a bounded-capacity heterogeneous computing platform, see, e.g., Li and Antonio [11].

This paper extends the approach presented in Wosko and Nikodem [20]. It answers the open question of the complexity of the problem of computing exactly and efficiently approximating distributed QoS values for given schedules, when not only a single resource (time) is involved and a single objective (makespan, with its specific aggregation function) is estimated based on known relationships of stochastic variables. It is structured as follows: Sect. 13.2 provides assumptions and formal definitions; Sect. 13.3 outlines a procedure for constructing a Bayesian network (BN) corresponding to the known schedule and the aggregation functions of the QoS attributes of interest and for simplifying it, by eliminating variables when this can be done efficiently. Section 13.4 presents a proof that, given this representation, the problem is the same as exact inference in a BN, or, equivalently, 3-SAT, and therefore still #P-hard. Section 13.5 shows how some of the proven BN analysis techniques can be adopted for specific QoS estimation problems, when approximation is allowed. Notably, a limited number of samples can be generated for so defined BN structures without recourse to a full-scale Monte Carlo approach. This is the case, when certain QoS are required to satisfy additional constraints: these can be used to condition the other probabilities in the network. Section 13.6 presents conclusions and an outlook of possible future directions.

13.2 Problem Formulation

In the following definitions, the application and schedule model used in the subsequent reasoning is briefly outlined.

Definition 13.1 An application is a tuple $\mathbb{G} = (J, E)$, where $J = \{j_1, \dots, j_k, \dots, j_n\}$ is a set of $|J| = n \in \mathbb{N}^+$ jobs and $E \subset J \times J$ is a set of precedence constraints.

A precedence constraint between two jobs j_i, j_k , indicated by $j_i < j_k$, or e_{ik} , means that job j_i must be completed before job j_k can be started. A precedence constraint may be associated with a communication link between two jobs j_i, j_k , indicated as $l(j_i, j_k) = l_{ik}$.

Definition 13.2 Given a set of jobs J and a set of communication links L , entailed by an application \mathbb{G} , and given a set of processors P and communication channels $H \subseteq P \times P$, a (complete) schedule \mathbb{S} is an allocation of (all) jobs in J to (a subset of) P and of the required communication links in L to (a subset of) H .

Formally, let $p(j) \in P$ be the processor assigned to j and let $S(j, p) = S_j$, $C(j, p) = C_j \in \mathbb{R}^+ \cup \{0\}$ be the start and finish time of a job j on $p(j)$. Then a schedule is a tuple $\mathbb{S} = (\{S_j\}, \{C_j\}, \{p(j)\})$.

Definition 13.3 A schedule \mathbb{S} is feasible if and only if it satisfies the following conditions:

- release dates are respected: $S_j \geq r_j$
- precedence constraints and communication links, if applicable, are respected: $\forall (i, j) \in E \Rightarrow S_j \geq C_i + c_{ij} \mathbf{1}_{P \setminus \{p(i)\}}(p(j))$, where:

$$\mathbf{1}_{P \setminus \{p(i)\}}(p(j)) = \begin{cases} 1 & \Leftrightarrow p(j) = p(i) \\ 0 & \text{otherwise} \end{cases}$$

A schedule is additionally complete, if each job has (at least) one processor assigned to it: $\forall j \in J \Rightarrow \exists p(j) \in P$ ($J \rightarrow P$ is surjective).

In the above definition the communication link is assumed to be *optional* in the sense that there are schedules in which it is not enforced. Specifically, those are the schedules where the two connected jobs are scheduled to the same processor, which is expressed by the function $\mathbf{1}_{P \setminus \{p(i)\}}(p(j))$. As noted, it always implies a precedence constraint (clearly, the converse is not true) and in fact it generalizes the notion of a precedence constraint, by assigning it a set of real-valued functions (or probability distributions). In most scheduling models with communication, the communication links required by a schedule, unlike job assignments, are allowed to overlap in terms of time intervals. A full schedule (solution), when considering all activities, i.e., all jobs and the required (by that solution) communication links, is given by the tuple $\mathbb{S}_a = (\{S_a\}, \{C_a\}, \{p(a)\})$, where a is any activity and the rest of the notations is analogous to those in the above Definition 13.2, with $p(a)$ appropriately chosen from the set $P \cup H$.

Let ρ be a non-renewable resource and $w(\rho)$ be a quantity associated with this resource, with a value in $\mathbb{R}^+ \cup \{0\}$. In general, this value may be a constant known a priori (i.e., have a deterministic value), or be a variable with a known range or probability distribution, or a value unknown until the occurrence of an event. The symbol $w(\rho, a)$ indicates that the quantity of the resource intended is also associated with, and consumed by, an activity a . In the following, it is assumed that $w(\rho, a)$ is a random variable with an a priori known probability distribution, i.e., it may not be unknown. On the other hand, the case of deterministic values, which correspond to degenerate distributions, is omitted as uninteresting for the proposed approach.

In single-attribute scheduling there is only one non-renewable resource: $\rho = \rho_t$ typically represents the execution time of the activities (processing time for jobs, communication or data transfer delay for communication links). In multi-attribute schedules, which are of direct interest to this approach, there exists a set of non-renewable resources $\{\rho_i\}$ and, for each activity a (job or communication link), there is corresponding set of random variables $W(a) = \{w(\rho_i, a)\}$.

As it is well known, different resources have different aggregation functions $f_g(\rho)$. For example, the execution time ρ_t is additive along a sequence and conforms to the

max operator for a join of two parallel sequences. A sequence of two activities may be derived from the application, i.e., be due to a precedence constraint, or be the result of a scheduling decision, as for, e.g., two unrelated jobs allocated to the same processor one after the other. Quality and reliability, on the other hand, are often assumed to be aggregated along a sequence on a product basis and conform to the *min* operator for parallel sequences joint together: see, e.g., Zeng et al. [21]. The sequences in this case are only those derived from the precedence constraint set entailed by the application, not by the juxtaposition of unrelated tasks on a processor, which is irrelevant.

Finally, different objective functions also have their corresponding aggregation functions $f_o(\gamma_k)$, that take as arguments the values of $\{f_g(\rho_i)\}$.

13.3 Solution Approach

In Sect. 13.1 a substantial amount of research dedicated to the question of the complexity and the estimation of the makespan in the PERT problem is examined. This problem, which constitutes originally the core PERT problem in a stochastic environment, is superseded in any context where multiple stochastic attributes are relevant for the determination of a schedule and the evaluation of its quality. The complexity of the computation of the (distributed) multi-attribute quality of service associated with an *a priori* given schedule, i.e., with a solution to a scheduling derived by an offline algorithm, has been an open question. This section provides a method to determine this complexity, which is based on an algorithm (shown in Algorithm 6) that transforms any offline schedule abiding to the assumptions formulated in Sect. 13.2 into a Bayesian network. A Bayesian, or belief network (BN, also known as probabilistic or causal network, or knowledge map) is a directed acyclic graph having random variables as nodes and representing a joint distribution of such variables (see, e.g., Pearl [15] or Russel et al. [18]).

The time complexity of the algorithm can be evaluated as follows. The first two nested loops generate as many operations as the product of the size of the resource set and the size of the mapping set (the number of the partial schedules in the complete schedule). The second pair of nested loops produces as many operations as the total number of different aggregations, derived for all the objective functions, of the partial schedules contained in the offline schedule.

With some particular objective functions such as makespan and (weighted) flow time, the BN obtained can be partially compacted by merging nodes that represent truly independent random variables that are individually irrelevant and have to be added for the evaluation of the objective function. In other words, in many cases the schedule can be reformulated in a reduced and re-usable representation, by compacting all the linear chains of tasks that are aggregations of independent random variables: this is trivial for chain-additive aggregation functions, such as makespan or flowtime, as mentioned, while it requires integration or its discrete equivalent for other aggregations.

Algorithm 6: Construction of a Bayesian network from a schedule.

```

1 input: a schedule  $\mathbb{S}_a = (\{S_a\}, \{C_a\}, \{p(a)\}) = \{M(a)\}$ ; a set of resource aggregation
   functions  $\{f_g(\rho_i)\}$ ; a set of objective aggregation functions  $\{f_o(\gamma_k)\}$ 
2 output: a Bayesian network
3 foreach resource  $\rho_i$  do
4   foreach activity mapping  $m(a) = (S_a, C_a, p(a))$  do
5     add a network node referring to  $m(a)$ 
6     derive the set  $M_{parent}(a, \rho_i)$  of the parent mappings of  $m(a)$  under the resource
       aggregation  $f_g(\rho_i)$ 
7     add a network link to all nodes referred to by  $M_{parent}(a, \rho_i)$ 
8 foreach objective function  $\gamma_k$  do
9   add a network node  $n(\gamma_k)$  referring to  $\gamma_k$ 
10  derive the set  $N_o(\gamma_k)$  of the parent nodes of  $n(\gamma_k)$  under the resource aggregation  $f_o(\gamma_k)$ 
11  foreach node in  $N_o(\gamma_k)$  and resource  $\rho_i$  do
12    add a network link to  $n(\gamma_k)$ 
13    derive the set  $N_{parent}(N_o(\gamma_k), \rho_i)$  of the parent nodes of  $N_o(\gamma_k)$  under the resource
       aggregation  $f_g(\rho_i)$ 
14    add a network link to all nodes referred to by  $N_{parent}(N_o(\gamma_k), \rho_i)$ 

```

13.4 Complexity of the Multi-Attribute QoS

In the previous section, an algorithm is presented that transforms a schedule computed by an offline algorithm into a Bayesian network that represents the stochastic events and their relationships relevant for the chosen set of objective functions as nodes and links in a Bayesian network. After applying the set of compacting operations briefly mentioned in Sect. 13.3, a network is obtained that is irreducible.

If the schedule can be represented by a Bayesian network, the determination of the probability distribution of the value for any node is of the same complexity as exact inference in a BN, or, equivalently, computing the number of satisfying assignments in a propositional logic, i.e., the 3-SAT problem (see, e.g., Russell et al. [18]).

In a multiply connected network, the inference is #P-complete, as established by Roth [17], following the determination, by Dagum and Luby [3], that such inference is at least NP-hard. Dagum and Luby also proved that even approximate inference is NP-hard: this is true for both absolute and relative inference, however *without any additional input*. The network generated by Algorithm 6 is multiply connected in the general case (the proof of this obvious fact is omitted: a simple example can be formulated to show it). Therefore, the stochastic multi-attribute QoS computation problem is still (at worst) #P-complete, like the distributed PERT makespan problem. This proves the main claim of our paper.

Note that, for most realistic objective functions and all multi-attribute objective functions, the constructed BN is, in fact, multiply connected. Therefore, the BN representation does not make easier the *exact* computation of the distributed QoS (nor does it make easier the absolute or relative approximation in the sense of Dagum

and Luby's definitions). However, it makes other, sample-based approximation techniques developed for Bayesian networks applicable, with the prospect of a computationally more efficient estimation of the QoS.

13.5 Approximation

In the previous section, the complexity of the exact distributed QoS computation is determined by an application of the appropriate results pertaining to the complexity of exact inference in Bayesian networks. As briefly mentioned, sample-based approximate inference methods seem, however, to provide a way of reducing the complexity of the problem of approximating the value range of the distributed QoS.

Clearly, this problem, given the semantics of the employed Bayesian network representation, may be formulated as follows: to provide estimates for the probabilities of interest, i.e., those associated with the nodes in the network that correspond to the values of the objective functions of interest. An unformed randomized sampler, such as a pure Monte Carlo algorithm, would produce such estimates with an accuracy dependent on the number of samples generated for the independent variables. Obviously, this approach is either inefficient or inaccurate but it is also unlikely to be superseded by a better method in the absence of any additional input.

The situation changes, if such input is available. The BN representation makes it possible to employ evidence-based inference algorithms, that are the true strength of the BN research field in general. Our idea is simple, yet novel: the problem of estimating distributed QoS values for a schedule is usually defined in a context, from which *additional constraints* may be derived, that, in turn, may be *treated as evidence* for the purpose of the sampling algorithm. For instance, the quality (in the sense of the example in Sect. 13.2) of the schedule needs to be estimated under the assumption that the total makespan is contained in a range of values. In addition, given the BN representation, the knowledge of the particular scheduling problem may be easily employed to restrict the values of the quality attribute for particular nodes "inside" the network, representing the intermediate task assignments in the schedule. This paves the way for the application of particularly efficient evidence-driven approximate BN inference algorithms, such as the Gibbs sampler [5] or another variant of the Markov Chain Monte Carlo method. To the best of the authors' knowledge, a comparable estimation method does not exist for any standard, unprocessed representation of a schedule such as the one in Sect. 13.2, which is the typical output of a scheduling algorithm.

13.6 Conclusions and Future Work

In this work, we investigated the problem of the evaluation of quality of service (QoS) values for a multi-attribute multi-objective (MAMO) schedule computed by an offline scheduling algorithm, given that its inputs are subject to stochastic variability.

We applied a novel approach of manageable complexity that transforms the input schedule, given the set of objective functions and attribute aggregation functions, into a Bayesian network (BN).

The results are twofold. Firstly, given this representation, it is easy to prove that the MAMO QoS computation problem is still, in the worst case, #P-complete, like the familiar PERT stochastic makespan problem. Secondly, we have shown how, by identifying the concepts of schedule constraint and evidence variable in a Bayesian network, efficient evidence-based approximate inference algorithms developed for BNs can be used to estimate the QoS values of interest.

Acknowledgments The authors wish to thank Smart Services CRC for partially funding this work.

References

1. Allahverdi, A., Gupta, J.N., Aldowaisan, T.: A review of scheduling research involving setup considerations. *Omega* **27**(2), 219–239 (1999)
2. Allahverdi, A., Ng, C., Cheng, T.E., Kovalyov, M.Y.: A survey of scheduling problems with setup times or costs. *Eur. J. Oper. Res.* **187**(3), 985–1032 (2008)
3. Dagum, P., Luby, M.: Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artif. Intell.* **60**(1), 141–153 (1993)
4. Drozdowski, M.: *Scheduling for Parallel Processing*, 1st edn. Springer Publishing Company, Incorporated, New York (2009)
5. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 721–741 (1984)
6. Graham, R.L.: Bounds for certain multiprocessing anomalies. *Bell Syst. Tech. J.* **45**(9), 1563–1581 (1966)
7. Graham, R.L., Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.: Optimization and approximation in deterministic sequencing and scheduling: a survey. *Ann. Discret. Math.* **5**, 287–326 (1979)
8. Hagstrom, J.N.: Computational complexity of PERT problems. *Networks* **18**(2), 139–147 (1988)
9. Hagstrom, J.N.: Computing the probability distribution of project duration in a PERT network. *Networks* **20**(2), 231–244 (1990)
10. Kwok, Y.-K., Ahmad, I.: Static scheduling algorithms for allocating directed task graphs to multiprocessors. *ACM Comput. Surv. (CSUR)* **31**(4), 406–471 (1999)
11. Li, Y.A., Antonio, J.K.: Estimating the execution time distribution for a task graph in a heterogeneous computing system. In: *Proceedings of 6th Heterogeneous Computing Workshop, (HCW'97)*, pp. 172–184. IEEE (1997)
12. Ludwig, A., Möhring, R.H., Stork, F.: A computational study on bounding the makespan distribution in stochastic project networks. *Ann. OR* **102**(1–4), 49–64 (2001)
13. Munier, A., Queyranne, M., Schulz, A.S.: Approximation bounds for a general class of precedence constrained parallel machine scheduling problems. *Integer Programming and Combinatorial Optimization*, pp. 367–382. Springer, Berlin (1998)
14. Niño-Mora, J.: Stochastic scheduling. In: Floudas, C.A., Pardalos, P.M. (eds.) *Encyclopedia of Optimization*, 2nd edn, pp. 3818–3824. Springer, New York (2009)
15. Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo (1988)
16. Pinedo, M.: *Scheduling: Theory, Algorithms, and Systems*. Springer Science+ Business Media, New York (2012)
17. Roth, D.: On the hardness of approximate reasoning. *Artif. Intell.* **82**(1), 273–302 (1996)

18. Russell, S.J., Norvig, P., Canny, J.F., Malik, J.M., Edwards, D.D.: *Artificial Intelligence: A Modern Approach*, vol. 74. Prentice Hall, Englewood Cliffs (1995)
19. Schmidt, C.W., Grossmann, I.E.: The exact overall time distribution of a project with uncertain task durations. *Eur. J. Oper. Res.* **126**(3), 614–636 (2000)
20. Wosko, M., Nikodem, J.: Estimation of QoS in stochastic workflow schedules. In: *Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014*, 12th February 2014, pp. 75–76 (2014)
21. Zeng, L., Benatallah, B., Ngu, A.H., Dumas, M., Kalagnanam, J., Chang, H.: QoS-aware middleware for web services composition. *IEEE Trans. Softw. Eng.* **30**(5), 311–327 (2004)

Chapter 14

Multi-Party System Authentication for Cloud Infrastructure by Implementing QKD

Roszelinda Khalid, Zuriati Ahmad Zukarnain, Zurina Mohd Hanapi
and Mohamad Afendee Mohamed

Abstract In this paper, we will highlight the enhances technique to authenticate multi-party system in cloud infrastructure. We propose an enhanced tight finite key scheme for Quantum Key Distribution (QKD). With this technique we believe it can provide a secure channel between a cloud client to establish a connection between them by applying quantum theories. As a final result it shows, that it can reduce the possibility of losing a private key by producing a high efficient key rate and attack resilient. Key length is an important security measures.

14.1 Introduction

Cloud infrastructure requires new ways of providing security. This includes deployment of Public Key Infrastructure (PKI). PKI is actually a mechanism where they exchange key using certificates and revocation list have the capabilities to authenticate users in the cloud infrastructure [1]. In security triad confidentiality, integrity and authentication are the major point in this cloud infrastructure. So looking into that circumstances, we found that in PKI authentication, the public key cryptography only provide computational security. This is because PKI is based on Asymmetric Key Cryptography. It vulnerable to security threats such as eavesdropping, man in the middle attack, masquerade and etc. Which means, attacker easily actuate the persons private key. This is prone to information disclosure. On the other hand is the loss of private key may be irreplaceable. In this case, the received messages will not be decrypt if the private key is missing. This circumstances, highlight the importance of authentication technique involving multiple users. Hence, in any communication mechanism is a must for us to ensure the safeness. Therefore, we propose the

R. Khalid (✉) · Z.A. Zukarnain · Z.M. Hanapi · M.A. Mohamed
Faculty of Computer Science and Information Technology, University Putra Malaysia,
UPM Selangor, Selangor, Malaysia
e-mail: roszelinda@hotmail.com

Z.A. Zukarnain
e-mail: zuriati@upm.edu.my

Multi-party Quantum Key Distribution (MQKD) protocol together with finite key technique. In this paper [2] they introduce quantum secure authentication with classical key that provides a solution of quantum readout. Basically the quantum readout happen in physical keys that easily can be exploit. Thus, by approaching the classical key, they still need to combine the classical channel and quantum channel. As what we all know, quantum cryptography is providing a platform to distribute a key and not to transmit the message. The crucial part is how we are going to confirm that the key cannot be tampered by unauthorized party. Padmavathi et al. in [3] proposed to use group signature base. The group signature base will be verify by third party so called as Trusted Center to distribute the key among members. It seems like the implementation of Public Key Infrastructure. However their research still lacking in terms of attack resilient. Of all the review we had made, authentication is base of our research. Authentication is a process determining whether someone or something is declared to be. As a classical approach authentication scheme is based on client server architecture. Furthermore quantum cryptography could be part of this authentication process. The authentication phase will happen in quantum key distribution. In the previous paper [4] have shown the significant of the distribution of final key in two conditions. Which are with attack scenario and without attack scenario. It is proved the proposed method in [4] could be improve more. On the other hand, each user in the communication channel will be authenticate via their secret key [5]. Multi-party Quantum Key Distribution is actually a key distribution protocol that establish a common key among a number of users. Our proposed solution is to authenticate Multi-party Quantum Key Distribution protocol tied up with an enhanced tight finite key scheme that formally introduces by Tomamichel [6]. This may eliminate the percentage of missing a private key. We acknowledge this can be done by yield high efficient key rate and attack resilient into the proposed method.

14.1.1 Authentication Mechanism

Authentication is a well studied area. In basic idea authentication is a process of determining whether someone or something is, in fact, who or what it is declared to be. Today's transmission security relies on the unproven computational security. In every communication task, transmission of the data or information the security triad will be the main criteria to look into. It must have confidentiality, integrity and availability. One standard cryptographic task is authentication. This is an important task to be done prior to communication that guarantees that the user identification and the origin of data is genuine because, if a malicious user masquerades as a legitimate user, the key distribution schemes and encryption schemes will be easily compromised. In Table 14.1 it shows the existing authentication schemes. Authentication scheme known as digital challenge response. Throughout this paper we are proposing an adaptation of digital cryptography challenge response mechanism on quantum channel. The protocol shown in the table is basically used broadly in a classical approach.

Table 14.1 Digital challenge-response authentication schemes (adopted from [7])

Protocol	Characteristics	Mechanism	Advantages	Limitation
Challenge-handshake authentication protocol (CHAP)	Authenticates a user or network host to an authenticating entity	Shared secret key, one way hash function, three way handshake	Protection against: replay attack	distribution of secret key
CRAM-MD5	SMTP mail agent authentication	Hash function, fresh concatenation, fresh random challenge	Resist to replay attack	Lack of mutual authentication, Storage of password, Vulnerable to dictionary attack
Kerberos	Network authentication protocol over insecure channel	Issue tickets, trusted third party	Mutual authentication, resist to replay attack and eavesdropping	Single point of failure, vulnerable to man-in-middle attack. Time constraints
Otway-Rees protocol	Network authentication protocol over insecure channel	Usage of one session identifier server	Resist to replay attack and eavesdropping	Vulnerable to intercept and resend attack
Needham-Schroeder protocol	Network authentication protocol over insecure channel	Shared secret key, server	Establish session key and mutual authentication	Key distribution
Wide mouth frog protocol	Network authentication protocol over insecure channel	Global clock, server, shared secret key, BAN logic	Resist replying attack and eavesdropping. Detection of modification	Key distribution and required trusted server
CAPTCHA, reCAPTCHA	Network user identification non-cryptographic scheme	Images	Widely used in webmail	Availability
Distance-bounding protocol	Cryptographic protocols that enable a verifier V to establish an upper bound on the physical distance to a prover P	Delay time, radio frequency implementation	This approach was a breakthrough to thwart relay attacks by measuring the round trip time of short authenticated messages	The computation is based on the fact that electro-magnetic waves travel nearly at the speed of light, but cannot travel faster
Physical unclonable function or PUF	PUF is a function that is embodied in a physical structure and is easy to evaluate but hard to predict	Hardware implementation of hash function	Resist to spoofing attacks	Practical implementation and generality

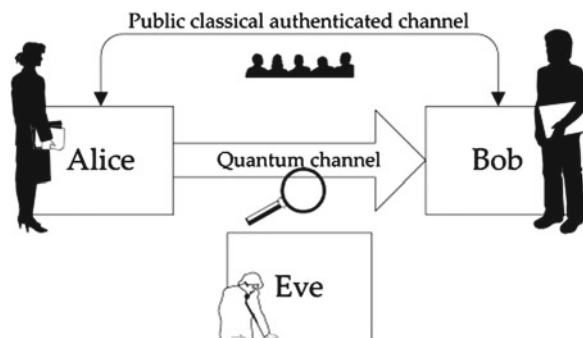
14.1.2 Quantum Cryptography

By definition cryptography itself can be describe as the art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret key can decipher or decrypt the message into plain text. Encrypted messages can sometimes be broken by cryptanalysis, also called code breaking, although modern cryptography techniques are virtually unbreakable. The difference between traditional cryptography and quantum cryptography is in quantum cryptology it only rely on physics and not mathematics. But unlike traditional cryptology methods, encoding and decoding information or messages, quantum cryptology depends on physics, not mathematics.

Quantum cryptography uses our current knowledge of physics to develop a cryptosystem that is not able to be defeated—that is, one that is completely secure against being compromised without knowledge of the sender or the receiver of the messages. The word quantum itself refers to the most fundamental behavior of the smallest particles of matter and energy: quantum theory explains everything that exists and nothing can be in violation of it. Quantum cryptography is different from traditional cryptographic systems in that it relies more on physics, rather than mathematics, as a key aspect of its security model. As in Fig. 14.1, it explain the composition of quantum channel and a public authenticated channel. Distribution of the key will happen in quantum channel where else the transmission on information still in the public channel. Because of this composition we can identify the existing of eavesdropping.

Essentially, quantum cryptography is based on the usage of individual particles waves of light photon and their intrinsic quantum properties to develop an unbreakable cryptosystem essentially because it is impossible to measure the quantum state of any system without disturbing that system. It is theoretically possible that other particles could be used, but photons offer all the necessary qualities needed, their behavior is comparatively well-understood, and they are the information carriers in optical fiber cables, the most promising medium for extremely high-bandwidth communications.

Fig. 14.1 Quantum key distribution comprises a quantum channel and a public classical authenticated channel. As a universal convention in quantum cryptography, Alice sends quantum states to Bob through a quantum channel. Eve is suspected of eavesdropping on the line



Back to cryptography, we will imagine the things that play with big numbers. Then it will involve the process of encryption and decryption. We want to hide our conversation or our information from the people that we might neglect them to know. After a several decade previous researchers are introducing steganography. Why are they doing this? All is to hide the information. In a huge network, the concept of public key infrastructure is in used. Previously in a huge network industry is using a public key infrastructure also being called as PKI. It is a platform that enables users to securely exchange data or information via the use of public and private cryptographic matching key. Trusted authority will play the role to distribute the key among the certified user. Basically unsecure public network such as the Internet to securely and privately exchange data and money through the use of a public and a private cryptographic key pair that is obtained and shared through a trusted authority in the PKI. The public key infrastructure provides for a digital certificate that can identify an individual or an organization and directory services that can store and, when necessary, revoke the certificates. Although the components of a PKI are generally understood, a number of different vendor approaches and services are emerging. Meanwhile, an Internet standard for PKI is being worked on. How the public key infrastructure work is shown in table below (Table 14.2).

Quantum cryptography is the only known method for transmitting a secret key over distance that is secure in principle and based on the laws of physics. Current methods for communicating secret keys are all based on unproven mathematical assumptions. These same methods also are at risk of becoming cracked in the future, compromising today’s encrypted transmissions retroactively. This matters very much if you care about long-term security. Yet in [8] the researcher introducing the Information-theoretically Secure (ITS) authentication. They believe there is a need in Quantum Key Distribution. Their study is on security of an ITS authentication scheme proposed by Wegman and Carter, in the case of partially known authentication key. This scheme uses a new authentication key in each authentication attempt, to select a hash function from an Almost Strongly Universal hash function family. The partial knowledge of the attacker is measured as the trace distance between the authentication key distribution and the uniform distribution; this is the usual measure in QKD. Then they provide direct proofs of security of the scheme, when using partially known key, first in the information-theoretic setting and then in terms of witness indistinguishability as used in the Universal Composability (UC) framework. The result shows is interesting because it can increase the level of security and the percentage of error rate correction is less.

Table 14.2 How public key infrastructure work

Work	Who	Type of key
Send an encrypted message	Use the receiver’s	Public key
Send an encrypted signature	Use the senders’s	Private key
Decrypt and encrypted message	Use the receiver’s	Private key
Decrypt and encrypted signature (and authenticate the sender)	Use the sender’s	Public key

14.1.3 How Quantum Key Distribution Work in Quantum Cryptography

In quantum key distribution there are several protocol that being test on the field. The pioneer is BB84 protocol that introduced back in the year of 1984 by Charles and Bennet. Figure 14.2 shows how the protocol work. In the initial states Alice uses a light source to create a photon. Then the photon is sent through a polarizer and randomly given one of four possible polarization and bit designations. The notation be given as Vertical (one bit), Horizontal (zero bit), 45 degree right (one bit), or 45 degree left (zero bit). After that the photon travels to Bob’s location. In this phase Bob has two beamsplitters which are diagonal and vertical/horizontal and two photon detectors. Bob randomly chooses one of the two beamsplitters and checks the photon detectors. The process is repeated until the entire key has been transmitted to Bob. Bob then tells Alice in sequence which beamsplitter he used. Alice compares this information with the sequence of polarizers she used to send the key. Alice tells Bob where in the sequence of sent photons he used the right beamsplitter. Now both Alice and Bob have a sequence of bits (sifted key) they both know. All in all, a pretty cool way of securely transferring an encryption key between two different locations. Figure 14.2 shows the scenario of quantum key distribution process.

14.1.4 Behind Quantum Cryptography Story

In the history of cryptography, quantum cryptography is a new and important chapter. It is a recent technique that can be used to ensure the confidentiality of information transmitted between two parties, usually called Alice and Bob, by exploiting the counterintuitive behavior of elementary particles such as photons. The physics of elementary particles is governed by the laws of quantum mechanics, which were

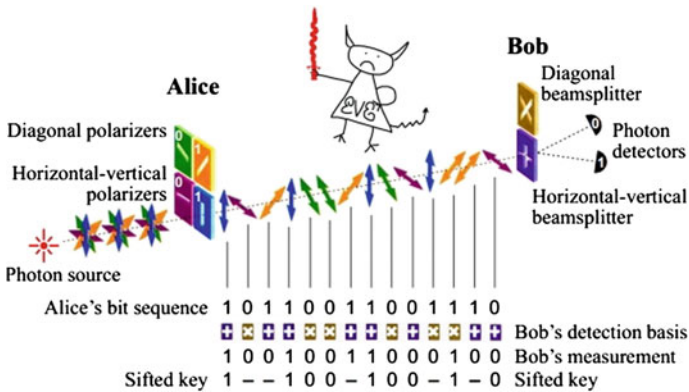


Fig. 14.2 Quantum key distribution scenario

discovered in the early twentieth century by talented physicists. Quantum mechanics fundamentally change the way we must see our world. At atomic scales, elementary particles do not have a precise location or speed, as we would intuitively expect. An observer who would want to get information on the particle’s location would destroy information on its speed and vice versa as captured by the famous Heisenberg uncertainty principle. This is not a limitation due to the observer’s technology but rather a fundamental limitation that no one can ever overcome. The uncertainty principle has long been considered as an inconvenient limitation, until recently, when positive applications were found [9]. In the meantime, the mid-twentieth century was marked by the creation of a new discipline called information theory. Information theory is aimed at defining the concept of information and mathematically describing tasks such as communication, coding and encryption. Pioneered by famous scientists like Turing and von Neumann and formally laid down by Shannon, it answers two fundamental questions: what is the fundamental limit of data compression, and what is the highest possible transmission rate over a communication channel?

Shannon was also interested in cryptography and in the way we can transmit confidential information. He proved that a perfectly secure cipher would need a secret key that is as long as the message to encrypt. But he does not say how to obtain such a long secret key. This is rather limiting because the secret key needs to be transmitted confidentially, e.g. using a diplomatic suitcase. If we had a way, say a private line, to transmit it securely, we could directly use this private line to transmit our confidential information. Figure below depicted the previous explanation (Fig. 14.3).

Since the seventies and up to today, cryptographers have found several clever ways to send confidential information using encryption. In particular, classical ciphers encrypt messages using a small secret key, much smaller than the message size. This makes confidentiality achievable in practice. Yet, we know from Shannon’s theory that the security of such schemes cannot be perfect. Shannon defined information as a mathematical concept. Nevertheless, a piece of information must somehow be stored or written on a medium and, hence, must follow the laws of physics. Landauer was one of the first to realize the consequences of the fact that any piece of information ultimately exists because of its physical support. Shannon’s theory

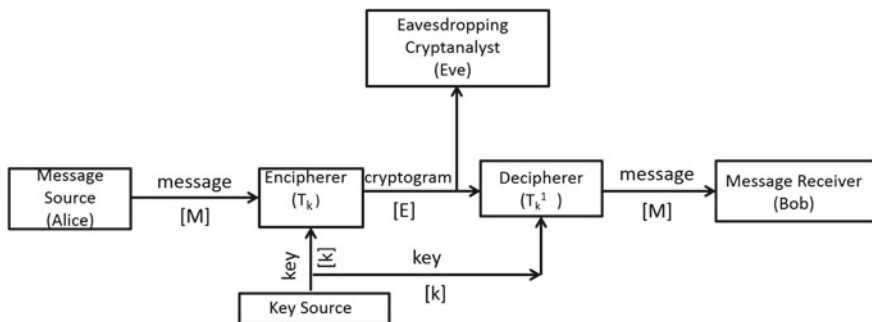


Fig. 14.3 Originated theory from Shannon

essentially assumes a classical physical support. When the medium is of atomic scale, the carried information behaves quite differently, and all the features specific to quantum mechanics must be translated into an information-theoretic language, giving rise to quantum information theory.

The first application of quantum information theory was found by Wiesner in the late sixties. He proposed using the spin of particles to make unforgeable bank notes. The spin of a particle satisfied the uncertainty principle: an observer cannot get all the information about the spin of a single particle; he would irreversibly destroy some part of the information when acquiring another part. By encoding identification information on bank notes in a clever way using elementary particles, a bank can verify their authenticity by later checking the consistency of this identification information. At the atomic scale, the forger cannot perfectly copy quantum information stored in the elementary particles; instead, he will unavoidably make mistakes. Simply stated, copying the bank note identification information is subject to the uncertainty principle, and thus a forgery will be distinguishable from a legitimate bank note.

Other applications of quantum information theory is quantum computer. Quantum computer uses quantum principles instead of the usual classical principles, can solve some problems much faster than the traditional computer. In a classical computer, every computation is a combination of zeroes and ones (i.e. bits). At a given time, a bit can either be zero or one. In contrast, a qubit, the quantum equivalent of a bit, can be a zero and a one at the same time. In a sense, processing qubits is like processing several combinations of zeroes and ones simultaneously, and the increased speed of quantum computing comes from exploiting this parallelism. Unfortunately, the current technologies are still far away from making this possible in practice.

Following the tracks of Weisner's idea, Bennett and Brassard proposed in 1984 a protocol to distribute secret keys using the principles of quantum mechanics called quantum cryptography or more precisely quantum key distribution [10]. By again exploiting the counterintuitive properties of quantum mechanics, they developed a way to exchange a secret key whose secrecy is guaranteed by the laws of physics. Following the uncertainty principle, an eavesdropper cannot know everything about a photon that carries a key bit and will destroy a part of the information. Hence, eavesdropping causes errors on the transmission line, which can be detected by Alice and Bob. Quantum key distribution is not only based on the principles of quantum physics, it also relies on classical information theory. The distributed key must be both common and secret. First, the transmission errors must be corrected, whether they are caused by eavesdropping or by imperfections in the setup. Second, a potential eavesdropper must know nothing about the key. To achieve these two goals, techniques from classical information theory, collectively denoted as secret-key distillation, must be used.

Unlike the quantum computer, quantum key distribution is achievable using current technologies, such as commercially available lasers and fiber optics. Furthermore, Shannon's condition on the secret key length no longer poses any problem, as one can use quantum key distribution to obtain a long secret key and then use it classically to encrypt a message of the same length. The uncertainty principle

finds a positive application by removing the difficulty of confidentially transmitting long keys. State-of-the-art ciphers, if correctly used, are unbreakable according to today's knowledge. Unfortunately, their small key size does not offer any long-term guarantee. No one knows what the future will bring, so if clever advances in computer science or mathematics once jeopardize today's ciphers' security, quantum key distribution may offer a beautiful alternative solution. Remarkably, the security of quantum key distribution is guaranteed by the laws of quantum mechanics. Furthermore, quantum key distribution guarantees long-term secrecy of confidential data transmission. Long-term secrets encrypted today using classical ciphers could very well become illegitimately decryptable in the next decades. There is nothing that prevents an eavesdropper from intercepting an encrypted classical transmission and keeping it until technology makes it feasible to break the encryption. On the other hand, the key obtained using quantum key distribution cannot be copied. Attacking the key means attacking the quantum transmission today, which can only be done using today's technology. In some reason we can conclude that quantum cryptography and quantum key distribution are similar. For others, however, quantum cryptography also includes other applications of quantum mechanics related to cryptography, such as quantum secret sharing. A large portion of these other applications requires a quantum computer, and so cannot be used in practice. On the other hand, the notion of key is so central to cryptography that quantum key distribution plays a privileged role.

14.1.5 Chronology of Cloud Infrastructure Security

There are some concerns that should not be taken lightly when moving to a cloud service. Once the data has been moved to a cloud provider, control over it has been lost. The user cannot tell where the data resides physically and cannot be fully confident the data is handled with care in a secure manner. Providers are very much aware of the complications and concerns of their services and works constantly to improve the quality and security of their services. Despite all these, researcher are eagerly looking for the best mechanism in order to provide a high level of security. As being discusses in earlier section, the wide implementation of security authentication in cloud is Public Key Infrastructure. However, there are three issues that can complicate the implementation of PKI on cloud. In PKI environment there are three factors to consider when designing the system. There are scalability, mobility and automation. A solution must be able to add more CAs which is Certificate Authority on demand, be relatively consistent in required time to sign certificates and always be available. Hence, the solution must support the CA operations being movable to another less strained server if the number of requested signatures increases beyond the limit of the Hardware Security Module or the service unexpectedly fails. To able to move all CA operations to another server, all data regarding that CA must be moved between databases and the private key has to be moved or be the same at the new location. However, there exists no sufficiently secure procedure to move private keys between

HSMs which is hardware security module autonomously. Therefore, the same private keys must be predefined in HSMs at all available locations of that CA. The ability to move the CA to another location and to bind private keys on demand provides scalability in the number of signatures the system can handle. The scalability of the number CAs at one location is relative to the number of keys the Hardware Security Module is able to store. Then the second issue is one essential requirement of a cloud based PKI is that one customer should only be able to see and use its own CAs. Consequently, there must be separation between CAs and customers. Only a number of predefined CAs can issue certificates to administrators due to the trust store in the application server. Other CAs issuing administrator certificates can be added but that requires restarting of the application server. The purpose of this is to give each customer a dedicated CA to issue certificates to its administrators.

14.2 Objective

Our objective is to design an efficient multi-party QKD scheme for multi-party system. We believe the use of QKD protocol such as BB84 could help enhance the security level for communication purposes. In [11] shows that BB84 protocol that being introduce by Bennet and Brassard in 1984 is secure and reliable. Even though there are other protocol appear after BB84 we believe it still cant beat it. So, through out this paper, the journey towards the realization of the proposed method we will use BB84 protocol. The main idea of the proposed scheme is the use multiparty key and reduced the utilization of quantum communication. Moreover, the distance and loss factors are handled using this multiparty key and one-way communication among the parties. Further, this research concentrates on analysis of key size by tight finite-key and security analysis with imperfect device assumption. We are looking forward to eliminate the possibility of losing a private key. This objective can be achieve by generating a high efficient key rate and attack resilient. A multi party quantum key distribution with tight finite key based on Shannon entropy and von Neumann theory is proposed. In contrast with scheme proposed by Tomamichle that is using the same theory as a base. The proposed scheme is also easier to implement and highly efficient without losing in security level. This quantum communication will involve a sender and multiple receivers.

14.3 Scope

This research is geared to authenticate multi-party Quantum Key Distribution protocol using a tight finite key scheme. This research focus on producing an appropriate key size using a quantum protocol that is practicable and doable within a cloud infrastructure. We will use our own enhance tight finite key using a theory from Von Neumann, Shannon Entropy and Shor's Algorithm in order to prevent any kind of

attack. Shannon entropy is the interpretation of information through measurement. Meanwhile Von Neumann usually used to measure the quantum information in entanglement state. However, both theories will give with BB84 protocol. By the way, for Shor's algorithm we use to generate prime numbers for creating a hash function and key. We simulate the cloud environment so that we can experiment the proposed scheme. The simulation platform is using *C#* programming language.

14.4 Proposed Methodology

We are considering to involve more parties in the key distribution process instead of two parties. This is because our method will be tied up in cloud infrastructure that may involve more than two parties. In other words we are going to improve key distribution process by adopting some classical concepts and quantum techniques. Therefore to carry out the objectives, we branched it into several phases. In initial phase, Alice as person who will initial a communication session will enter her secret key. This phase will acknowledge how many party will involve in this communication. Authentication process will take into place. Then in the next phase all the private key will convert into common key. System will automatically convert the value insert by Alice into hash value then convert into binary. The established algorithms in QKD such as BB84 will be used to convert all private keys into a common key. Then the private key will be shared among parties involved in the multi-party system. Before the connection establish among the party, each party will authenticate with multi party QKD protocol. We have to make sure that only 25 % error is allowed before the establishment. Comparing key with the raw key that give by Alice to the system and other party will bring to privacy amplification phase. Lastly if the condition is meet, the system will produce a secure key and distribute among the involving parties. The key distribution and creation is held in quantum channel. In order to achieve the high security level, challenge response to all user in first phase will be used. The execution of proposed method will execute in both public channel and quantum channel. As we mention earlier, the key distribution will be held in quantum channel and the message will transfer using a public channel. The security analysis for this proposed multi-party QKD authentication protocol is carried out by getting the value of final key length and attack resilient. This multi-party QKD authentication software tool using finite-key size will be simulated using a simulator develop with *C#* programming language.

14.5 Result

After experimenting with the proposed technique, we manage to get some positive result. By applying tight finite key in quantum key distribution, the problem of man in the middle attack can be solved. We categorized it as attack resilient. Comparing

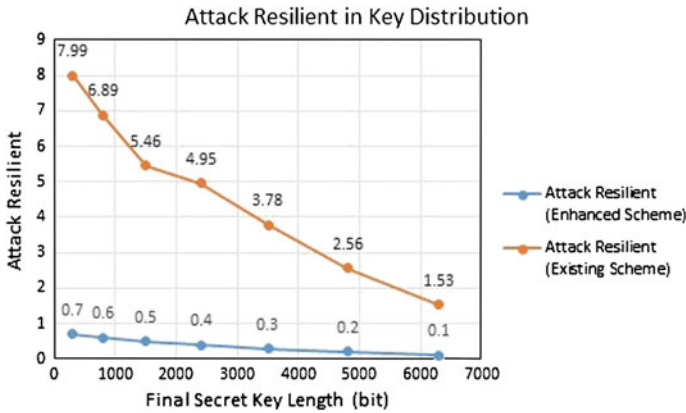


Fig. 14.4 Attack resilient using enhanced tight finite key scheme for quantum key distribution

the enhanced scheme with existing scheme there is a highly impact on the proposed method. To analyze the attack resilient value, we formulate a formula as below:

$$attack\ resilient = \frac{initial\ key\ length - final\ key\ length}{initial\ key\ length} \tag{14.1}$$

Figure 14.4 shows that the final secret key size do reflect the attack resilient. The key size below 200 bit can be successfully attack by intruders. The shorter key size, it is easier for intruder to steel the key. However with our enhance scheme, it is proven that there is an improvement in terms of attack resilient. We put the highest score that prone to attack is 10. From the graph it is merely shows that it is impossible for the intruder to steal of guess the final key. In this research we are trying to keep the initial key size until the end of the session which mean the connection establishment. This propose method is briefly being discussed in [12] during the 2nd Asia-Pacific Conference on Computer Aided System Engineering, APCASE 2014 (Fig. 14.4).

14.6 Conclusion

The main idea of presenting the quantum protocol in cloud infrastructure is to provide a safe platform of establish a communication between more than two parties. Therefore, in order to minimize damages or losses due to security threats, a reliable and robust key distribution protocol and communication channel is very much in demand. The proposed framework will be able to provide the needed capability in a timely manner. There are many appliance, application and devices available commercially, however the proposed framework is believed to be able to reduce any possible threat that may harm the data transmission. Many researchers are looking for the best method to secure the communication by being tempted by third party. The focus

mainly goes to key size. We are trying to find out and proof that our proposed method are manage to produce the suit key size that hard to break. This research is still on going.

Acknowledgments We acknowledge the generous help from Faculty of Computer Science And Information Technology, University Putra Malaysia. This work is supported by Malaysia Ministry Education under Fundamental Research Grant Scheme (FRGS).

References

1. Chouhan, D.S., Kharche, H.: Building trust in cloud using public key infrastructure. *Int. J. Adv. Comput. Sci. Appl.* **3**, 26–31 (2012)
2. Goorden, S.A., Horstmann, M., Mosk, A.P., Škorić, B., Pinkse, P.W.H.: Quantum-secure authentication with a classical key, p. 287 (2013). arXiv preprint [arXiv:1303.0142](https://arxiv.org/abs/1303.0142)
3. Padmavathi, V., Madhavi, M., Nagalakshmi, N.: An approach to secure authentication protocol with group signature based quantum cryptography. *Int. J. Innov. Technol. Explor. Eng.* **2**, 105–107 (2013)
4. Khalid, R., Zulkarnain, Z.A.: Enhanced tight finite key scheme for quantum key distribution (QKD) protocol to authenticate multi-party system in cloud infrastructure. *Appl. Mech. Mater.* **481**, 220–224 (2013)
5. Yuan, H., Zhou, J., Zhang, G., Yang, H., Xing, L.: Efficient multiparty quantum secret sharing of secure direct communication based on bell states and continuous variable operations. *Int. J. Theor. Phys.* **51**(11), 3443–3451 (2012)
6. Tomamichel, M., Lim, C.C.W., Gisin, N., Renner, R.: Tight finite-key analysis for quantum cryptography. *Nat. Commun.* **3**, 634 (2012)
7. Buhari, A., Ahmad, Z.Z., Subramaniam, S.K., Zainuddi, H., Saharudin, S.: A quantum based challenge response user authentication scheme over noiseless channel. *Int. J. Netw. Secur. Appl.* **4**(6), 67–79 (2012)
8. Abidin, A., Larsson, J.-Å.: Direct proof of security of Wegman-Carter authentication with partially known key. *Quantum Inf. Process.* **13**(10), 2155–2170 (2013)
9. van Assche, G.: *Quantum Cryptography and Secret Key Distillation*. Cambridge University Press (2012). 978-0-521-86485-5
10. Bennett, C.H., Brassard, G., et al.: Quantum cryptography: public key distribution and coin tossing. In: *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, vol. 175(150), p. 8 (1984)
11. Shor, P.W., Preskill, J.: Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.* **85** 441–444 (2000)
12. Roszelinda, K., Ahmad, Z.Z.: Multi-party system authentication for cloud infrastructure by implementing QKD. In: *2nd Asia-Pacific Conference on Computer Aided System Engineering-APCASE*, pp. 52–53 (2014)

Chapter 15

SmartCloud Orchestrator—the First Implementation for Education in the World at WrUT

Jerzy Greblicki, Jerzy Kotowski, Mariusz Ochla and Jacek Oko

Abstract SmartCloud Orchestrator provides a consistent, flexible, and automated way of integrating the cloud with customer data center policies, processes, and infrastructures across various IT domains, such as backup, monitoring, and security. It proposes the intuitive, graphical tool to define and implement business rules and IT policies. One can connect the aspects of different domains into a consistent orchestration of automated and manual tasks to achieve the business goals. SmartCloud Orchestrator is based on a common cloud platform that is shared across IBM's Cloud offerings. This common cloud stack provides a common realization of the core technologies for comprehensive and efficient management of cloud systems. Last year, at the end of November, Wrocław University of Technology, due to co-operation with IBM Corporation, put this tool into the motion for education purposes, between others.

15.1 Orchestrator—General Information

Orchestrator is a workflow management solution for the data center. Orchestrator lets to automate the creation, monitoring, and deployment of resources in our environment. Cloud orchestration solutions from IBM are designed to reduce the IT management complexities introduced by virtual and cloud environments and accelerate cloud service delivery, allowing enterprises to quickly respond to changing business needs.

J. Greblicki · J. Kotowski (✉) · J. Oko
Wrocław University of Technology, Wrocław, Poland
e-mail: jerzy.kotowski@pwr.wroc.pl

J. Greblicki
e-mail: jerzy.greblicki@pwr.wroc.pl

J. Oko
e-mail: jacek.oko@pwr.wroc.pl

M. Ochla
IBM Poland, Warsaw, Poland
e-mail: mariusz_ochla@pl.ibm.com

IT administrators perform many tasks and procedures to keep the health of their computing environment up-to-date and their business running. Tasks might include the diverse activities. Individual tasks and subtasks are automated, but typically, not for the whole process [1–3].

By using Orchestrator, we can carry out the following tasks: Automate processes in the data center, regardless of hardware or platform, automate your IT operations and standardize best practices to improve operational efficiency and connect different systems from different vendors without having to know how to use scripting and programming languages.

Orchestrator provides tools to build, test, debug, deploy, and manage automation in your environment. The standard activities defined in every installation of Orchestrator provide a variety of monitors and tasks with which it is possible to integrate a wide range of system processes.

Any IT organization can use Orchestrator to improve efficiency and reduce operational costs to support cross-departmental objectives. Orchestrator provides an environment with shared access to common data. By using Orchestrator, an enterprise can evolve and automate key processes between groups and consolidate repetitive manual tasks. It can automate cross-functional team processes and enforce best practices for incident, change, and service management.

At Pulse 2013, The Premier Cloud Conference, IBM announced the Beta of IBM SmartCloud Orchestrator, based on OpenStack, Topology and Orchestration Specification for Cloud Applications (TOSCA) and Open Services for Lifecycle Collaboration (OSLC). SmartCloud Orchestrator constitutes IBM's new unified and open cloud management platform, consisting of three main layers. The infrastructure services layer is based on OpenStack for provisioning, configuring and managing storage, compute and network resources. The platform services layer includes virtual machine image lifecycle management capabilities and pattern services. The latter refers to IBM's so-called patterns of expertise, which include exact deployment and management instructions for the entire business service. The orchestration services layer is based on IBM's Lombardi acquisition, offering an easy-to-use business process management solution. IBM announced that later in 2013 the platform services layer and the services orchestration layer will both support the TOSCA standard. SmartCloud Orchestrator supports OSLC for continuous delivery across heterogeneous development environments and is able to deploy workloads to a software-based private cloud, IBM's integrated PureSystems or public clouds such as Amazon EC2 or IBM's SmartCloud Services.

15.2 IBM Smarter Cloud Orchestrator at Wrocław University of Technology

Future education will be not only widely available, but also much smarter than today. Due to a rapid growth of technologies, among which cloud computing plays a key role, the education supporting systems will use all available information in order to

personalize the process of teaching. IBM says it is already happening and predicts a fast development of such technologies within 5 years.

The revolution of process of teaching will affect all levels of education, not only academia. The adoption of new mentality as well as the development of necessary systems will obviously take some time, however it is particularly academia's role to lead the process [4, 5].

All the changes in education will not be possible without a flexible system that would support cooperation between all types of schools and interactions with businesses. IBM Smarter Cloud Orchestrator (SCO) at Wrocław University of Technology (WrUT) is a base on which such systems for the entire region of Lower-Silesia will be built. The first phase of the project consists of implementation of an IT Cloud for education at the University with the architecture and technology already prepared to support future extension.

A private cloud for education, a flexible cloud for the university vital administration systems now, the system built on IBM SCO will become a core engine for Lower Silesian Educational Cloud. As IBM says “The classroom of the future will learn about each individual student...” WrUT has implemented SCO to build a flexible cloud system that will support the idea of modern and personalized education. For this moment let us state that virtual platform of the system is based on four servers IBM Flex System x240. At our University System SmartCloud Orchestrator was installed on virtual machines with operational system Red Hat Enterprise Linux 6.3 x64. The particular components of Cloud System are as follows:

- Smart Cloud Orchestrator.
- IBM Image Construction and Composition Tool.
- IBM Virtual Image Library.
- IBM Process Center.
- IBM Process Portal.

15.3 Lower Silesia Educational Cloud

This chapter presents the concept developed by Wrocław University of Technology on the execution and implementation of pilot of the services called “Lower Silesia Educational Cloud”—hereon called LSEC—in Polish “Dołnoślaska Chmura Edukacyjna”—DCE. The proposed concept assumes to adapt processing solutions in the cloud computing to the specifications and requirements of the educational sector. Next there is a description of the various functional blocks showing the hardware infrastructure. On the basis of the infrastructure virtualization layer (software layer) we will build, both hardware resources, to ensure access to computing power, and resources to ensure access to space.

15.3.1 Analysis of the State

Institutions and companies implementing IT systems performing defined goals often face the problem of non-productive use of hardware resources and human resources for implementation/deployment of the system [6, 7].

The information system reflects business processes in the institution/company. The business process in turn is a set of interrelated activities, which lead to the realization of a particular objective. Hallmarks of a business process include among others: duration, frequency and complexity.

Workstations operating in schools are also affected by hypertrophy, inconsistencies, bad management and maintenance of IT infrastructure. Most schools build their teaching facilities (in the field of computer science) based on incremental purchases that are not correlated with any unified concept, leading to the building of a cohesive work environment.

An inconsistent environment significantly affects the unit costs associated with the maintenance and management of infrastructure. In addition, the life cycle of a workstation in correlation with the time spent developing software (new versions of application software usually increase hardware requirements criteria) is a disadvantage economically, and affects any justification of the continued functioning of schools in the currently adopted model.

Another problem identified in the education sector is the lack of interoperability between different schools. Lack of this interoperability leads even to the creation of specific barriers preventing the free transfer of knowledge and good practice between schools/learning centers—schools, instead of drawing standard and proven solutions, develop their own.

The above-mentioned problems lead to hypertrophy of the infrastructure held by individual schools and maintenance problems[2, 5].

Another very commonly reported problem in the evaluational units is the collection, processing and dissemination of personal data in accordance with the Act on Personal Data Protection, along with the Regulation. Indicated Act applies to educational units, a number of restrictions and requirements, the implementation of which translates directly into the need to involve both its own means and resources.

15.3.2 Target LSEC Model

Model realization of Lower Silesia Educational Cloud should be considered at three levels:

1. Infrastructure—including the necessary ICT resources (hardware, licenses, and network resources) to implement the project, providing the required availability, reliability and quality. Infrastructure means all the hardware and software components implementing specific functionality in Cloud-Computing (IA as Papas,

SaaS, VI). Cloud Computing involves deliberate “loose” logical infrastructure linking the physical infrastructure, which foundation can be accomplished in several locations (Data Processing Centers).

2. Content and services—including the most important from the point of view of the end-user: range of teaching and learning materials and methods (e-services) its presentation, a group of services that will support the teaching process (e-labs, e-classrooms) and collaborate with existing resources and applications, and a group of services and mechanisms to support the management of educational units. Lower Silesia Educational Cloud will provide access to hardware and software resources (including Virtual Desktop Infrastructure) build within the frames of the project and access to produced e-services embedded in the Cloud Computing model. All e-services implemented under the project will be built around the core which guarantees mutual connection between them—this core will be a personalization access system based on a unique and secure identification (smart card) that provides visual and electronic identification of the holder. It is expected to implement the following groups of e-services:
 - a. (technology services) knowledge centers, infrastructure that provides the hardware re-sources, infrastructure allowing virtualization of workstations and laptops, public key infrastructure, electronic signature and
 - b. (Functional services) knowledge centers, access control, registration of school attendance, library card, travel card functionalities, access control to printing and copying, school locker key.
3. Maintenance and operation—including the organization of maintenance and operation of these two planes (i.e., infrastructure, services and content). In the classical model, an institution or company that implements the system has to bear the costs associated with the purchase of servers, storage arrays, network infrastructure (LAN, WAN, SAN). Often provided by the above-mentioned infrastructure, the potential (resource computing power, the resource space, etc.) is not fully used—mainly in the initial phase of development of services. In the next period further development of services can force investment in additional infrastructure that is capable of handling an unexpected rise in popularity of the service. The cycles linked to the growing demand for computing power service and storage space are in correlation with the provision of these resources as the service stations increasingly get popular.

The solution to the problems described above is to use cloud computing to create new services in cloud computing and in the transformation of already deployed services to the Cloud Computing model. Benefits to resolve the above problems are mainly due to the seven characteristics of cloud computing:

- Scalability
- Accessibility
- Measurability
- Ease of deployment

- Performance
- Security
- Savings

15.3.3 Project Organization

The concept involves the construction of Lower Silesia Educational Cloud in the following phases (while deciding on the shape and size of the target of the project is Phase 2):

Phase 1—Needs Analysis

Phase 2A—Implementation of a pilot

Phase 2B—Implementation of basic services (selected geographical area subjective)

Phase 3—Implementation of basic services (the rollout of basic services in the area of the Lower Silesia Province)

Phase 4—Implementation of the target model for multi-service (along with the model of self-provisioning).

15.4 Virtualisation System for Education—New Technologies Research at the Wrocław University of Technology

In our university we developed, as part of the research a system DVSE (Distributed Virtualisation System for Education). The basic idea is the ability to run VM tasks on computers (workstation) in the background and making them available for remote users. From the perspective of the remote user's job is done in virtualization system (such as the classical virtualizations task in CC). In the same time, the local user preserve all the functionality associated with the work locally (e.g., low-level access to the hardware, and work on the GPU). This approach requires effective algorithms for system monitoring and scheduling.

Methodology known from business VM scheduling problems do not fit to educational tasks. In light of this there is a need of novel design and analysis methodology for private Cloud for education. As it was mentioned, methods from industry do not apply to educational area (i.e. problem of performance parallelism). Main difference is life cycle of virtual machine.

The key problem in the task of planning the distribution of virtual machines is to ensure the specified performance at a specified time (i.e. working hours, time classes for students). Unfortunately we cannot handle the worst-case calculation time. Because in our model VM must be deployed on a single host and it is not allow for a virtual machine to use resources from several hosts also host changes are prohibited when machine is running. It is possible to run two and more than two virtual machines on one host.

To obtain a useful tool for the problem solving we put into the motion metaheuristic approach. Namely, for our purposes we reworked an optimization procedures based on the idea of Genetic Algorithm and Harmony Search Algorithm. Both approaches need for the very beginning an appropriate coding procedures to conduct necessary operators easy.

15.5 Summary and Future Plans

Cloud computing is a useful tool in the modern Information Technology. IBM has invited Wroclaw University of Technology to join its two centers of Cloud excellence.

We consider also virtualization (part of Cloud Computing) as an interesting tool for education as well as for research purposes and also topic of our researches. In this paper we have presented mathematical model of virtualization for universities educational tasks. Virtual machines deployment for educational purposes has other than business type restrictions.

One of them is an association of elite higher education institutions working in the Cloud Computing technologies called the IBM Cloud Academy. Wroclaw University of Technology and IBM Poland plan to continue collaboration in the field of education, PHD programs and research. IBM plans to initiate similar programs with other institutes, faculties and data centers in Poland, as well as to integrate with local business communities in the process.

References

1. Kotowski, J.: Internships as an application of cloud computing solutions for education at universities. In: *Computer Aided Systems Theory—EUROCAST 2013: 14th International Conference*, Las Palmas de Gran Canaria, Spain, 10–15 February 2013
2. Brzozowska, A.J., Greblicki, J., Kotowski, J.: Cloud computing in educational applications: methods of virtual desktops deployment. In: *Computer Aided Systems Theory—EUROCAST 2011: 13th International Conference*, Las Palmas de Gran Canaria, Spain, 6–11 February 2011
3. Greblicki, J., Kotowski, J.F.: Analysis of the Properties of the Harmony Search Algorithm Carried Out on the One Dimensional Binary Knapsack Problem. *Lecture Notes in Computer Science*, vol. 5717, pp. 697–704 (2009)
4. Cichon, A., Szlachcic, E., Kotowski, J.F.: Differential evolution for multi-objective optimization with self adaptation. In: *Proceedings of 14th IEEE International Conference on Intelligent Engineering Systems, INES 2010*, Las Palmas of Gran Canaria, Spain, 5–7 May 2010, pp. 165–169. IEEE, COP
5. Brzozowska, A.J., Greblicki, J., Kotowski, J.: State encoding and minimization methodology for self-checking sequential machines. In: *Computer Aided Systems Theory—EUROCAST 2011: 13th International Conference*, Las Palmas de Gran Canaria, Spain, 6–11 February 2011
6. Kotowski, J., Wilkocki, M.: *Dynamic Tivoli Data Warehouse*. Komputerowe wspomaganie badań naukowych XIX/red. Jan Zarzycki. Wroclaw: Wroclawskie Towarzystwo Naukowe (2012) (in polish)
7. Klempous, R., Kotowski, J.F., Nikodem, J., Ulasiewicz, J.: Optimization algorithms of operative control in water distribution systems. *J. Comput. Appl. Math.* **84**, 81–99 (1997)

Chapter 16

Cloud Computing—Effect of Evolutionary Algorithm on Load Balancing

Shahrzad Aslanzadeh, Zenon Chaczko and Christopher Chiu

Abstract In cloud computing due to the multi-tenancy of the resources, there is an essential need for effective load management to ensure an efficient load sharing. Depends on the structure of the tasks, different algorithms could be applied to distribute the load. Workflow scheduling as one of those load distribution algorithms, is specifically designed to schedule the dependent tasks on available resources. Considering a job as an elastic network of dependent tasks, this paper describes how evolutionary algorithm, with its mathematical apparatus, could be applied as workflow scheduling in cloud computing. In this research, the impact of Generalized Spring Tensor Model on workflow load balancing, in context of mathematical patterns have been studied. This research can establish patterns in cloud computing which can be applied in designing the heuristic workflow load balancing algorithms to identify the load patterns of the cloud network. Furthermore, the outcome of this research can help the end users to recognize the threats of tasks failure in processing the e-business and e-since data in cloud environment.

Keywords Cloud computing · Load balancing · Resource management

16.1 Introduction

Rapid improvement in today's technologies enabled businesses to grow more quickly. Cloud computing as one of the new emerging technology provides real time services for enterprises without binding them to their organizations. Accessing to the Internet with any devices that can connect to the Internet authorized the businesses to

S. Aslanzadeh (✉) · Z. Chaczko · C. Chiu
Faculty of Engineering & IT, University of Technology, Sydney, Australia
e-mail: Shahrzad.Aslanzadeh@uts.edu.au

Z. Chaczko
e-mail: Zenon.Chaczko@uts.edu.au

C. Chiu
e-mail: Christopher.Chiu@uts.edu.au

access their information at any time [1]. As an example Dropbox is one of these popular services that will accredit the access to the information while it can be easily synchronized and updatable at anytime [2]. By applying the virtualization techniques, cloud providers can minimize the costs of resource management process. As Creeger [3] highlights, cloud users can pay to access multiple resources at any time specially in peak hours, which will help them to grow more quickly, and rolling out among their competitors.

Cloud computing is the enhanced generation of grid computing. It refers to clusters of computers with the ability of dynamic provisioning in geographically distributed networks which is customisable on users requirements. Despite of the mentioned characteristics, cloud computing has one more advantages over grid computing. The ability of virtualization enables elasticity and scalability in cloud computing which can be considered as prominence of that over grid computing. Therefore according to virtualization concept, cloud computing can be explained as the collaboration of scalable and elastic virtualized resources that can be provisioned dynamically over the internet. Moreover cloud computing was named as utility computing. Cloud computing can offer variety of services on infrastructure, platform and software. It can bring profits for businesses by saving more money on their IT infrastructures. Cloud computing is a fifth utility after water, gas, electricity and telephone. It allows users to use its services according to their demands and without any constraint [4].

The overall aim of this research is to visualize the magnitude and direction of the load between workflow tasks and jobs in cloud computing. The visualization will mainly help in monitoring the cloud load to increase the availability of the resources while minimizing the response time. Moreover the visualization will be useful in terms of identifying the anomalies and threats in workflow applications which will lead to effective decision making.

In this research a Evolutionary workflow scheduling algorithm will be investigated to identify the load patterns within cloud network. The patterns to be investigated shall highlight the interconnectivity between tasks and jobs and shall enhance the recovery plan upon failure.

Reviewing the literature variety of load balancing algorithms have been proposed to balance the load in cloud computing by representing the static and dynamic movement of the tasks. But still there is a shortage of effective visualization tool to capture the anticipatory behaviour of the workflow tasks which can project the interactions and dependencies between workflow tasks.

16.2 Load Balancing in Cloud Computing

Cloud computing is composed of several different resources, interconnected to each other to form a network or a grid. These resources should be flexible and dynamic in terms of usage and allocations. In cloud computing, load balancing is one of the major techniques that has a dramatic impact on resource availability. The term availability, was always a main concern in cloud- computing. Fundamentally, availability explains

the ubiquitousness of the network information in case of resource scaling. Load balancing could be illustrated, as proper strategy for task scheduling that will lead to balanced load distribution in cloud networks. It is an important key to improve the network performance. Moreover load balancing algorithm can minimize the response time while utilizing the resource usage. The lack of proper load management can create traffics due to the long waiting time for accessing the resources. Today most of the cloud vendors are trying to use automated load balancer to enable the users, scale the numbers of their resources automatically. Promoting the availability and performance of the cloud system highlights the main goal of the automated load balancers. To design an effective load balancing algorithm, Dillon [5] suggested the following strategies:

- Load balancing algorithm should be smart enough to make load balancing decisions in a right time;
- Depends on behaviour of the application load balancer should be able to gather the information locally and globally;
- Load balancer should be designed in a centralized or distributed pattern. If the load balancer is centralized then there is a less opportunity for scalability purposes;
- Local load balancers are costing less, but the information provided by global load balancer is more accurate.

Reviewing the literature, different load balancing algorithms have been proposed to utilize the available resources. Efficient load balancing algorithm should be robust, and simple enough to be compatible with variety types of applications. The following points are defining a standard framework to design an effective load balancing algorithm [5]:

- *Complexity*: The algorithm should not be too complex as complexity will add more overhead on the system;
- *Scalability*: The algorithm should be scalable enough to manage all the existing services, if the network scaled up/down;
- *Fault tolerance*: The algorithm should be able to manage the load, even if any failure occurs in the network;
- *Performance and makespan*: The load balancing should be able to optimize the response time to enhance performance.

Load balancing methodologies are categorised into two main groups:

1. **Static load balancing** This method is mainly designed for homogenous and stable environments. Static load balancing algorithm cannot handle the dynamic load changes and that's why it cannot be used in real time systems.
2. **Dynamic load balancing** With this algorithm load balancer will manage the load dynamically at the run time. These algorithms are more flexible and will consider different attributes of the system before managing the load. Each of the static or dynamic algorithms could be divided into 4 different categories:
 - a. *Centralized versus Distributed*: In centralized model, scheduler has information about all the resources. In this model generally there is more controlling

over the resources and the implementation is much easier. However, in case of scalability and fault tolerance, centralized load balancing is not fully efficient. In distributed load balancing, there is no central controller for monitoring the nodes. Multiple schedulers can be used to help in scheduling the tasks. Distributed load balancing is suitable for scalable networks and it will support elasticity.

- b. *Preemptive algorithm versus non-preemptive*: Preemptive algorithm will allow jobs to be interrupted. As an example, if a low priority job changes to be a high priority job, then preemptive algorithm could be so practical. On the other hands, in non-preemptive task scheduling methods, no interruption is allowed until all the scheduled processes are completed.
- c. *Mediate versus batch mode*: In immediate mode, jobs will be assigned as soon as they arrive. So there is no waiting time for them. In batch mode, jobs will be grouped base on mapping criteria then each group will be assigned to proper resources for processing.
- d. *Independent versus workflow*: Workflow tasks are describing the tasks with some sort of dependencies. As an example finishing time of one task can be the start time of the other tasks. Most of the workflow tasks are represented with DAG graphs or Petri nets, and other language modelling tools such as XPDL and XML. On the other hand with independent scheduling approach, tasks could be assigned independently without considering the prerequisites for implementing a specific task.

Therefore we can summarise the load balancing benefits as follows:

- Load will be distributed evenly
- Processing time will be minimized
- Resource utilization will be maximized
- Availability of the system will be increased
- Performance will improve
- Resource will be more utilized
- Resource consumption will be minimized

Considering the benefits of the load balancing, there are some challenges that need to be addressed within load balancing concept.

Throughput Most of the loads balancing algorithms are trying to complete the highest numbers of tasks in given period of time. High throughput is an essential component to ensure the better performance of the system.

Response time The time that the load balancer needs to assign the tasks on available resources, for smaller response time, the performance is higher.

Overhead For each load balancing algorithm there is an associated overhead which is created by the process communications, tasks allocation and processor operation. The large overhead can impact the performance of the system.

Fault tolerance The load balancing algorithm should be able to find any node failure while allocating the tasks on available resources. **Scalability** As one of the specification of the cloud computing is related to its scalability, therefore each of the load balancing algorithm should be able to scale up/down base on the status of the network.

16.3 Proposition of Evolutionary Algorithms in Cloud Computing

16.3.1 Elastic Workflow Scheduling Model

Elasticity could be highlighted as one of the main characteristics of the cloud. In load balancing, elasticity could be explained as the ability of the system to allocate and reallocate the resources dynamically. Similarly in workflow load scheduling, elasticity could be interpreted in context of resource allocations for dependent tasks. In our proposed workflow load balancing, inspired from Fig. 16.1, tasks could be seen as an elastic network of mass spring model, in which each task is connected to its dependent with a spring. The rigidity of the springs will capture a standard pattern that could explain the impact of the task dependencies level on load balancing [6]. Hooks law, advised in 17th century by Robert Hook, elaborates that the expansion of a spring is proportional with the force that was imposed on it. The constant factor K highlights the elasticity ratio of the materials. In Hooks law elasticity is referring to the ability of the elastic body to return to its original status, after it was bent, stretched or squeezed [7]. If force is greater than the elasticity limitation of the material, it will cause a permanent deformation. Figure 16.2 is illustrating the spring behavior under force [8].

Fig. 16.1 Network of mass spring model

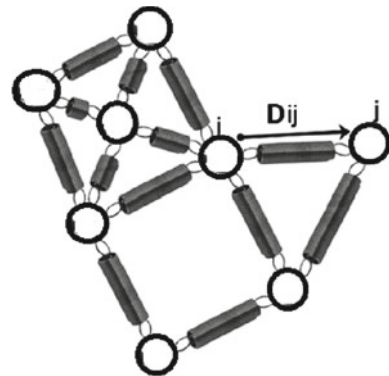
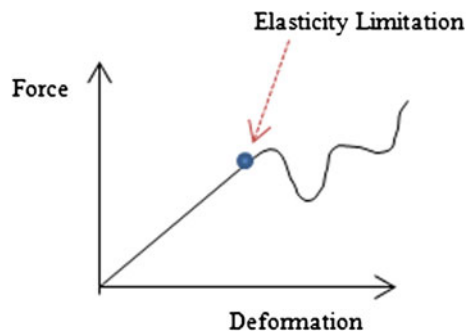


Fig. 16.2 Elasticity behavior



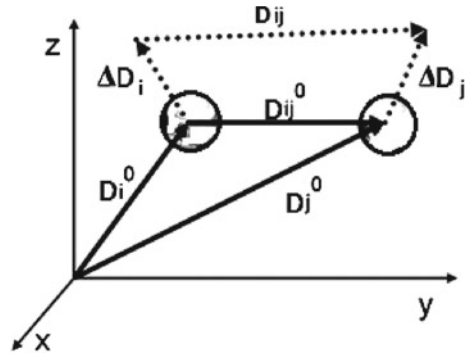
16.3.2 Generalized Spring Tensor Model in Cloud Computing

Various computational algorithms have been suggested to analyze the dynamics and complexity of the cloud. Among these algorithms, Coarse-gained algorithms could be highlighted as one of the important methods, developed to study the complication of the elastic networks. Elastic Network Model (ENM) is one of the coarse-gained models that particularly used in science to study the fluctuation and magnitudes of the proteins movement. Anisotropic network model (ANM) and Gaussian Network Model (GNM) are two different types of ENM algorithm [9, 10]. GNM is designed to analyze the B-factor values in network of proteins. B-Factor explains the protein dynamics by reflecting the magnitude of the protein fluctuation in related position. Moreover the fluctuation of the atoms is considered as Gaussian distributed along X, Y, and Z vector [10]. Furthermore, ANM model is able to evaluate the direction of the fluctuation in network of elastic nodes. In contrast with GNM model, ANM cannot explore the magnitude of the proteins fluctuations. In theory, ANM is complying with Generalized Hooks law which is featuring the relation between the stress and strain in 3 dimensions. Based on Hooks principal, imposing force P can cause deformation on elastic bodies. Inspiring from Hooks law, ANM removed the required energy minimization, and proposed a simpler version of Hooks law which could calculate the fluctuation direction of the load. As the fluctuation in this case is only limited to longitude axis of nodes i and j , therefore the magnitude of this fluctuation is senseless [11]. To overcome the above mentioned limitations of ANM and GNM algorithms, generalized spring tensor (STeM) was considered by Bahar [9] who explored the fluctuation and direction of the proteins simultaneously. Using Go-like algorithm as the potential model, Generalized spring tensor (STeM), a coarse-gained algorithm, is able to manage the network complexity by analyzing the magnitude and direction of the load [6]. In this model the interaction between two nodes i and j will be investigated but not in a linear structure. Next stage explains the mathematical apparatus of the STeM algorithm in more details.

16.3.3 STeM Algorithm and Workflow Load Balancing

Using STeM algorithm, the proposed workflow load balancing will highlight the direction and magnitude of the load fluctuations in cloud network. To manage the load efficiently the dependency ratio of the tasks, which was illustrated by the rigidity of the springs should be defined. Figure 16.3 is modeling two dependent tasks of i and j in 3D environment, where $D_i = (x_i, y_i, z_i)$ and $D_j = (x_j, y_j, z_j)$ are the positions of i and j in 3 Dimensions. Based on Hookeans principal, Force P could move an elastic body from its relaxed position by the value of d to the new state. The transition between original state and the new state could be impacted by constant

Fig. 16.3 Mass spring network



factor k . Based on the quality of the elastic body, the force could cause shearing, stretching or compression. According to Eq. 16.1 for some elastic bodies, if the directions of the stress and the strain are the same, then the magnitude of the force will be proportional [11, 12].

$$P = kd \tag{16.1}$$

However, if the force and displacement are not in a same direction the relation between stress and strain could be explained as Eq. 16.2:

$$P = k(\alpha d_1 + \beta d_2) \tag{16.2}$$

K is a second order tensor and α and β are real numbers with shifting values. Hooks law could also connect the stress and strain in three dimensional objects. In this sense, k could be explained as a 3×3 matrix that will be multiplied by movement of d in three dimensions to represent the applied force [13, 14]. Therefore the force vector will be depicted by Eqs. 16.3 and 16.4.

$$p = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \times \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \tag{16.3}$$

$$p_i = k_{i1}d_1 + k_{i2}d_2 + k_{i3}d_3 = \sum_{j=1}^3 k_{ij}d_j \tag{16.4}$$

Adhering to the same principal, and inspiring from Go-Like model, STeM algorithm, will use Hessian matrix, which is composed of four 3×3 matrices. These matrices are resulted from contribution of bond bending, angel, torsional and non-local interaction between dependent task. Each of these hessian matrices could be calculated with Eq. 16.5.

$$H_{ij} = \begin{bmatrix} \frac{\partial^2 v_{1,(r,r_0)}}{\partial X_i \partial X_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial X_i \partial Y_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial X_i \partial Z_j} \\ \frac{\partial^2 v_{1,(r,r_0)}}{\partial Y_i \partial X_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial Y_i \partial Y_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial Y_i \partial Z_j} \\ \frac{\partial^2 v_{1,(r,r_0)}}{\partial Z_i \partial X_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial Z_i \partial Y_j} & \frac{\partial^2 v_{1,(r,r_0)}}{\partial Z_i \partial Z_j} \end{bmatrix} \quad (16.5)$$

By combining the value of the four matrices, Eq. 16.6 is showing the potential correlation between two nodes:

$$(\Delta r_i) \cdot (\Delta r_i) = \frac{3k_\beta T}{\gamma} (H_{3i-2,3j-2}^\top + H_{3i-1,3j-1}^\top + H_{3i,3j}^\top) \quad (16.6)$$

As a future work, by using the correlation result of dependent tasks formulated in Matlab, we will highlight the areas with high dependencies percentage. This pattern could be used to predict better resource management plan and performance enhancement in cloud.

16.4 Conclusion and Future Work

This paper is proposing a new workflow load balancing algorithm that could be used in an elastic cloud. STeM algorithm has been suggested as the potential algorithm that could improve the load management by explaining the magnitude and direction of the fluctuation between dependent tasks. The model will help finding the level of the dependencies between each task by acknowledging the magnitude and direction of the load. Considering the behavior of the depended tasks, this approach will explain a pattern for managing the load balancing more efficiently. The expected benefits of the proposed algorithm will improve load balancing technique which could result in better performance rate. Moreover as the pattern will define the level of the tasks dependencies, a better fault tolerance and risk management could be predictable.

References

1. Khiyaita, A., Zbakh, M., El Bakkali, H., El Kettani, D.: Load balancing cloud computing: state of art. In: Network Security and Systems (JNS2), pp. 106–109 (2012)
2. Sawant, S.: A genetic algorithm scheduling approach for virtual machine resources in a cloud computing environment (2011)
3. Vöckler, J.-S., Juve, G., Deelman, E., Rynge, M., Berriman, B.: Experiences using cloud computing for a scientific workflow application. *Condor* **300**, 15–24 (2011)
4. Zhang, C., De Sterck, H., Jaatun, M., Zhao, G., Rong, C.: CloudWF: a computational workflow system for clouds based on Hadoop. *Cloud Comput.* **5931**, 393–404 (2009)
5. Galante, G., de Bona, L.C.E.: A survey on cloud computing elasticity. In: IEEE Fifth International Conference on Utility and Cloud Computing (UCC), pp. 263–270 (2012)

6. Lin, T.-L., Song, G.: Generalized spring tensor models for protein fluctuation dynamics and conformation changes. *BMC Struct. Biol.* **10**(Suppl. 1), S3 (2010)
7. HowStuffWorks ‘Elasticity’: <http://science.howstuffworks.com/dictionaryphysics-terms/elasticity-info.htm>
8. Hooke’s Law elasticity limitation: <http://www.clickandlearn.org/physics/sph4u/hookeslaw.htm>
9. Bahar, I., Rader, A.J.: Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **15**(5), 586–592 (2005)
10. Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., Bahar, I.: Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **80**(1), 505–515 (2001)
11. Relation, S.: Generalized Hooke’s Law (2009)
12. Yang, G., Kabel, J., Rietbergen, B.V.A.N., Odgaard, A., Huiskes, R.I.K., Cowin, S.C.: The Anisotropic Hooke’s law for cancellous bone and wood. *J. Elast.* **2138**, 125–146 (1999)
13. Aweya, J., Ouellette, M., Montuno, D.Y., Doray, B., Felske, K.: An adaptive load balancing scheme for web servers. *Int. J. Netw. Manag.* **12**(1), 3–39 (2002)
14. Gaussian network model—Wikipedia, the free encyclopedia: <http://en.wikipedia.org/wiki/Gaussiannetworkmodel>

Part III
Computer-Based Systems

Chapter 17

Crossed Linear Arrays Using Doppler Radar Beamforming for Detecting Single Moving Targets

Jiajia Shi and Robin Braun

Abstract In machine vision and robotic system, moving object tracking is an important issue. Much research effort has been put to subtract moving objects from background and find out the location, speed and features of the objects. Variety of image processing methods and variety of sensors have been applied. Especially, radar and acoustic radar system have a good capability in detecting the position and speed of objects. Usually, long-range, high accurate radars are not equipped into robotic system due to the large size and high production cost. In this chapter, we first review radars which are used to detected human, vehicle and other objects in our daily lives. Then we proposed a low cost crossed linear arrays using acoustic Doppler radar beamforming method. Designs and methods are both described in theory and practical hardware. The results show that the proposed method has a good ability to detect a single moving target.

17.1 Introduction

Moving object tracking is a key to many intelligent robotic applications. This is generally achieved through the application of Machine Vision techniques to optical images. The level of computing power required can be very high, particularly in the case of images with large pixel numbers, and rapidly moving objects.

In applications where we are trying to detect single rapidly moving targets, such as in ball games like tennis, optical imaging can be very expensive and computationally demanding.

J. Shi (✉) · R. Braun
Centre for Real-Time Information Networks, University of Technology,
Sydney, Australia
e-mail: Jiajia.Shi-1@student.uts.edu.au

R. Braun
e-mail: Robin.Braun@uts.edu.au

In this paper, we describe the use of two “crossed” linear arrays with Doppler Radar and Beamforming to detect a single moving target. Because of its simplicity, it can be constructed cheaply, and can operate in both the ultrasonic bands as well as the RF/Microwave bands [11].

17.2 Previous Works

Much research has been done to build radars to track moving objects such as human bodies, vehicles and other objects in our daily life. Hewish [6] and Nebabin and Sergeev [9] distinguish human and different kinds of vehicles in battlefields by using Doppler effect inducing by radical movements. Geisheimer et al. [3] use the fully coherent, continuous-wave (CW) radar operating near 10.5 GHz to record the radar signature corresponding to the walking human gait. van Dorp and Groen [12] use a kinematic walking model developed by Thalmann [1] to calculate. Gurbuz et al. [5] used single-channel, airborne, synthetic aperture radar (SAR) for human target detection and identification. Human targets are differentiated from other detected slow-moving targets by analysing the spectrogram of each potential target. Otero [10] proposed a continuous wave radar for human gait recognition. Comparing with Geisheimer [3] who use a continuous chirplet transform and find a new FFT method. Liu et al. [8] applied pulsed Doppler range control radar into fall detection. He employed mel-frequency cepstral coefficients (MFCC) to represent the Doppler signatures of various human activities such as walking, bending down, falling, etc.

Besides radio wave radar, there are many acoustic Doppler radars. Kalgaonkar [7] proposed an acoustic Doppler Sensor (ADS) based technique for gait recognition. The device is built at very low cost using off-the-shelf 40 kHz acoustic devices and low-frequency audio-range sampling. Zhang [13, 14] also develop an acoustic continuous micro Doppler-radar to detect the gait signatures of both human and four-leg animals in indoor and outdoor environments. Georgiou [4] proposed a design and the collection of a multi-model data corpus for cognitive acoustic scene analysis. Most of the radars are either expensive or the detection rate is not high enough. Some radar only subtract features without knowing the position of objects.

17.3 Theoretical Basis

17.3.1 Doppler Radar Processing

The Doppler effect is that the frequency increases when the object approaching the receiver, while the frequency reduce when the object leaving the receiver. The amount

of frequency shift is proportional to the object's velocity towards the radar receiver. The Doppler shift frequency (F_D) is given by:

$$F_D = V \times \frac{f}{C} \cos \phi \quad (17.1)$$

where f = transmitter frequency in hertz, C = velocity (m/s), V = velocity of the target (m/s), ϕ = angle between object direction and radar beam.

17.3.1.1 Subtract Doppler Signal

There are many ways to subtract Doppler frequency from carrier frequency. The method widely used in analogue radar is to multiple the receiver signal with the transmitted signals. $X_T(t)$ is the transmitted signal, $X_R(t)$ is the received signal, and F_D is the Doppler frequency.

$$X_T(t) = \cos(2\pi \cdot f \cdot t) \quad (17.2)$$

$$X_R(t) = \cos(2\pi \cdot (f + F_D) \cdot t) \quad (17.3)$$

So when we multiple the signals we will get $X_m(t)$

$$\begin{aligned} X_m(t) &= \cos(2\pi \cdot (f + F_D) \cdot t) \cdot \cos(2\pi \cdot f \cdot t) \\ &= \frac{1}{2}(\cos(2\pi \cdot (2 \cdot f + F_D) \cdot t) + \cos(2\pi \cdot F_D \cdot t)) \end{aligned} \quad (17.4)$$

$X_m(t)$ is a signal mainly combined of 2 signal. If the carrier frequency f is much higher than F_D , $2 \cdot f + F_D$ is much higher than F_D , we can easily subtract F_D by adding a low pass filter.

When we use digital signal processing method, we usually adopt Analogue to Digital Convertor (ADC), then process the signal digitally. Similar to analogue processing method, we can also multiple signal with carrier frequency sample it. Then the sub-sampling is applied instead of low pass filter.

$$y[n] = \sum_{k=0}^{K-1} x[nM - k] \cdot h[k] \quad (17.5)$$

where $h[k]$ is impulse response.

According to different process order, we can use three methods as in Fig. 17.1.

In the first sub-sampling method, according to Nyquist sampling theorem, we need sample at a frequency higher than double of carrier frequency. In the second sub-sampling method, we only need sample at a frequency higher than double of highest Doppler frequency.

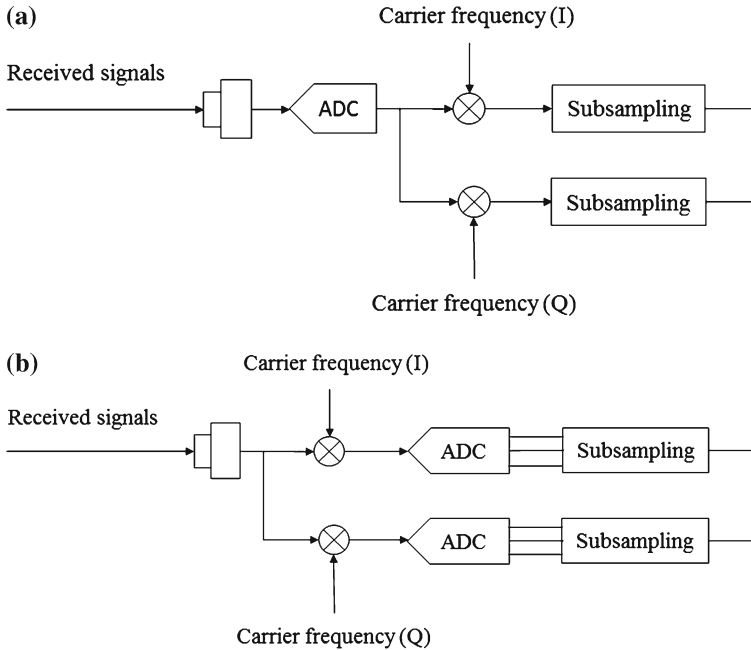


Fig. 17.1 Sub-sampling of Doppler signals

If we transfer the signal into frequency domain using Fourier transform, we can also subtract the Doppler signal, which is the frequency difference between received signal and transmitted signal. For real time object tracking, we need short-time Fourier transform (STFT), which is a time-frequency processing method. Here we analyze continuous-time STFT. A nonzero window function $\omega(t)$ is required for only a short period of time.

$$STFT\{x(t)\}(\tau, \omega) \equiv X(\tau, \omega) = \int_{-\infty}^{\infty} x(t)\omega(t - \tau)e^{-j\omega t} dt \quad (17.6)$$

If the window function is too wide, we can get very limited measurement of Doppler signal, otherwise, the resolution of each measurement is low.

17.3.2 Beamforming

Beamforming is also called spatial filtering. A SAR radar need mechanically steer beams to acquire information from each direction. Beamforming radar can electronically steer beams by adding delay to each sensor according to the direction.

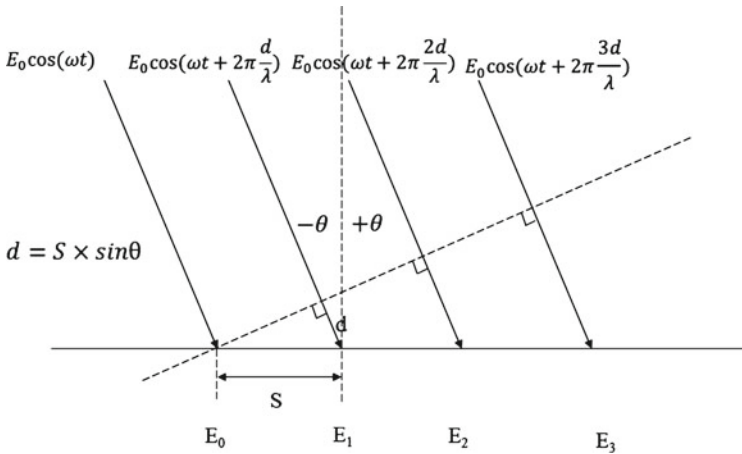


Fig. 17.2 Beamforming calculation

The basic idea can be seen from Fig. 17.2

$$\begin{aligned}
 s_q(t) &= E_q \cos(\omega t + n \times 2\pi d / \lambda) = E_q \cos(\omega t + q \times 2\pi n \cdot S \cdot \sin / \lambda) \\
 &= E_q \cos(2\pi f \cdot (t + q \cdot S \cdot \sin \theta / V))
 \end{aligned}
 \tag{17.7}$$

where, $S_q(t)$ = Received signals, E_q = Maximum amplitude of received signal of q th sensor after attenuation, w = Angular frequency of carrier waves, S = Space between sensors, λ = Wavelength of carrier waves.

17.3.2.1 Time Domain Beamforming

To process signal in computers, we need first sample the signals.

$$s_q[n] = E_q \cos(\omega(n + q \cdot S \cdot \sin \theta f_s / (V \cdot f))) \tag{17.8}$$

In a traditional beamforming which $S = \lambda/2$

$$s_q[n] = E_q \cos(\omega(n + q \cdot \pi \cdot \sin \theta)) \tag{17.9}$$

where f_s is the sample frequency. According to Nyquist sample theorem, $f_s \geq 2 \cdot f$

The basic idea of beamforming is adding delay to each sensor and sum the received signals.

If we add time delay $-q \cdot \pi \cdot \sin \theta$ to sensor q , the sum of all sensors is

$$y(n, \theta) = \sum_{q=0}^{Q-1} E_q \cos(\omega(n + q \cdot \pi \cdot \sin \theta - q \cdot \pi \cdot \sin \theta)) = \sum_{q=0}^{Q-1} E_q \cos(\omega n)
 \tag{17.10}$$

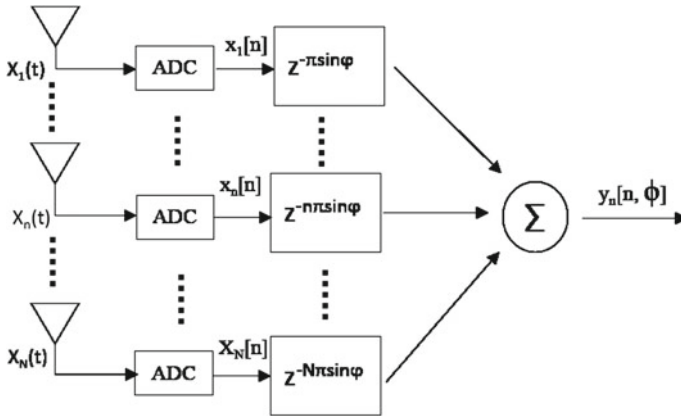


Fig. 17.3 Time domain beamforming

If $\varphi = \theta$,

$$Y(n, \varphi) = \sum_{q=0}^{Q-1} E_q \cos(\omega n) \tag{17.11}$$

And the absolute value of $Y(n, \varphi)$ will achieve the maximum.

The flowchart of time domain beamforming processing is shown in Fig. 17.3.

As n, q are integers, $\pi \cdot \sin \theta \cdot f_s / f$ is integer, which means beams can be quantized.

We assume $a = \pi \cdot \sin \theta \cdot f_s / f = 0, \pm 1, \pm 2, \dots$, and the number of value a represent beam numbers.

$$\theta = \arcsin\left(\frac{a}{\pi} \cdot \frac{f}{f_s}\right). \tag{17.12}$$

If $f_s = 2 \cdot f$, only 13 beams can be detected. We can increase the sample frequency to increase beam numbers.

If we increase sample frequency, the device cost will increase sharply and data loads expand quickly. There are several improved method to increase beam number without changing the hardware, such as interpolation beamforming.

17.3.2.2 Frequency Domain Beamforming

In frequency domain, according to Nyquist sample theorem, if $f_s / f > 2$, we can recover the original signals. So we do not need to increase sample frequency to increase beam numbers. In Discrete Fourier Transform, the time delay can be transformed into a phase shift. That is

$$x[n - \Delta n] \leftrightarrow X[k] \cdot e^{-j2\pi \Delta n k / N} \tag{17.13}$$

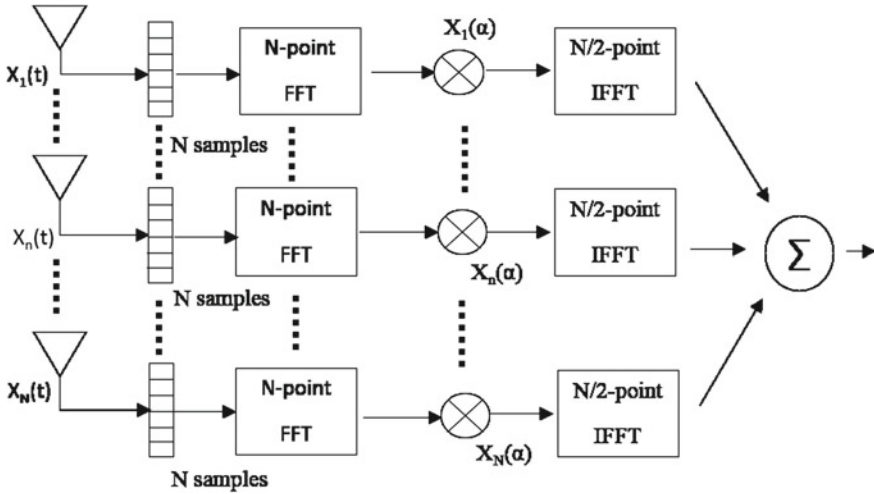


Fig. 17.4 Frequency domain beamforming

So $s_q[n]$ can be transformed into frequency domain. We assume sample length is N .

$$S_q[k] = \frac{1}{N} \sum_{n=0}^{N-1} s_q[n] \cdot e^{j2\pi nk/N} \tag{17.14}$$

So the delay and sum of the received signals $Y[k, \varphi]$ is

$$Y[k, \varphi] = \sum_{q=0}^{Q-1} S_q[k] \cdot e^{-j2\pi k \cdot q\pi \sin \varphi / N} \tag{17.15}$$

If $\varphi = \theta$, the absolute value of $y[\varphi]$ achieve the maximum. The flowchart of frequency domain beamforming is shown in Fig. 17.4.

17.3.2.3 Beamforming Processing on Doppler Signals

We assume Doppler frequency is F_D . So the sampled received signal of q th sensor is

$$s_q(t) = E_q \cos(2\pi(f + F_d)(t + q \cdot S \cdot \sin \theta / V)) \tag{17.16}$$

We can use sub-sampling to subtract the Doppler signal.

$$d_q(t) = s_q(t) \cdot \cos(2\pi ft) \tag{17.17}$$

If we sub sample $d_q(t)$, and sample frequency $2F_D \leq F_s < f$

$$ds_q[n] = \frac{E_q}{2} \cos(2\pi F_D n + 2\pi(f + F_D)qS \sin \theta / V) \tag{17.18}$$

Similar to beamforming, the frequency domain beamforming of Doppler signal is

$$DS_q[k] = \frac{1}{N} \sum_{n=0}^{N-1} ds_q[n] \cdot e^{j2\pi nk/N} \quad (17.19)$$

So the delay and sum of the received signals $Y[k, \varphi]$ is

$$Y[k, \varphi] = \sum_{q=0}^{Q-1} S_q[k] \cdot e^{-j2\pi k \cdot (f + F_D) \cdot q\pi \sin \varphi / (N \cdot f)} \quad (17.20)$$

If $\varphi = \theta$, the absolute value of $y[\varphi]$ achieve the maximum.

17.4 Hardware Implementation at Ultrasonic Frequencies

17.4.1 System Structure

Thanks to the development of electronic industry, high quality web-cams are very cheap. On other hand, FPGA board or Micro-controller board such as Raspberry Pi can be executed like a computer. Compare to electromagnetic wave, the frequency of ultrasonic wave is much lower, so the processing of acoustic beamforming and Doppler effects is much easier. For piratical application, the system should be neither too big nor too small, the 40kHz ultrasonic transducer is suitable for our project.

In some application, if the moving object is small, such as balls or birds, we need both horizontal and vertical positions of the objects. Two dimensional sensor arrays are usually used to solve this problem. The system structure is complicated and the computation load is high. If we adopt M sensors in horizontal direction and N sensors vertical direction, we totally need M by N sensors. The amount of ADCs and amplifies are also large. If we just use two crossed line linear sensors array, we can reduce total sensor number to $M + N - 1$ with the same space resolution in both direction. The beamforming processing in both horizontal or vertical directions can be processed separately to reduce computation load. The deficit of this design is that only one moving object can detected in one time. In most of applications, there usually exits one moving object. If we use optical image as a supplement of radar image, we can easily deal with ambiguities when there are multiple targets.

The layout of system structure is shown in Fig. 17.5.

17.4.2 Ultrasonic Transducer

There are two types of ultrasonic transducers available: Magnetostrictive and Piezoelectric.

Piezoelectric transducers use the piezoelectric effect and certain materials possess and convert electrical energy into mechanical energy. Magnetostrictive transducers work on the magnetostrictive effect and convert electrical energy into a magnetic

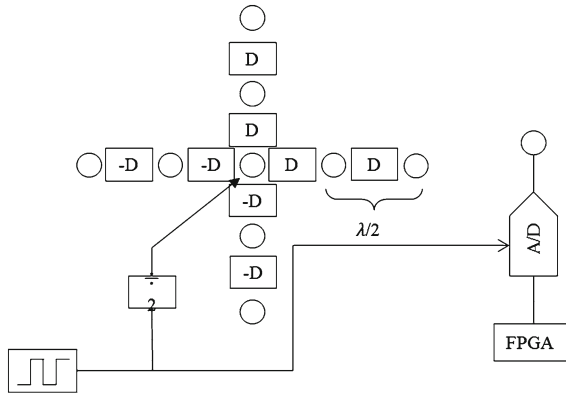


Fig. 17.5 Layout of sensors

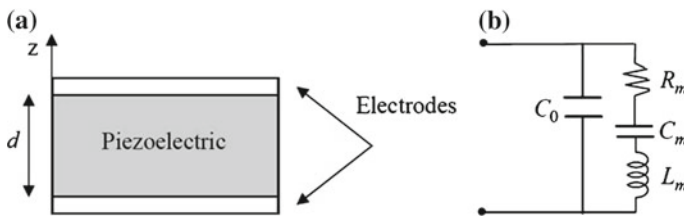


Fig. 17.6 Piezoelectric equivalent circuit [2]

field and then into mechanical energy. The electrical behavior of ultrasonic ceramic piezoelectric transducers may be represented by an equivalent circuit made up of inductors, capacitors and resistors and may be represented by the following simplified equivalent circuit which is shown in Fig. 17.6.

According to our project, we chose 400SR120 from Midas components Ltd as the receiver and 400PT120 as transmitter. They are transducer pairs with aluminum housing.

When we set the frequency at 40 kHz, the wavelength is very similar to a 35 GHz electromagnetic wave, which bandwidth is often used in radio wave Doppler-Radars. The wavelength is equivalent to minimum distinguish space. So the 40 kHz acoustic radar has the similar performance in minimum distinguish space with 35 GHz radio wave radar.

17.5 Controller and Programming

17.5.1 Altera DE1 Board and Cyclone II FPGA Chips

For cost and performance balance, we use Altera DE1 with Cyclone II FPGA chip: EP2C20F484C7N. The board is built in with a 512-Kbyte Static RAM memory chip,

organized as $256\text{ K} \times 16$ bits, and a 8-Mbyte Synchronous Dynamic RAM memory chip.

17.5.2 Signal Generator and ADC System Diagram

In the FPGA, executable blocks and a 32 bit Nios II micro-processor are built to control the whole system. The micro-controller connects to JTAG port and communicates with PC through USB-Blaster. USB-Blaster circuit is based on a FT245BM chip and a CPLD. The parallel output data from the FT245BM FIFO is serialized by the CPLD to drive the JTAG pins on the attached FPGA. The highest packets transmitting mode is ‘byte-mode’, where each byte from the host generates 8-bits of JTAG TDI activity. JTAG bit rate is 6MHz with USB Blaster. The best data rate will be 1 MB/s divided by 8, or about 1 Mbps. Therefore, in the first design, we cannot send the data in real-time. One solution is that we store the data in SRAM. Then we write the SRAM to PC. The transmitted signal are created by FPGA blocks. To ensure the strength of returned signals, we use a XOR gate to generate two inverse 5 V signals. Connect both the inverse signals to the two pin of ultrasonic transmitter. As the ultrasonic transducer is equivalent to capacitors, the actual wave is 40kHz sine wave at 10 V.

17.5.3 FPGA Programming

Several executable blocks are built in FPGA, which are shown in Fig. 17.7:

- DAC chip: MCP3001 which has 10-bit resolution, 200 ksp/s sampling rate at 5 V, SPI serial interface (modes 0,0 and 1,1). It requires Data, CS and CLK to convert data and transport data.

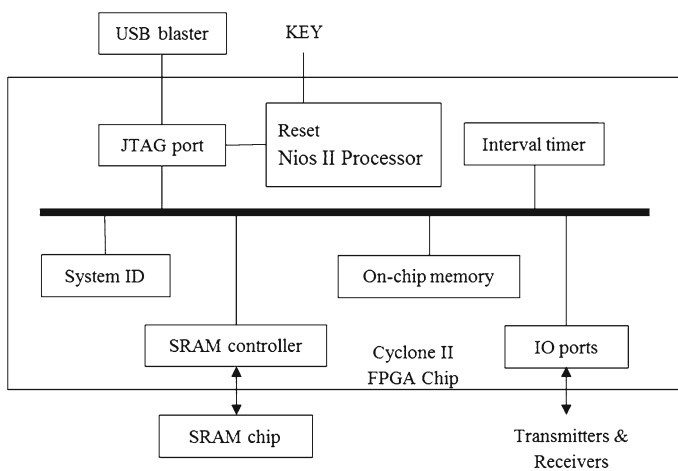


Fig. 17.7 Blocks inside FPGA

- Signal: Create the 40kHz pulsed wave as specified in the system description.
- Clock_ADC: The ADC clock sends to the ADC. It is also the data transfer clock between ADC and DE1 board. As it must more than 14 times of ADC sample rate (in our case, it is 160kHz). He we set it at 2.5MHz.
- CS_generator: The Chip Select clock sent to ADC to control the sample rate. The duty cycle rate should be approximately 92% that allow the ADC to send data. Also, this is the enable to data convert and memory storage. It is at 160kHz at our case.
- Delay: according to data sheet, use delay to make the waveform accurately what they want.
- ADC_s2p: convert the serial data from ADC to Parallel data according to data format.
- Timer: counter the time from each pulse and record the time when an object is detected.
- Memory: Store the sample data to SRAM and use SOPC to access the data and send them to host-computer through JTAG and USB_blaster.
- An SRAM Controller provides a 32-bit interface to the static RAM (SRAM) chip on the DE1 board. This SRAM chip is organized as $256\text{K} \times 16$ bits, but is accessible by the Nios II processor using word (32-bit), halfword (16-bit), or byte operations. The SRAM memory is mapped to the address space 0x08000000 to 0x0807FFFF.
- Nios II micro-processor: Altera Nios II processor a 32-bit CPU. In the first stage of our project, it contains 1 input: 10bit data, which receive the data from ram; 2 output: 1 bit enable, which enable the whole system and control when to write to SRAM and when to read from SRAM; 19 bit address, which help the host-computer to read data from SRAM.
- PLL block: according to the data sheet, the SRAM clock should be double of data rate, so a PLL block is required.

17.6 Simulation and Experiments

In a crossed linear array, one line of sensors are aligned in horizontal direction and one line of sensors are aligned in vertical direction. If we process received signals from all sensors, the computation load is high. The fast method is processing two linear array separately and multiple the possibilities to create a possibility image. The highest value in the specified location means the highest possibility that there is an object.

We capture the signals through the designed system and process the captured data in Matlab.

The processing of this method is very fast if we use desktop to calculate. In Fig. 17.8, we can find the one object is detected and the secondary lobe is quite small comparing to the peak.

The shortage is that only one object can be detected. If there are 2 targets in the detection distance, 4 possible objects will be detected. If there are 3 targets in the detection distance, 9 possible objects will be detected which is shown in Fig. 17.9.

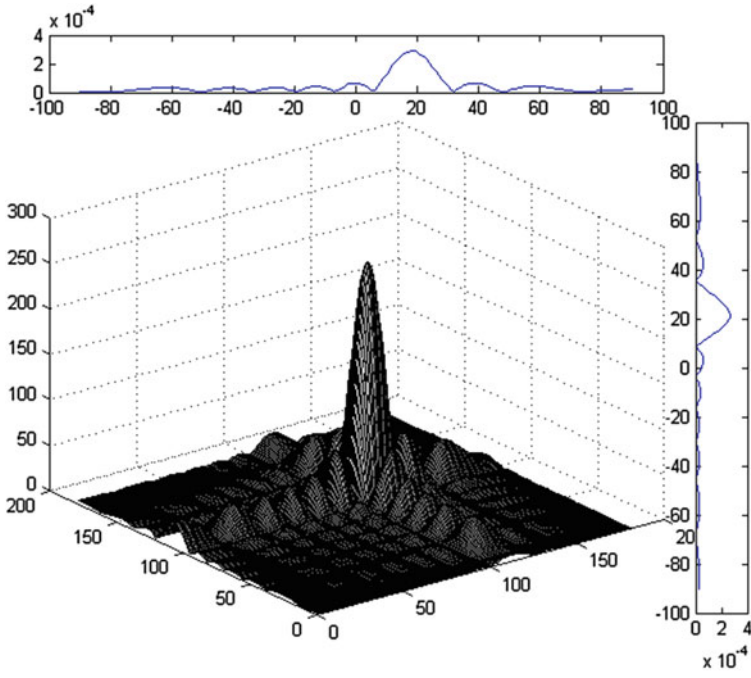


Fig. 17.8 2 dimensional beamforming

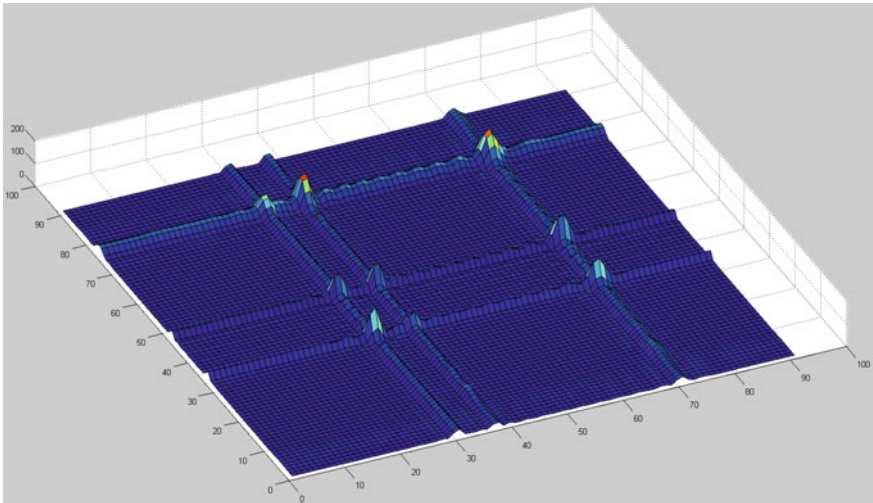


Fig. 17.9 Ambiguity of cross of sensors

17.7 Conclusion

In this paper, we first investigate existing portable low-cost radar system as well as the principle of the Doppler radar and beamforming processing method. Then we proposed a new designed cross-lined radar combining Doppler and beamforming processing method. The results show the processing speed is high and the performance is good in detecting one objects. Ambiguity exists when detecting multiple objects.

References

1. Boulic, R., Thalmann, N.M., Thalmann, D.: A global human walking model with real-time kinematic personification. *Vis. Comput.* **6**(6), 344–358 (1990)
2. Buccella, C., De Santis, V., Feliziani, M., Tognolatti, P.: Finite element modelling of a thin-film bulk acoustic resonator (FBAR). *COMPEL: Int. J. Comput. Math. Electr. Electron. Eng.* **6**, 1296–1306 (2008)
3. Geisheimer, J.L., Marshall, W.S., Grenaker, E.: A continuous-wave (CW) radar for gait analysis. In: *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 834–838 (2001)
4. Georgiou, J., Pouliquen, P., Cassidy, A.: A multimodal-corpus data collection system for cognitive acoustic scene analysis. In: *2011 45th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6 (2011)
5. Gurbuz, S., Melvin, W., Williams, D.: Detection and identification of human targets in radar data. In: Ivan, K. (ed.) *Proceedings of the SPIE Signal Processing, Sensor Fusion, and Target Recognition XVI*, vol. 6567, 65670I (2007)
6. Hewish, M.: The last line of defense: reducing the vulnerability of fixed installations. *Jane's Int. Def. Rev.* **32**(10), 28–30 (1999)
7. Kalgaonkar, K., Smaragdis, P., Raj, B.: Sensor and data systems, audio-assisted cameras and acoustic Doppler sensors. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2–3 (2007)
8. Liu, L., Popescu, M., Skubic, M., Rantz, M.: Automatic fall detection based on Doppler radar motion signature. In: *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp. 222–225 (2011)
9. Nebabin, V.G., Sergeev, V.V.: *Methods and techniques of radar recognition*. Moscow Izdatel Radio Sviaz 1 (1984)
10. Otero, M.: Application of a continuous wave radar for human gait recognition. In: Ivan, K. (ed.) *Proceedings of the SPIE Signal Processing, Sensor Fusion, and Target Recognition XIV*, pp. 538–548 (2005). http://72.52.208.92/~gbpprorg/mil/cavity/05_0330.pdf
11. Shi, J., Braun, R.: Crossed linear arrays using Doppler radar beam-forming for detecting single moving targets. In: *Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014*, p. 118 (2014)
12. van Dorp, P., Groen, F.C.A.: Human walking estimation with radar. In: *IEE Proceedings of Radar, Sonar and Navigation*, vol. 150, pp. 356–365 (2003). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1249155
13. Zhang, Z., Pouliquen, P.O., Waxman, A., Andreou, A.G.: Acoustic micro-Doppler radar for human gait imaging. *J. Acoust. Soc. Am.* **121**(3), 110–113 (2007)
14. Zhang, Z., Pouliquen, P.O., Waxmant, A., Andreou, A.G.: Acoustic micro-Doppler gait signatures of humans and animals. In: *41st Annual Conference on Information Sciences and Systems, CISS'07*, 14–16 March 2007, pp. 627–630 (2007)

Chapter 18

Image Construction Using Low Cost Airborne Beamforming

Jiajia Shi and Robin Braun

Abstract Images usually refer to video images which capture the density of light in each pixels. Instead of density of light, there are many other measurements we can use to sense the objects. For example, radar can sense the intensity of the reflected microwave or acoustic signal at each location. In this chapter, we review and describe new methods for using airborne beamforming radar to create images. By using radar technology, the images so created can have new information, such as velocity and distance. In this paper, we first review the characteristic of optical imaging. Then we proposed an image construction method using low cost airborne beamforming and compare the relationship between optical image and the created radar image. The performance of this method is also investigated. This kind of radar image can be used as a supplement of optical imaging to increase processing speed and accuracy.

18.1 Introduction

Beamforming is widely used for the detection of targets in remote sensing applications [2, 5, 10, 17]. However it is not widely used for the creation of images. In particular, optical imaging has become so widespread that other kinds of imaging, for other purposes does often not gain much attention. By imaging, we understand the creation of two dimensional representations of a “scene” in the field of view. This representation is generally a collection of values related to the scene at a particular location. In optical imaging this is an ordered collection of values representing the light intensity being emitted, or reflected from a portion of the scene. We generally call this a pixel.

J. Shi (✉) · R. Braun
Centre for Real-Time Information Networks, University of Technology,
Sydney, Australia
e-mail: Jiajia.Shi-1@student.uts.edu.au

R. Braun
e-mail: Robin.Braun@uts.edu.au

The question arises as to whether measurements of other factors than light intensity would be useful. For example, if the imaging was to be part of a machine vision or robotic system, then the value in a pixel would be used to indicate the presence or absence of an object at that location. In such a case, a radar image may be more useful. The pixel in the radar image would indicate the intensity of the reflected microwave or acoustic signal at that location. There are many other factors about the object at the location of the pixel that may be very useful. In particular, we may want to know the radial velocity of that object, or its distance from the receiver.

Optical imaging is not well adapted to provide us with that information. However, electromagnetic and acoustic signals reflected off the object have the potential to provide this kind of information. For example, Doppler Radar can measure radial velocity [14, 16].

Radar was developed as a means of observing an object in circumstances where it could not be seen. Perhaps it was too distant, or hidden behind clouds, or perhaps it was night time. It was not developed for the purposes of creating an image. All of the radar technology that we know was optimized for detecting these remote objects. Even beamforming radar, which was developed so as to remove the need for a moving or rotating antenna was developed for this purpose.

In this paper, we review and describe new methods for using beamforming to create images. By using radar technology, the images so created can have new information, such as velocity and distance. We have done extensive work using ultrasonics, which we use as the basis of this paper [13].

18.2 Video Camera Image Construction

18.2.1 Image Capture

Replacing film cameras, digital video cameras now dominate the image capturing market due to the cost efficiency and easy connection to computer. Usually, digital video cameras have these basic parts: lenses, diaphragm, shutter, CCD or CMOS. Specifically, Lenses focus the light rays coming from the objects into a coherent image; diaphragm determines amount of light to be entered; shutter quick open and close to control the time of exposure; CCD or CMOS sensors scales the projected light to generate corresponding electric charges or voltages. To represent color images, red, green and blue light is captured and scaled separately.

18.2.2 Digital Process

To convert charges or voltages to light intensity values, Analogue to Digital Converter (ADCs) are used to get value in binary form. In a digital image, a pixel is the combination of RGB values in a particular location. Then the ordered collection of two dimensional values are input to memory of microprocessors and stored as image file.

18.2.3 Resolution

The most important measurement to determine the quality of digital camera is image resolution. Resolution can be described in many ways and the higher resolution means the more image details. Several types of resolutions are used to measure image. The pixel resolution is the number of pixels in horizontal and vertical directions. Spatial resolution is capability to measure minimum distance between lines. Comparing with pixel resolution, spatial resolution depends on properties of the system creating the image, such as diffraction, aberration, imperfect focus and atmospheric distortion. Also, radiometric resolution is capability to distinguish differences of intensity. For a typical of computer image file, intensity is 256 levels, which are presented by 8 bits data.

18.2.4 Frame Rate

Human eyes can capture up to 12–13 images per second [11]. When the displaying frame rate is higher than the rate people can sense, people see moving pictures. Digital cameras usually capture the pictures at 24, 25 and 30 fps.

However, in the application of image processing or robotic system, we need much higher image rates. For example, in a tennis course, the speed of ball can achieve 263 km/h, that is 70.056 m/s. If we capture the picture at 30 fps. The distance between ball's positions in two continuous images should be about 2.43 m. In this case, we can not tell whether the ball is inside or outside borders.

18.2.5 Computation Load

In a typical RGB single digital camera system, the resolution is 640×480 and the frame rate is 25 fps. Each pixels have 3 bytes data. The total data rate is 23 MB/s. For full HD video with resolution of $1,920 \times 1,080$, the data rate is 155.52 MB/s.

Optical imaging needs 3 main steps: objects subtraction, feature detection and classification. Some pre-processing methods such as noise reduction are also need in case of low quality images. Each steps require loads of computation. Real-time image processing can only be applied in some specified environment, such as face recognition or finger recognition. In most circumstance, we post processing images.

18.3 Existed Radar Imaging

Recently, many kinds of sensor imaging have been developed, especially radio or ultrasonic sensors [1].

Radar image is a great alternative of optical image in robotic system. Radar can direct sense the exist of objects in pixel location as well as the distance and velocity. Synthetic Aperture Radars (SARs) [3, 6, 7, 15, 18] are widely used to create long distance pictures of earth or ocean surfaces. SARs are mounted on a moving platform such as an aircraft while ISARs are mounted on fixed platforms to scan rotated objects. SARs use relative motion between an antenna and its target region to provide distinctive long-term coherent-signal variation. The data are processed by Fast Fourier Transform. Huge computation load is required and real-time signal processing is a big challenge. Usually, data are preprocessed after capturing.

Medical ultrasonography [4] is an ultrasonic wave based medical device to sense muscles and organ structures. Specifically, B-mode ultrasonography can create a two dimensional image by transmitting and receiving ultrasonic wave through a linear array ultrasonic sensors. Doppler ultrasonography adopt Doppler effect in measuring and visualizing blood flow. 3D-mode ultrasonography can sense the three dimensional structure of human tissue and organ.

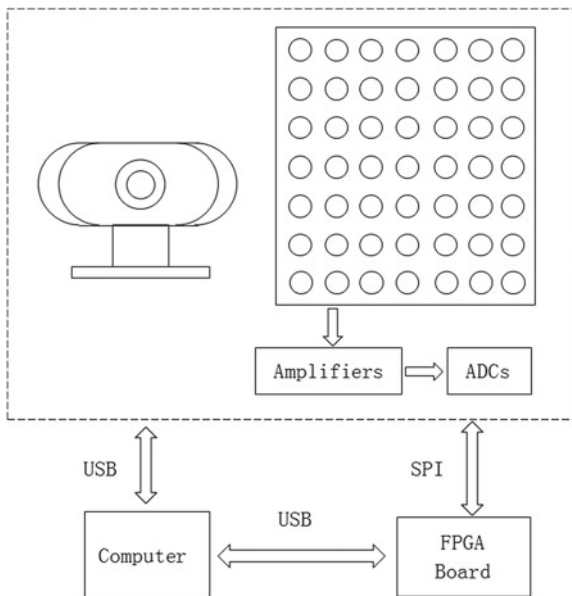
By now only a limited research has been done on creating airborne radar imaging in a vision distance. The reason is that other types of image are important when optical images can not be obtained. People manually determine what they see in an optical image. In a machine vision or robotic system, information captured by radar or ultrasonic device can provide complementary information to optical image to reduce the computation load and improve the detection accuracy. Also, object velocity can also be used to predict the trajectory of the object. Some research has been done to fuse radar and image. Roy et al. [12] use a system contain a continuous wave Doppler radar and a calibrated video camera to monitor the traffic. In their work, he use image to help Doppler radar to detect multiple target as an continuous Doppler radar cannot detect the location of objects. Kalman filter is used to fuse the image and radar. Ji and Prokhorov [8] proposed a classification system that incorporates information from a video camera and a long-range radar system. They use a classifier called Multilayer In-place Learning Network (MILN) to combine two types of information and short the processing time. Few research has been done on the fusion of optical image and radar image in pixel level.

18.4 Proposed Algorithm

18.4.1 System Structure

To compare the optical image and the proposed radar image, a web-cam and a beamforming pulsed Doppler radar are set side by side. The camera provides RGB values and beamforming pulsed Doppler radar provides location of objects, distance and speed of the pixel. A PC and a FPGA board are used to control the system and process the signals. The system structure is shown in Fig. 18.1.

Fig. 18.1 System structure

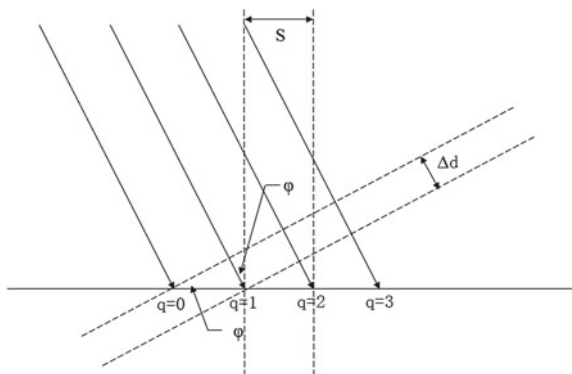


18.4.2 Beamforming

A beamforming device can electronically steer the beam without mechanical part. Therefore the steer speed is fast and the equipment is durable. The basic idea of beamforming is so called “delay-sum”, that is to add assumed time delays to the received signals from each sensors and sum the delayed signals. The angle relate to highest absolute value of the sum means the highest possibility that there are objects.

To create radar image with beamforming, we start with a linear of sensor arrays which is shown in Fig. 18.2.

Fig. 18.2 A linear of sensor arrays



To simplify our calculation, we use the approximations that the array operates in the far field of the transmitters. In the far field, the space between sensors is much smaller than the distance between objects and transmitter. So the angles of arrival signals are approximately equal for each sensor in a linear array and the returned signals come approximately parallel. Also, as the time difference of received signal from different sensor is very small, we assume the maximum amplitudes of received signal from different sensors are the same.

In Fig. 18.2, we notice that signals of $q = 0$ and $q = 1$ form a right angled triangle with the wave front. We also notice that the angle of arrival φ , which is the beam direction is also the acute angle in the triangle. We therefore note that

$$\Delta d = S \sin(\varphi) \quad (18.1)$$

Given that the velocity of the signal is v , we note that the signal arriving at $q = 0$ is delayed over the signal at $q = 0$ by

$$\Delta t = \frac{\Delta d}{v} \quad (18.2)$$

In general, the signal is delayed by

$$\Delta t_q = \frac{q \Delta d}{v} = \frac{q S_h \sin(\varphi)}{v} \quad (18.3)$$

Given a frequency in radians per second as ω_f , the delay translates into a phase shift as follows

$$\Delta \phi = q \frac{S \omega_f \sin(\varphi)}{v} = q K \sin(\varphi) \quad (18.4)$$

where $K = \frac{S \omega_f}{v}$,

To avoid aliasing, we need to sample at twice the highest the frequency, so the spacing between sensor are usually half wavelength, that is $\lambda/2$. As $v = f \lambda$ and $\lambda = v/f = 2\pi v/\omega_f$, $S = \frac{\lambda}{2}$.

$$\Delta \phi = q \pi \sin(\varphi) \quad (18.5)$$

Therefore, each of the signals from the different sensors can be written as

$$s_q(t) = A \cos(\omega_f t + q \pi \sin(\varphi)) \quad (18.6)$$

Let us assume we transform these signals to the sample domain at twice the highest frequency, which is $2\omega_f$. We get a sequence of samples from each sensors as follows:

$$S_q[n] = A \cos(n\pi + q \pi \sin(\varphi)) \quad (18.7)$$

The Discrete Fourier Transform can be written as

$$S_q(\omega) = AF(\cos(n\pi))e^{-j\omega q\pi \sin(\varphi)} \quad (18.8)$$

Here we need to add delay to each sensors according to assumed angle of arrival α varying from $-\pi/2$ to $\pi/2$.

$$X_q(\omega, \alpha) = AF(\cos(n\pi))e^{-j\omega q\pi \sin(\varphi)} e^{j\omega q\pi \sin(\alpha)} \quad (18.9)$$

$$x_q(n, \alpha) = F^{-1}(X_q) \quad (18.10)$$

We define $y(\alpha) = \sum_{n=1}^N x(n, \alpha)$, when $y(\alpha)$ achieve the maximum value, α is the angle of arrival.

In two dimensional sensors array. We assume a two dimensional array with $M \times N$ sensors.

$$\begin{bmatrix} q_{0,0} & q_{0,1} & \dots & q_{0,n} & \dots & q_{0,N} \\ q_{1,0} & q_{1,1} & \dots & q_{1,n} & \dots & q_{1,N} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ q_{m,0} & q_{m,1} & \dots & q_{m,n} & \dots & q_{m,N} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ q_{M,0} & q_{M,1} & \dots & q_{M,n} & \dots & q_{M,N} \end{bmatrix} \quad (18.11)$$

We can get the similar equations,

$$\Delta\phi = n \frac{S_h \omega_f \sin(\varphi)}{v} + m \frac{S_v \omega_f \sin(\theta)}{v} = nK_1 \sin(\varphi) + mK_2 \sin(\theta) \quad (18.12)$$

If both in vertical and horizontal direction, spacing $S_h = S_v = \frac{\lambda}{2}$.

$$\Delta\phi = n\pi \sin(\varphi) + m\pi \sin(\theta) \quad (18.13)$$

Therefore, each of the signals from the different sensors can be written as

$$s_{m,n}(t) = A \cos(\omega_f t + n\pi \sin(\varphi) + m\pi \sin(\theta)) \quad (18.14)$$

The Discrete Fourier Transform can be written as

$$S_{m,n} = AF(\cos(n\pi))e^{-j\omega m\pi \sin(\varphi)} e^{-j\omega n\pi \sin(\theta)} \quad (18.15)$$

Here we need to add delay to each sensors according to assumed angle of arrival α and β varying from $-\pi/2$ to $\pi/2$.

$$X_{m,n}(\omega, \alpha, \beta) = AF(\cos(n\pi))e^{-j\omega n\pi \sin(\varphi)}e^{-j\omega m\pi \sin(\theta)}e^{j\omega n\pi \sin(\alpha)}e^{-j\omega m\pi \sin(\beta)} \quad (18.16)$$

$$x_{m,n}(n, \alpha, \beta) = F^{-1}(X_{m,n}) \quad (18.17)$$

We define $y(\alpha, \beta) = \sum_{n=1}^N x(n, \alpha, \beta)$, We can create an image based on $y(\alpha, \beta)$. α and β are axis. The value of $y(\alpha, \beta)$ is possibility that whether there existed an object.

18.4.3 System Parameters

In practical application, equipments should be neither too big nor too small. For cost consideration, we use 40 kHz as carrier frequency because 40 kHz MEMS ultrasonic transducer is cheap and easy to get. We suppose the speed of sound in air is 340 m/s, and the spacing between sensors are $340/40000/2 = 0.00425$. For a 2 dimensional linear array with 169 sensors (13×13). The sensor array size is similar to 0.06×0.06 , which is very portable.

We need to sample the analogue signal using A/D converter. If we sample at higher frequency, the accurate is also higher, but the cost of hardware increase sharply, and the computation load and memory consumption is much higher. According to Nyquist theorem, a band limited signal must be sample at least twice the carrier frequency to avoid aliasing. In case of noise, we sample at 160 kHz.

18.4.4 Resolution

The half power beam width is defined as the angle between the half-power (-3 dB) points of the main lobe. A one-dimensional beamforming linear sensor array can detect objects from 0 to 180° . If we make the beamwidth lower than 4° . We can have more than 45 horizontal pixels in radar image. In Fig. 18.3, we use Matlab to calculate the beamwidth and set -3 dB to edge of beamwidth. In Table 18.1, we can find the relationship between number of sensor and resolutions. When the beamwidth is lower than 4° , we need 13 sensors. When the resolution is lower than 2° , we need 25 sensors.

In vertical one dimensional beamforming, the relationship of sensor number and angle resolution is similar.

Generally, the more sensors are used, the more accurate the results is; the bandwidth is lower and main lobe to side lobe rate is higher. However, the computation and memory will also increase sharply.

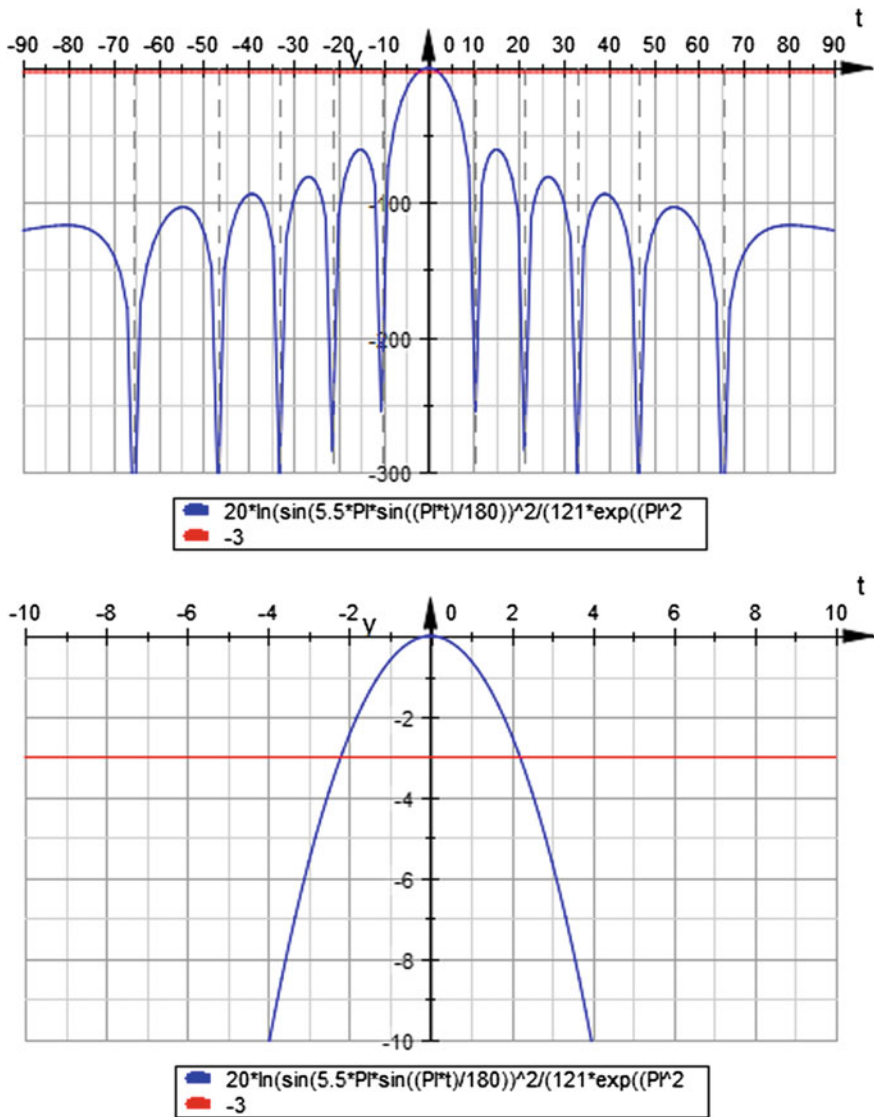


Fig. 18.3 Beamforming width

18.4.5 Frame Rate

In an optical image, the frame rate is 24, 25 or 30 frames per second and in our system, frame rate varies upon system modes. In continuous wave mode using Fast Fourier Transform, the frame rate can be calculated as follows.

Table 18.1 Relationship between sensor number and resolution

Number of sensor	Resolution (°)	Number of sensor	Resolution (°)
3	16.1	4	12.1
5	9.7	6	8.1
7	6.9	8	6.0
9	5.4	10	4.8
11	4.4	12	4.0
13	3.7	14	3.5
15	3.2	16	3.0
17	2.85	18	2.69
19	2.54	20	2.42
21	2.30	22	2.20
23	2.10	24	2.01
25	1.93	26	1.86

$$F = f/N \tag{18.18}$$

where f is the carrier frequency and N is the FFT series length. If 40kHz ultrasonic wave are transmitted and real time data is captured, and 512 points sample series have a proper resolution in distinguishing frequency, the frame rate is around 79 fps, which is relatively high for detecting moving objects. We can increase carrier frequency or reduce the Fourier series length to increase the frame rate.

18.4.6 Pixel Reflection

The VGA resolution is defined as 640×480 . For a APS-C camera, the relationship between angle of view and focal length is shown in Table 18.2.

A surveillance camera usually has wide angle length. we suppose the horizontal angle is 128 and vertical angle is 96. According to the resolution calculated above.

Table 18.2 Relationship between angle of view and focal length

Lens focal length (mm)	Angle of view in APS-C camera
10	107
18	74
50	30
100	15
250	6

If we use 24×24 sensors, we can have a radar image 64×48 . That is around 10 pixels in image is related to 1 pixel in radar.

18.4.7 Detection Distance

The attenuation of sound in medium depends on the following equation:

$$A = A_0 \cdot e^{-\alpha Z} \quad (18.19)$$

$$\alpha = \frac{0.1151}{v} U_t \quad (18.20)$$

In this expression A_0 is the unattenuated amplitude of the propagating wave at some location. The amplitude A is the reduced amplitude after the wave has traveled a distance z from that initial location. The quantity α is the attenuation coefficient of the wave traveling in the z -direction. v is the velocity of sound in meters per second and U_t is in decibels per second. In our application, the effective detection distance is about from 1 to 50 m.

18.4.8 Coordinate

Many coordinate systems have been used to present object position. Cartesian coordinate system is the most used system in which three perpendicular planes are chosen and the three coordinates of a point are the signed distances to each of the planes. Other types of system includes Polar coordinate system. In radar system with beamforming processing, angle of arrival is used to present the location of objects both in horizontal and vertical direction. Therefore a non-conventional coordinate system [9] which angle of arrivals as the main parameters is used in our method. The system is shown in Fig. 18.4.

In our coordinate system, the position of point P on the spherical surface is characterized by angle α , β and the distance R . α is the angle between the projection of the vector R to the plane $x0z$ and axis $0z$, while β is the angle between the projection of the vector R to the plane $y0z$ and axis $0z$, and R is the distance between the point P and the origin of the coordinate. The transform from $P(x_p, y_p, z_p)$ to $P(\alpha, \beta, R)$ is given as follows.

$$\alpha = \arctan \frac{x_p}{z_p} \quad (18.21)$$

$$\beta = \arctan \frac{y_p}{z_p} \quad (18.22)$$

$$R = \sqrt{x_p^2 + y_p^2 + z_p^2} \tag{18.23}$$

Inversely, the transform from $P(\alpha, \beta, R)$ to $P(x_p, y_p, z_p)$ is given as follows.

$$x_p = \frac{R \tan \alpha}{\sqrt{1 + \tan^2 \alpha + \tan^2 \beta}} \tag{18.24}$$

$$y_p = \frac{R \tan \beta}{\sqrt{1 + \tan^2 \alpha + \tan^2 \beta}} \tag{18.25}$$

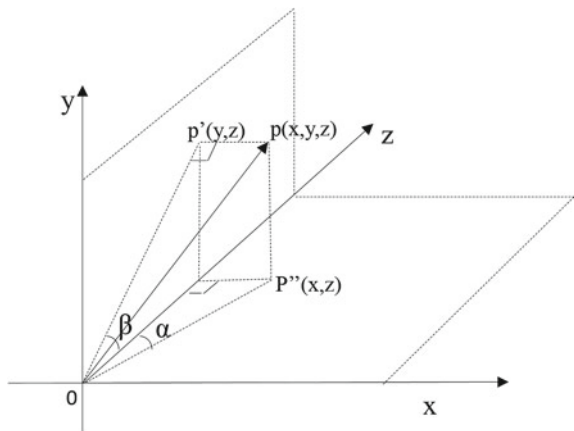
$$z_p = \frac{R}{\sqrt{1 + \tan^2 \alpha + \tan^2 \beta}} \tag{18.26}$$

18.5 Simulation Results

The main purpose of simulation is to investigate how sensor number and sensor location affects the created images. We suppose 4 or 5 objects are allocated in front of the 2 dimensional ultrasonic sensor arrays. The speed of sound is 40kHz and space between sensors are 0.00425 m.

In Fig. 18.5, 4 distanced objects are allocated. The 2 dimensional beamforming image can clear detect the objects and the secondary lobe is relatively small comparing the peak. If we increase the objects to 5 and we can see from Fig. 18.6, two objects on the top left are very closed to each. If 9×9 sensors are used, two objects cannot be separated, the first lobes are very close and the secondary lobes are overlapped. If Fig. 18.7, if we increase the sensor number to 15×15 , we can distinguish the two objects, though the secondary lobes still overlapped with each other. In Fig. 18.8, if

Fig. 18.4 Coordination



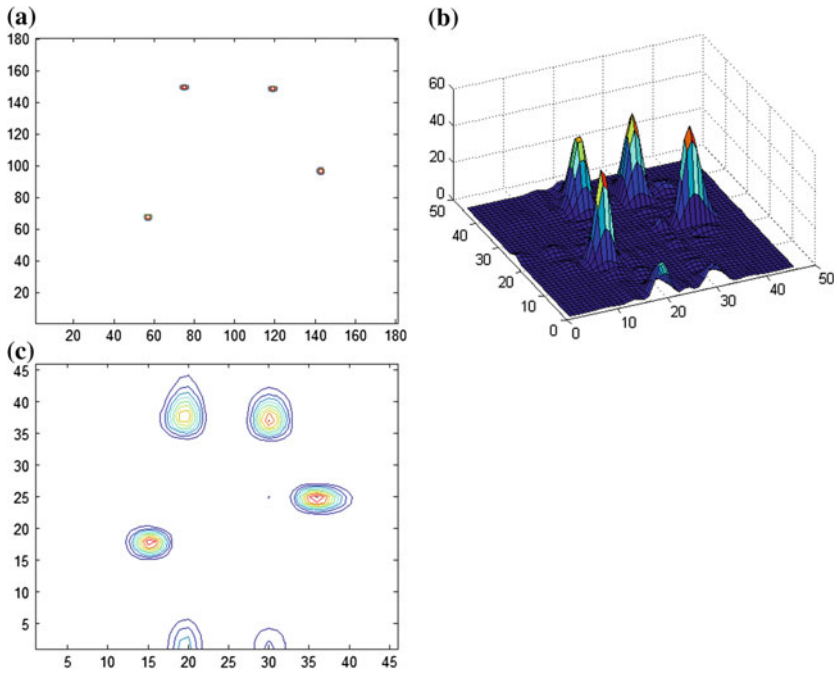


Fig. 18.5 Beamforming image with $M = 9, N = 9, 4$ objects

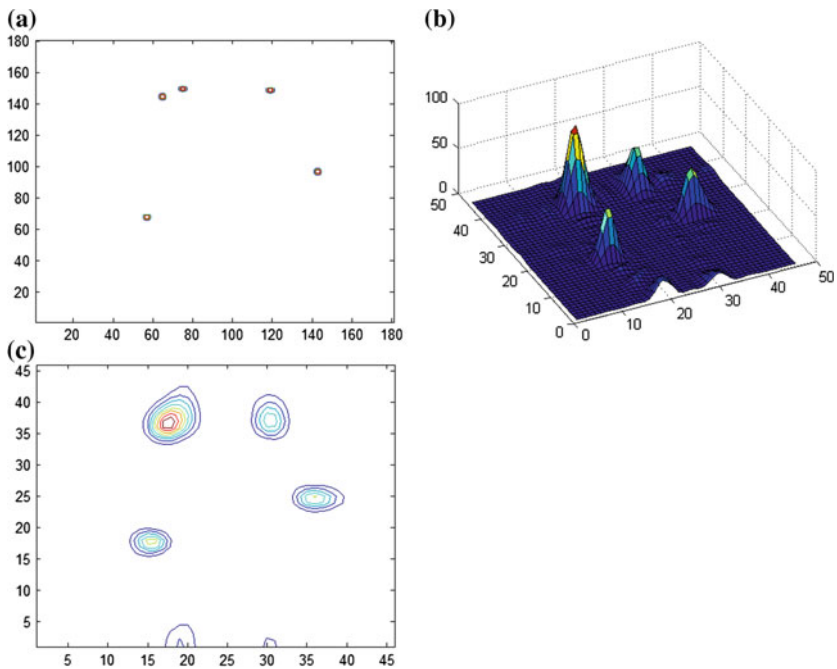


Fig. 18.6 Beamforming image with $M = 9, N = 9, 4$ objects

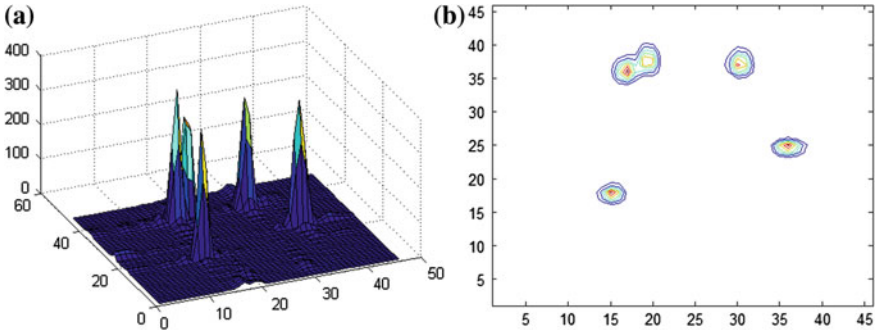


Fig. 18.7 Beamforming image with $M = 15, N = 15, 5$ objects

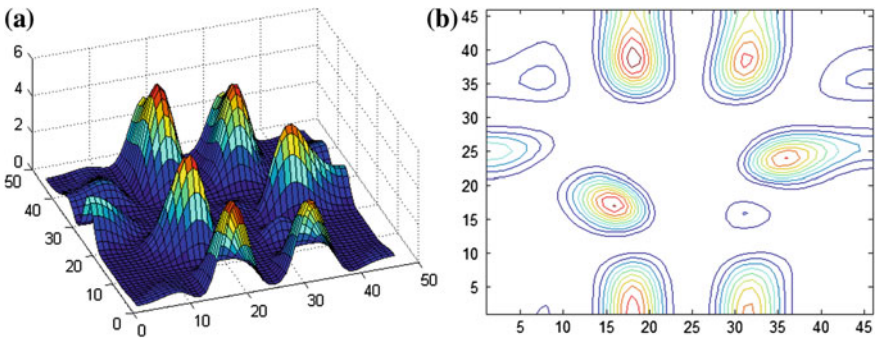


Fig. 18.8 Beamforming with $M = 15, N = 7, 4$ objects

we use 15×7 sensors, we can find that we have ambiguity in y direction, and the secondary lobes are relatively high comparing to first lobes, which the resolution in y direction is low.

18.6 Conclusions

In this paper, we compare the optical image with radar image using 2 dimensional low cost airborne beamforming. In machine vision and robotic system, radar image has better sense to real objects. The mathematical and practical basis of radar image is described. The results show the proposed image can provide more information to traditional optical which can accelerate the image processing speed and accuracy.

References

1. Ausherman, D.A., Kozma, A., Walker, J.L., Jones, H.M., Poggio, E.C.: Developments in radar imaging. *IEEE Trans. Aerosp. Electron. Syst.* **20**(4), 363–400 (1984)
2. Axelsson, S.R.J.: Noise radar for range/Doppler processing and digital beamforming using low-bit ADC. *IEEE Trans. Geosci. Remote Sens.* **41**(12), 2703–2720 (2003)
3. Çetin, M., Karl, W.C.: Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization. *IEEE Trans. Image Process.* **10**(4), 623–631 (2001)
4. Cole, C.R., Maslak, S.H., Petrofsky, J.G.: Method and apparatus for Doppler receive beamformer system (1996)
5. Fujisaka, T., Kirimoto, T., Ohashi, Y.: Digital beam forming radar system (1991)
6. Hasselmann, K., Hasselmann, S.: On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion. *J. Geophys. Res.: Oceans* (1978–2012) **96**(C6), 10713–10729 (1991)
7. Hasselmann, S., Brüning, C., Hasselmann, K., Heimbach, P.: An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra. *J. Geophys. Res.: Oceans* (1978–2012) **101**(C7), 16615–16629 (1996)
8. Ji, Z., Prokhorov, D.: Radar-vision fusion for object classification. In: 2008 11th International Conference on Information Fusion, pp. 1–7 (2008)
9. Kažys, R., Jakevičius, L., Mažeika, L.: Beamforming by means of 2D phased ultrasonic arrays. *Ultragarsas (Ultrasound)*, Kaunas, *Technologija* **1**(29), 12–15 (1998)
10. Krieger, G., Gebert, N., Moreira, A.: Multidimensional waveform encoding: a new digital beamforming technique for synthetic aperture radar remote sensing. *IEEE Trans. Geosci. Remote Sens.* **46**(1), 31–46 (2008)
11. Read, P., Meyer, M.P.: *Restoration of Motion Picture Film*. Butterworth-Heinemann, London (2000)
12. Roy, A., Gale, N., Hong, L.: Fusion of Doppler radar and video information for automated traffic surveillance. In: 12th International Conference on Information Fusion. FUSION'09, pp. 1989–1996 (2009)
13. Shi, J., Braun, R.: Image construction using beam forming. In: Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014, p. 121 (2004)
14. Skolnik, M.I.: *Radar Handbook*. McGraw-Hill, New York (1970)
15. Ulaby, F.T., Kouyate, F., Brisco, B., Williams, T.H.L.: Textural information in SAR images. *IEEE Trans. Geosci. Remote Sens.* **24**(2), 235–245 (1986)
16. Waldteufel, P., Corbin, H.: On the analysis of single-Doppler radar data. *J. Appl. Meteorol.* **18**(4), 532–542 (1979)
17. Xu, L., Li, J., Stoica, P.: Adaptive techniques for MIMO radar. In: Fourth IEEE Workshop on Sensor Array and Multichannel Processing, pp. 258–262. IEEE (2006)
18. Zebker, H.A., Goldstein, R.M.: Topographic mapping from interferometric synthetic aperture radar observations. *J. Geophys. Res.: Solid Earth* (1978–2012) **91**(B5), 4993–4999 (1986)

Chapter 19

Control of Hand Prosthesis Using Fusion of Biosignals and Information from Prosthesis Sensors

Andrzej Wolczowski and Marek Kurzynski

Abstract In this chapter we present an advanced method of recognition of patient's intention to move the multi-articulated prosthetic hand during grasping and manipulation of objects in a skillful manner. The proposed method is based on a 2-level multiclassifier system (MCS) with base classifiers dedicated to particular types of biosignals: electromyography (EMG) and mechanomyography (MMG) signals, and with combining mechanism using a dynamic ensemble selection scheme and probabilistic competence function. To improve the precision and reliability of prosthesis control, the feedback signal derived from the prosthesis sensors is used. We present two original concepts of using such a signal. In the 1st method, the feedback signal is treated as a source of information about a correct class of hand movement and competence functions of base classifiers are dynamically tuned according to this information. In the 2nd one, classification procedure is organized into multi-stage process based on a decision tree scheme and consequently, feedback signal indicating an interior node of a tree allows us to narrow down the set of classes. The performance of MCS with feedback signal were experimentally tested on real datasets. The obtained results show that developed methodology can be practically applied to design a control system for dexterous bioprosthetic hand.

19.1 Introduction

The general concept of bioprosthetic hand control is as follows: the signals generated by the active muscles (usually EMG signals) of the upper limb stump equipped with prosthesis are measured, and after analysis and classification converted into decisions controlling the prosthesis finger motion [1–3, 8, 13]. The latest publications in this field draw attention to the importance of sensory feedback signals from the

A. Wolczowski (✉) · M. Kurzynski
Wroclaw University of Technology, Wroclaw, Poland
e-mail: andrzej.wolczowski@pwr.edu.pl

M. Kurzynski
e-mail: marek.kurzynski@pwr.edu.pl

prosthesis to the man for the better precision of prosthesis manipulation and prehensile movements [5, 6].

The efficiency of the realized control process depends on the quality of signal representation (i.e., feature vectors) subjected to recognition, that is quality of biosignals acquisition and analysis process, but also on assumed number of recognized motion classes (hand gestures) and required level of recognition reliability, (wherein these values are inversely proportional to each other—the higher level of recognition reliability required the lower allowable number of recognized classes—and vice versa) [10, 18]. To improve the reliability and precision of control (i.e., reduce uncertainty of motion classes recognition while increasing the cardinality of the repertoire of motion commands), the process of classification should be helped by additional information derived from the process of prosthesis interaction with a grasped object, supplied through prosthesis sensors. Most frequently such information relates to posture (arrangement of the fingers) and the pressure forces of the fingers on the object [15].

19.2 The Biosignal-Based Prosthesis Control System

The idea of biosignals-based prosthesis control is shown in Fig. 19.1. Natural hand movements are accompanied by electrical and mechanical signals produced by the working muscles. These signals, after amputation of the hand can be captured and after analysis and classification used to control the movements of the prosthesis.

Functional diagram of natural hand

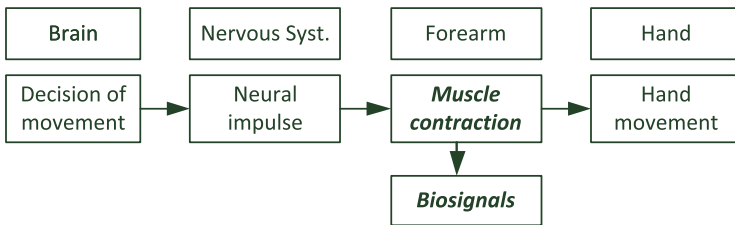


Diagram of prosthetic hand operation

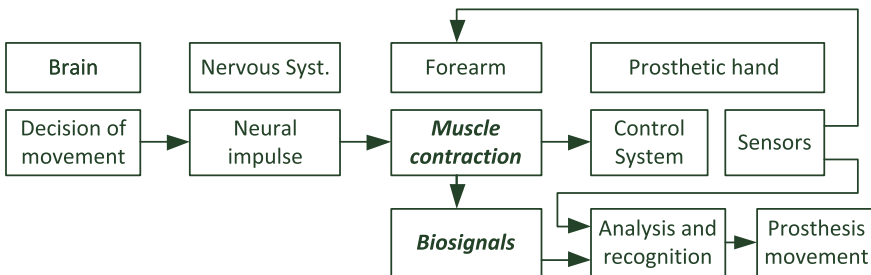


Fig. 19.1 Idea of biosignals-based prosthesis control

19.2.1 Myoelectric and Myomechanic Signals Phenomenon

Myoelectric phenomenon consists in variation of electrical potentials along cellular membranes of contracting muscle fibers, in relation to the surrounding tissue. This variation propagates through the body and can be registered on its skin as the EMG signal. Since in the working muscle the distinct fiber groups are successively activated and relaxed, the created myopotentials have an alternate character, and while propagating in the tissue, they are subject to spatial filtration. The EMG signal measured at a given point is the superposition of the potentials produced by all active fibers, and depends on spatial configuration of the fibers [7].

Muscle activity is accompanied as well by mechanical vibrations of the fiber groups, which (like myopotentials) propagate through the surrounding tissue to the body surface, and they can be registered by a microphone sensors. These vibrations are called MMG signals [9].

The distinct types of movements call for various activities of muscles, so the various spatial configuration of active fibers. Hence, the distinct finger movements correspond to the distinct classes of the EMG and MMG signals.

The biosignal-based control takes advantage of the fact that the vast majority of finger driving muscles are located in the forearm, which means that after a hand amputation they are left in the stump. The activity of these muscles still depends on the patients will, so biosignals accompanying this activity can be used to control prosthesis motion.

19.2.2 The Reliability Issue of Recognition of Prosthesis Motion Control Decisions

A key problem in the control of contemporary prosthetic hands, based on the biosignals recognition, is the limited number of reliably recognized classes of prosthesis movements. This results from the unavoidable decrease in the recognition reliability with increasing cardinality of recognized classes for the same amount of available information [14].

Possible solutions consist in:

1. extraction of more information from the acquired biosignals. This approach boils down to search for and choose the best methods for the analysis of this type of signals.
2. Obtaining more information from the combination of information from various kinds of biosignals. For example, muscle activity is accompanied by both electrical signals (EMG) and mechanical signals (MMG), which contain complementary information about the muscles contraction.
3. Using advanced classifier (2-level multiclassifier) systems dedicated to particular registered biosignals, working in the dynamic ensemble selection scheme with measures of competence of base classifiers.

4. Considering in the motion control process the information concerning the prosthesis interaction with a grasped object, supplied by prosthesis sensors. As a result, for the same motor command the prosthesis fingers can be set at different positions depending on the shape and size of object to be grasped.
5. Supporting the process of biosignals recognition by information from the prosthesis sensors. In this approach, information about the type of preliminary contact of the prosthesis with grasped object, influence the choice of the final decision controlling a grip (modification during the object capture).
6. Increasing the precision of movement commands issued by enables the patient “to feel” the movement of prosthesis fingers and their pressure on the grasped object.

The developed bioprosthesis control system includes most of the above mentioned ideas. It uses both EMG and MMG signals (according to proposition 2) as well as information from prosthesis sensors according to proposition 4. To recognize the intended hand movements (according to proposition 3) the multiclassifier system, working in a hierarchical structure, was applied.

19.2.3 Biosignal Acquisition and Analysis

Biosignal acquisition and analysis processes influence essentially on the reliability of recognition of prosthesis motion control decisions. The acquisition process should take into account the nature of the measured signals and their measurement conditions. A quality of the obtaining information depends on the ratio of the measured signal power to the interfering signal power, defined as Signal to Noise Ratio (SNR). For the non-invasive methods of measurements carried out on the surface of the patient's body, it is difficult to obtain a satisfactory SNR [14]. The noise amplitude usually exceeds many times the amplitude of the signal. For the EMG signals the amplitude of voltages induced on the patient body as a result of the influence of external electric fields, may exceed more than 1,000 times the value of useful signals. To overcome this difficulty a differential measurement system was applied. The system encompasses two signal electrodes placed above the examined muscle and an reference electrode placed as far as possible above electrically neutral tissue. Signals obtained from signal electrodes are subtracted from each other and amplified. The common components, including surrounding noise, are thus excluded and the useful signal is amplified. In the case of MMG signals the basic problem is to isolate the microphone sensor from the external sound sources along with the best acquisition of the sound propagating in the patient's tissue.

After the acquisition stage, the recorded signals have the form of strings of discrete samples. Their size is the product of measurement time and sampling frequency. For a typical motion action, that gives a record of size between 3 and 5 thousand of samples (time of the order of 3–5 s, and 1 kHz sampling). This primary representation of the signals hinders the effective classification and requires the reduction of

dimensionality. This reduction leads to a representation in the form of a signal feature vector. To determine the algorithm of features extraction, the database records were divided into 256 ms frames and then analyzed in time and frequency using Short Time Fourier Transform (STFT). Figures 19.2 and 19.3 show the exemplary results.

As we can see, the MMG histogram has two amplitude peaks: at the beginning and at the end of the movement, and relatively low amplitude in the middle while the EMG histogram shows a peak in the middle of the movement time span. The analyses of histograms for the tested movements allowed selecting the localization of the best signal features (the best points in time and frequency) securing the best differentiation of the movements.

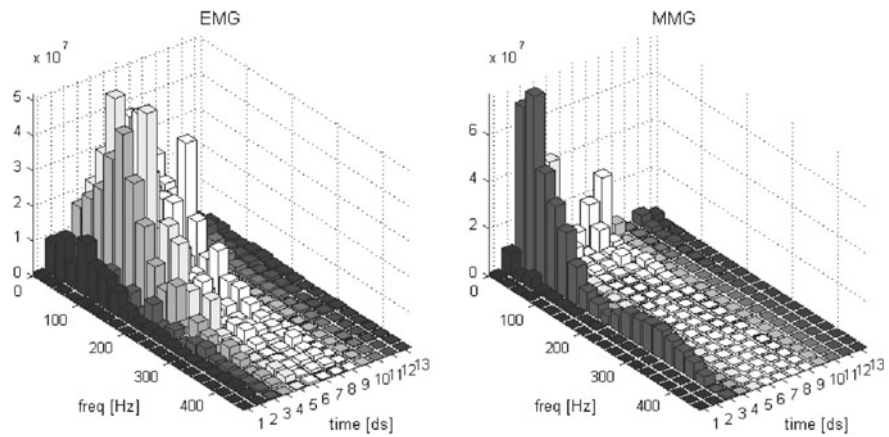


Fig. 19.2 Histograms obtained by Short Fourier Transform for EMG signal (left) and for MMG signal (right)

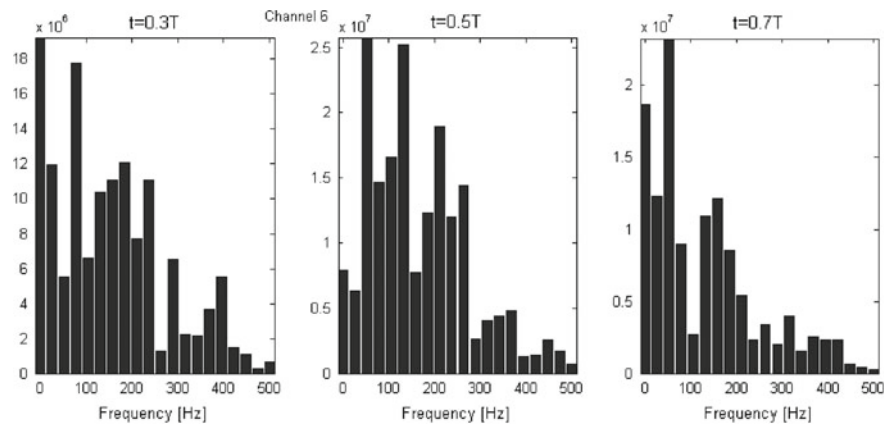


Fig. 19.3 The 3 (most representative) exemplary time slices

The resulting algorithm has the following form:

- Step 1.** Extract from the recorded EMG and MMG signals representing the specified movements the 256 sample segments. Each extracted segment has new time span ($t \in [0, T]$);
- Step 2.** Apply the STFT to each segment (Fig. 19.2);
- Step 3.** Choose as signal features the values from the STFT product corresponding to the k (most representative) time slices(Fig. 19.3);
- Step 4.** Repeat steps 2 and 3 for every channel;
- Step 5.** Use all the obtained (in steps 2 and 3) values as elements of the feature vector representing the analyzed signal segment.

This procedure allows creating input vectors with an adjustable size. The structure of this feature vector used as an input in the classifier is given by:

$$\left(A_{t_1}^{CH_i}, A_{t_2}^{CH_i}, \dots, A_{t_k}^{CH_i} \right)_{i=1,2,\dots,n}, \tag{19.1}$$

where k is the number of time slices and n denotes the number of signal channels.

19.2.4 The Classifier Structure

The block diagram of the 2-level multiclassifier system is shown in Fig. 19.4. It consists of two multiclassifier subsystems dedicated to individual biosignals [17].

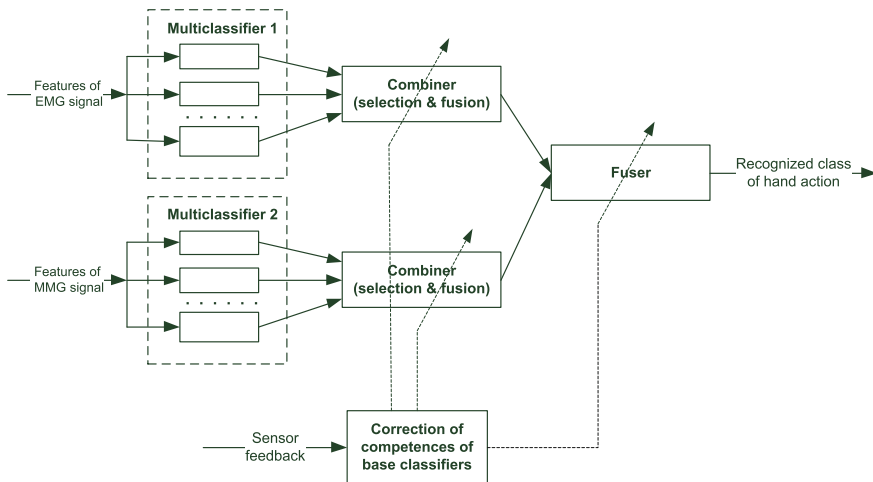


Fig. 19.4 Block-diagram of the proposed multiclassifier system

This means that each of them performs the classification on the basis of a different set of features. Both subsystems have been implemented according to the scheme of dynamic selection of classifier ensemble (DES). In this scheme, first base classifiers are evaluated using the so called competence function constructed in the learning procedure from validation set. Then for each classified object (class of hand grip) the competent classifiers are selected (with competence higher than a given threshold) and their decisions are combined by weighted majority voting mechanism at the level of classifying function. A similar mechanism of fusion of both sub-multiclassifiers has been applied on the second level. In addition, the classification procedure is adaptively tuned based on signals from bioprosthesis sensors, which are a source of information about the actual realized movement class. This is done by adjusting the competence functions of base classifiers involved in the recognition, according to the following rules: the value of classifier competence for recognized movement is increased (decreased) if the analyzed movement is classified correctly (wrongly).

In the next subsection the multiclassifier applied in the first level of the system is presented in detail.

19.2.5 The MCS with Dynamic Ensemble Selection Scheme and Competence Measure

In the multiclassifier system (MCS) we assume that a set of trained classifiers $\Psi = \{\psi_1, \psi_2, \dots, \psi_L\}$ called base classifiers is given. A classifier ψ_l is a function $\psi_l : \mathcal{X} \rightarrow \mathcal{M}$ from a feature space \mathcal{X} to a set of class labels $\mathcal{M} = \{1, 2, \dots, M\}$. Classification is made according to the maximum rule

$$\psi_l(x) = i \Leftrightarrow d_{li}(x) = \max_{j \in \mathcal{M}} d_{lj}(x), \quad (19.2)$$

where $[d_{l1}(x), d_{l2}(x), \dots, d_{lM}(x)]$ is a vector of class supports (classifying function) produced by ψ_l . The value of $d_{lj}(x)$, $j \in \mathcal{M}$ represents a support given by the classifier ψ_l for the fact that the object x belongs to the j th class. Without loss of generality we assume that $d_{lj}(x) \geq 0$ and $\sum_j d_{lj}(x) = 1$.

The ensemble Ψ is used for classification through a combination function which, for example, can select a single classifier or a subset of classifiers from the ensemble, it can be independent or dependent on the feature vector x (in the latter case the function is said to be dynamic), and it can be non-trainable or trainable [4]. The proposed multiclassifier system uses dynamic ensemble selection (DES) strategy with trainable selection/fusion algorithm. The basis for dynamic selection of classifiers from the pool is a competence measure $C(\psi_l|x)$ of each base classifier ($l = 1, 2, \dots, L$), which evaluates the competence of classifier ψ_l at a point $x \in \mathcal{X}$, i.e. its capability to correct activity (correct classification). For the training of competence it is assumed that a validation set

$$\mathcal{V} = \{(x_1, j_1), (x_2, j_2), \dots, (x_N, j_N)\}; \quad x_k \in \mathcal{X}, \quad j_k \in \mathcal{M} \quad (19.3)$$

containing pairs of feature vectors and their corresponding class labels is available.

The competence $C(\psi_l|x)$ of the classifier ψ_l for the object x is calculated using the potential function model with Gaussian potential function:

$$C(\psi_l|x) = \sum_{x_k \in V} C_{src}(\psi_l|x_k) \exp\left(-d(x, x_k)^2\right), \quad (19.4)$$

where $C_{src}(\psi_l|x_k)$ is the source competence calculated for an object x_k taken from V and $d(x, x_k)$ is the Euclidean distance between the objects x and x_k .

The source competence $C_{src}(\psi_l|x_k)$ of the classifier ψ_l at a point $x_k \in \mathcal{X}$ from the set (19.3) is defined as [16]:

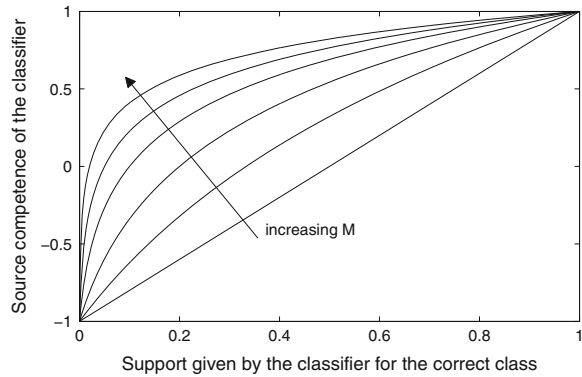
$$C_{src}(\psi_l|x_k) = 2 \cdot d_{jk}(x_k)^{\frac{\log(2)}{\log(M)}} - 1. \quad (19.5)$$

The values of the function $C_{src}(\psi_l|x_k)$ lie within the interval $[-1, 1]$, where the interval limits -1 and 1 describe absolutely incompetent and absolutely competent classifier, respectively. The function $C_{src}(\psi_l|x_k)$ was defined in such a way because it satisfies the following criteria:

- it is strictly increasing, i.e. when the support $d_{jk}(x_k)$ for the correct class increases the competence $C_{src}(\psi_l|x_k)$ also increases,
- it is equal to -1 (evaluates the classifier as absolutely incompetent) in the case of zero support for the correct class, i.e. $d_{jk}(x_k) = 0 \Rightarrow C_{src}(\psi_l|x_k) = -1$,
- it is negative (evaluates the classifier as incompetent) in the case where the support for the correct class is lower than the probability of random guessing, i.e. $d_{jk}(x_k) \in [0, \frac{1}{M}) \Rightarrow C_{src}(\psi_l|x_k) < 0$,
- it is equal to 0 (evaluates the classifier as neutral or random) in the case where the support for the correct class is equal to the probability of random guessing, i.e. $d_{jk}(x_k) = \frac{1}{M} \Rightarrow C_{src}(\psi_l|x_k) = 0$,
- it is positive (evaluates the classifier as competent) in the case where the support for the correct class is greater than the probability of random guessing, i.e. $d_{jk}(x_k) \in (\frac{1}{M}, 1] \Rightarrow C_{src}(\psi_l|x_k) > 0$,
- it is equal to 1 (evaluates the classifier as absolutely competent) in the case of maximum support for the correct class, i.e. $d_{jk}(x_k) = 1 \Rightarrow C_{src}(\psi_l|x_k) = 1$.

The source competence also depends on the number of classes in the classification problem. This dependence is shown in Fig. 19.5 where the value of the source competence is plotted against the support given by the classifier for the correct class.

Fig. 19.5 The source competence plotted against the support for the correct class for different number of classes ($M = 2, 3, 5, 10, 20, 50$)



19.2.6 Sequential Classification Procedures

To improve the precision and reliability of prosthesis control, the feedback signal derived from the prosthesis sensors was used. In our method, a classification procedure is organized into multistage process based on a decision tree scheme, and consequently a feedback signal indicating an interior node of a tree allows us to narrow down the set of classes. The basis for the use of feedback signals as an aid in the classification process (by narrowing the set of recognized grips) are:

- a decomposition of each single gripping movement onto the sequence of the so called elementary action;
- considering the classification as a multi-stage process described by the form of a decision tree diagram.

Human hand can be considered as a goal-oriented system $\langle H \rangle$, (controlled by the nervous system $\langle H, 0 \rangle$). The system operation consist in grasping the objects and/or manipulating them (e.g. capture, squeezing/pumping a rubber bulb, gripping sprayer, pressing of its trigger, gripping a cell-phone and pressing of its keys). In the operation process of $\langle H \rangle$, you can distinguish several separate grasping phases (the elementary actions, repeated regardless of the type of object to be gripped). These activities can be formally described as elements of a set:

$$\langle H \rangle \longleftrightarrow A = \{a_1, a_2, \dots, a_k\}. \quad (19.6)$$

The realization of each action a_i is a separate stage in the $\langle H \rangle$ system operation. Observing these states, by classifying successive segments of multimodal biosignals, is the basis of multistage control of artificial hand [11]. Considering some consecutive actions jointly (in the same phase of grasping process), in the simplest case, we can distinguished 7 types of macro-stages (macro-actions) in the process of grasping with a hand [12]:

a_0 —rest position (starting point for the grasp preparation; the fingers stay at rest half-closed arrangement, are motionless and passive);

a_1 —grasp preparation (the fingers are opening or closing, taking the posture depending on the shape of the object that is observed visually and the knowledge K about the method of grasping it);

a_2 —grasp closing (precedes the grasp—the fingers move seeking a contact with the grasped object, taking the posture depending on the object);

a_3 —grabbing (finger posture adapts to the shape and size of the object, finger contact with the object varies from random to appropriate for a given object; finally the fingers pressing on the object with the force resulting from the observation of his behavior and knowledge K);

a_4 —maintaining the grasp (with force adjustment—the fingers increase/decrease the squeeze depending on object deformation and slip);

a_5 —releasing the grasp (the fingers move with a velocity dependent on the knowledge of the object behavior, e.g. small velocity for an object with an unstable balance);

a_6 —transition to the rest position (the fingers move toward the rest position).

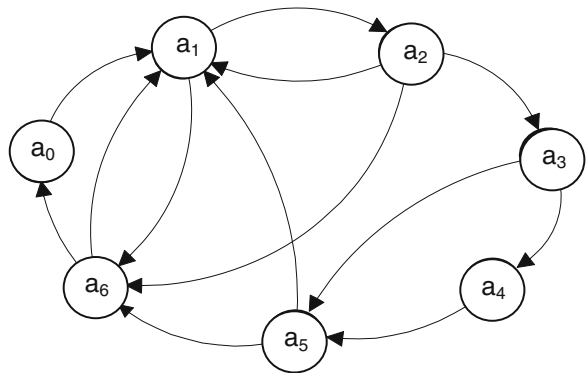
Between the states there exist some relations of admissible sequencing which enable the possible variants of the grasping process:

$$\Omega = \{a_i, a_j\}, \quad a_i, a_j \in A. \quad (19.7)$$

These relations are shown in Fig. 19.6.

Presented sequencing relationship of elementary action are valid regardless of the type of performed grasp (regardless of object being grabbed). However, as follows from the description above, the way of realization of individual actions (excluding the rest position a_0 and the transition to the rest position a_6 connected with it) depends essentially on the type of performed grasp (i.e. on the object being grabbed). Next, different realization of the same actions (different way of moving the finger in the same grasping stage) implicates the different courses of biosignals. Picking out these differences in connection with the knowledge of the actions sequencing relation constitutes the essence of multistage decision control based on the decision tree. Multistage classification enables a better distinction of the intention of specific type grasp execution. For the purpose of control, the described sequencing

Fig. 19.6 Admissible sequencing of the activities in the grabbing process



relation should be expressed in the form of the tree and should differentiate simultaneously individual actions depending on the type of grasp. Considering 10 types of grasp (according to Schlesinger classification) and the sequencing relation presented above, the tree dimension is quite large. To simplify the problem one should take the following constrains: the number of the distinguished grasps should be reduced and the sequencing relation should be limited to the elementary sequencing: $a_0-a_1-a_2-a_3-a_4-a_5-a_6-a_0$.

Let us consider the grasping of following objects: a pen and a credit card (standing in a container), a computer mouse and a cell phone (laying on the table), and a kettle and a tube (standing on the table) (Fig. 19.7).

The sensory feedback from the contact of prosthesis with the object (touch/force on palm distinguished points and arrangement/posture of fingers) introduces new quality to the control model. The first type of information can be represented by the vector:

$$f^i = \langle f_1^i, f_2^i, \dots, f_n^i \rangle, f_j^i \in \{0, 1, \dots, k\} \tag{19.8}$$

which coordinate f_j^i describes the value of the touch/force object to the j th point of the hand (in the i th grip). Coordinates take discrete values from 0 to k , wherein 0 represents no touch and 1—hand contact with the object at the immeasurable strength.

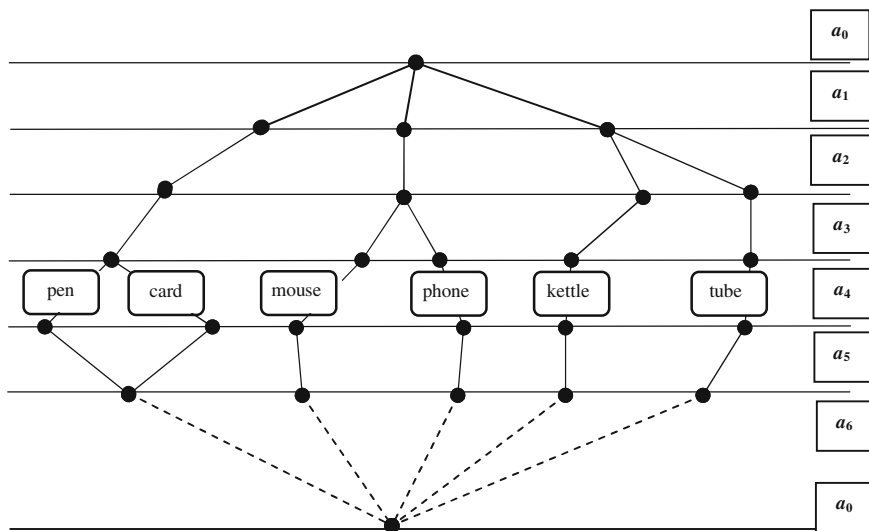


Fig. 19.7 Decision tree for recognition of intent sequence; a_0 —rest stage; a_1 —grasp preparation (3 different types of a finger gape); a_2 —grasp closing (different for the kettle and the tube), a_3 —grasping (different for the mouse and the phone—grasp and turn), a_4 —maintaining the grasp (modification of grasp for the pen); a_5 —releasing the grasp (equal for the pen and the card and different for the remaining items); a_6 —transition to the rest position (different for all items)

The second type of information is expressed vector:

$$\alpha^i = \langle \alpha_1^i, \alpha_2^i, \dots, \alpha_s^i \rangle, \alpha_k^i \in [0, \Pi/2], \quad (19.9)$$

where α_r^i denotes the angular position of the hand-wrist (in i th grip). The concatenation of these vectors characterizes a lump/shape and elasticity of the touched/gripped object, specifying the type of implemented gripping g^i (as well as its execution phase):

$$v^i = \langle f^i, \alpha^i \rangle \longleftrightarrow g^i. \quad (19.10)$$

At the beginning of the grasping process, all the coordinates of the vector f^i are zero. Information from the v^i begins to be significant only in the grasping phase, when it comes into contact of prosthesis finger with the object, i.e. when successive coordinates are beginning to change their values from 0 to 1 (or higher from the interval $[1, k]$). However in the grasping phase, information v^i does not describe explicitly gripping which will be realized. On the other hand, definitely narrows the set of possible grasps, therefore can significantly assist the ongoing decision process.

19.3 Experimental Investigations

In experiments conducted in the Matlab environment, 6 types of grips (tripoid, pinch, power, hook, column and mouse grip) were considered (Fig. 19.8). Biosignals were registered using 3 EMG electrodes and 3 MMG microphones located on a forearm. The dataset set consisted of 400 measurements, i.e. pairs EMG and MMG signals segment/movement class. The values from STFT product corresponding to the $k = 3, 4, 5$ most representative time slices were considered as feature vector. The training and testing sets were extracted from each dataset using two-fold cross-validation. Half of objects from the training dataset were used as a validation dataset and the other half were used for the training of base classifiers.

The performance of the systems constructed (MC—, MC+—without using and with using the signals from the bioprosthesis sensors) were compared against the following three benchmark MCSs [4]: the single best (SB); majority voting (MV); DCS—local accuracy system (LA). Three experiments were performed which differ in the biosignals used for classification (EMG, MMG, both signals). The experiments were conducted using the set of the following ten base classifiers: (1–2) linear (quadratic) classifier based on normal distributions with the same (different) covariance matrix for each class, (3) nearest mean classifier, (4–6) k -nearest neighbours classifiers with $k = 1; 5; 15$, (7) naive Bayes classifier (8) decision-tree classifier with Gini splitting criterion, (9–10) feed-forward backpropagation neural network with 1 hidden layer (with 2 hidden layers).

Classification accuracies (i.e. the percentage of correctly classified objects) for methods tested are listed in Table 19.1 (k denotes the number of time slices per signal



Fig. 19.8 Types of grips

Table 19.1 Classification accuracies of classifiers compared in the experiment (description in the text)

Classifier/Mean accuracy (%)					
k	SB	MV	LA	MC-	MC+
EMG signals					
3	77.2	74.5	78.3	78.5	79.2
4	85.7	83.2	85.1	85.4	84.9
5	90.5	92.6	91.8	93.1	93.3
MMG signals					
3	47.8	43.5	46.8	45.9	45.6
4	52.4	51.2	50.6	54.2	54.5
5	65.8	63.9	65.4	67.2	67.7
MMG and EMG signals					
3	82.5	81.8	83.1	84.3	84.8
4	92.7	92.1	91.9	93.8	93.2
5	95.9	95.1	94.7	96.8	97.1

The best score for each dataset is highlighted

channel). The accuracies are average values obtained over 10 runs (5 replications of two-fold cross validation).

19.4 Conclusion

The following conclusions result from the experiments:

1. Both MC– and MC+ systems produced the best results in 6 out of 9 cases;
2. There are no statistically significant differences between MC– and MC+ systems;
3. Both EMG and MMG signals lead to the highest classification accuracy.

Presented concept of the MC system for the bioprosthetic hand control has a preliminary character. The particular components of the proposed system are at different levels of development. The DES system using an original procedure of fusion/selection based on competence measure of base classifiers is a mature tool, which has proved to be an effective approach in the problem of biosignals recognition. The remaining important components, especially the two level structure of MC system and the algorithm of competence correction using a feedback information from bioprosthetic sensors require both theoretical analysis and experimental investigations. The problem of deliberate human impact on the mechanical device using natural biological signals generated in the body can be considered generally as a matter of human—machine interface. The results presented in this paper significantly affect the development of this field and the overall discipline of signal recognition, thereby contributing to the comprehensive development of civilization. But more importantly, these results will also find practical application in the design of dexterous prosthetic hand—in the synthesis of control algorithms for these devices, as well as development of computer systems for learning motor coordination, dedicated to individuals preparing for a prosthesis or waiting for a hand transplantation [15].

Acknowledgments This work was financed from the National Science Center resources in 2012–2014 years as a research project No ST6/06168, and from Wrocław University of Technology as a statutory project.

References

1. Boostanl, B., Moradi, M.: Evaluation of the forearm EMG signal features for the control of a prosthetic hand. *Physiol. Meas.* **24**, 309–319 (2003)
2. De Luca, C.: Electromyography. In: Webster, J.G. (ed.) *Encyclopedia of Medical Devices and Instrumentation*, pp. 98–109. Wiley, New Jersey (2006)
3. Englehart, K.: Signal representation for classification of the transient myoelectric signal. Ph.D. Thesis, University of New Brunswick, Fredericton, New Brunswick (1998)
4. Kuncheva, I.: *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, New Jersey (2004)
5. Kurzynski, M., Wolczowski, A.: Dynamic selection of classifier ensemble applied to the recognition of EMG signal for the control of bioprosthetic hand. In: *Proceedings of 11th International Conference on Control, Automation and Systems*, Seoul, pp. 175–182 (2011)
6. Kurzynski, M., Wolczowski, A., Tito, A.: Control of bioprosthetic hand based on EMG and MMG signals recognition using multiclassifier system with feedback from the prosthesis sensors. In: *Proceedings of 2nd International Conference on Systems and Control*, Marrakesh, pp. 127–135 (2012)

7. Kurzynski, M., Wolczowski, A.: Classification of EMG signals in a system for training of bioprosthetic hand control in one side handless human. In: Proceedings of International Conference on Electrical Engineering and Computer Science EECS12, Szanghai, pp. 566–576 (2012)
8. Nishikawa, D.: Studies on Electromyogram to Motion Classifier. Ph.D. Thesis, Graduate School of Engineering, Hokkaido University, Sapporo (2001)
9. Orizio, C.: Muscle sound: basis for the introduction of a mechanomyographic signal in muscle studies. *Crit. Rev. Biomed. Eng.* **21**, 201–243 (1993)
10. Wolczowski, A.: Smart hand: the concept of sensor based control. In: Proceedings of 7th IEEE International Symposium MMAR, Miedzyzdroje, pp. 783–790 (2001)
11. Wolczowski, A., Kurzynski, M.: Control of Artificial Hand Via Recognition of EMG Signals, LNCS, vol. 3337. Springer, New York (2004)
12. Wolczowski, A., Kurzynski, M.: Control of dexterous hand via recognition of EMG signals using combination of decision-tree and sequential classifier. *Adv. Soft Comput.* **45**, 687–694 (2007)
13. Wolczowski, A., Krysztoforski, K.: Artificial hand control via EMG signal classification—experimental investigation of algorithms. In: Tchon, K., Warszawa, W.K.L. (ed.) *Progress in Robotics*, pp. 97–122 (2008)
14. Wolczowski, A., Kurzynski, M.: Human-machine interface in bio-prosthesis control using EMG signal classification. *Expert Syst.* **27**, 53–70 (2010)
15. Wolczowski, A., Kurzynski, M., Zaplotny, P.: Concept of a system for training of bioprosthetic hand control in one side handless humans using virtual reality and visual and sensory biofeedback. *Med. Inf. Technol.* **18**, 85–91 (2012)
16. Woloszynski, T., Kurzynski, M.: On a new measure of classifier competence applied to the design of multiclassifier systems. *Lect. Notes Comput. Sci.* **5716**, 995–1004 (2009)
17. Woloszynski, T., Kurzynski, M.: A probabilistic model of classifier competence for dynamic ensemble selection. *Pattern Recognit.* **44**, 2656–2668 (2011)
18. Zecca, M., Micera, S., Carrozza, M., Dario, P.: Control of multifunctional prosthetic hands by processing the electromyographic signal. *Crit. Rev. Biomed. Eng.* **30**, 459–485 (2002)

Chapter 20

Transcutaneous Bladder Spectroscopy: Computer Aided Near Infrared Monitoring of Physiologic Function

Andrew Macnab, Lynn Stothers, Babak Shadgan and Behnam Molavi

Abstract Transcutaneous bladder spectroscopy is a novel non-invasive optical technique that uses near infrared (NIR) light to detect changes in concentration of the tissue chromophores oxy and deoxy-hemoglobin in the anterior wall of the organ as it fills and empties. From the patterns, trends and magnitude of these changes alterations in tissue oxygenation and hemodynamics can be inferred, and, as these differ in health and disease, novel information is gained related to the pathophysiology underlying the causation and symptoms of voiding dysfunction. Following proof of concept and a series of clinical studies, evolution of NIR spectroscopy (NIRS) devices for bladder study now requires computer aided solutions to optimize device design and refine software development to further advance application of this technology. We review the concept of bladder spectroscopy, summarize the evolution of related NIRS hardware and software, and outline the challenges and potential for further computer aided development.

20.1 Introduction

Near infrared spectroscopy (NIRS) is an optical technique that uses near infrared (NIR) light to monitor changes in local blood volume and detect differences in tissue oxygen delivery, consumption, and utilization [1–4]. Photons of NIR light

A. Macnab (✉) · L. Stothers
Department of Urology, University of British Columbia, Vancouver, Canada
e-mail: amacnab@cw.bc.ca

L. Stothers
e-mail: lynn.stothers@ubc.ca

B. Shadgan
Center for Collaboration on Repair Discoveries, University of British Columbia,
Vancouver, Canada
e-mail: shadgan@alumni.ubc.ca

B. Molavi
Department of Electrical and Computer Engineering, University of British Columbia,
Vancouver, Canada
e-mail: molavi@pathonix.com

(700–1300 nm) readily pass through skin, subcutaneous tissue and bone, and scatter in tissue [5], where they are absorbed in varying amounts by the chromophore hemoglobin, depending on its oxygenation status (oxygenated O₂Hb or deoxygenated HHb). The chemical structure, color, and concentration of hemoglobin influences how NIR light of different wavelengths is absorbed. It is the unique relationship between the transparency of tissue to NIR light and the different extinction coefficient (absorption characteristics) of O₂Hb and HHb across the NIR spectrum that provide the basis for biomedical applications of NIRS [1, 5–7].

The principal role of computing in NIRS is the provision of software algorithms employed to convert the raw optical data into chromophore concentration changes, using a modification of the Lambert-Beer law [8, 9]. These algorithms also accommodate for a number of limitations posed by the nature of human tissue, including where the path length of NIR light and number of photons lost due to scattering are unknown [8]. Further computer-aided data processing is required to provide a real time graphic display which allows the trend data for O₂Hb, HHb and their sum total hemoglobin (tHb) to be viewed and compared. Most devices incorporate proprietary software that displays changes in the concentration of O₂Hb and HHb in this manner, and such graphic display allows changes in tissue oxygenation and hemodynamics to be inferred from the changes in chromophore concentration.

Biomedical applications of NIRS can already provide a number of unique measurements, and physiologic information relevant to better understanding of the basis for disease including parameters that, because they differ in health and disease, offer a means of diagnosis where appropriate algorithms are generated. Such applications represent a disruptive technology especially in medicine, because they involve non-invasive transcutaneous measurement using a non-toxic energy source, the information provided to scientists and clinicians is not available by other means, and importantly the technology is readily accepted by patients.

One novel biomedical application of NIRS uses transcutaneous NIR light transmission to investigate physiologic function in the anterior wall of the bladder. The data obtained offers additional parameters of relevance for evaluation of patients with voiding dysfunction that is not provided by other tests [10, 11]. In the initial bladder studies, continuous wave (CW) NIRS instruments with lasers as the NIR light source were used [11, 12]. A sensor incorporated in the patient interface with the NIR emitter detected the photons returning from tissue that were neither scattered nor absorbed by chromophores. Proprietary software algorithms are used for the initial conversion of the raw optical data into chromophore concentration changes from baseline [2, 9]. Additional software has been developed to improve the quality of the data displayed by filtering noise and other artifact, and to modify how the parameters are displayed in relation to one another and to specific time points in the study cycle, in order to optimize interpretation of the physiologic relevance of the data.

In biomedical applications employing NIRS the principal chromophores monitored are O₂Hb and HHb. The information provided is the change in concentration in O₂Hb and HHb over time relative to their initial baseline concentration, because the full extent of the tissue monitored and hence the total amount of hemoglobin present is always unknown *in vivo*. However, the computer-generated patterns and magnitude of

change in chromophore concentration and tHb obtained in real time allow variations in tissue oxygenation and hemodynamics to be inferred [3, 5, 12]. Such changes include: an increase or decrease in O₂Hb (an indirect measure of oxygen content); an increase or decrease in the sum of O₂Hb and HHb, tHb (a change in blood volume); a gradual decrease in O₂Hb and matching increase in HHb (hypoxia); and, an abrupt decrease in O₂Hb with simultaneous fall in tHb and increase in HHb (ischemia).

20.2 Near Infrared Spectroscopy (NIRS)

NIRS hardware and software are continuously being refined, resulting in important changes that have made monitoring easier, and improved the processing and display of data. Notable advances include: (1) the reduction in size of the instruments; (2) advanced options for NIR photon detection; (3) use of light emitting diodes (LEDs) as an alternative to lasers as the NIR light source; (4) parallel development of miniature, self-contained wireless devices made possible by LED technology; (5) improvement in the reliability and scope of software algorithms; and (6) configuration of multiple LED's to provide the optical geometry required for spatially resolved spectroscopy (SRS) [12, 13].

20.2.1 Instrumentation

Instruments have become progressively smaller and more portable due to multiple design improvements and availability of better components, including computers.

20.2.2 Options for Photon Detection

Photon detection has improved progressively. Initially the majority of instruments required photo multiplier tubes (PMT's). PMT's were large, and in addition required 2h to heat prior to monitoring, two elements that placed constraints on clinical monitoring in particular. Evolution of detection technology resulted in instruments designed with charge-coupled detectors or synchro-scan cameras, and then photo-diodes, avalanche initially and now silicon (Fig. 20.1).

20.2.3 Light Emitting Diodes

The advantages of LEDs led to design advances for multiple reasons, particularly their low cost and lower power consumption than lasers, and small size. Also, the

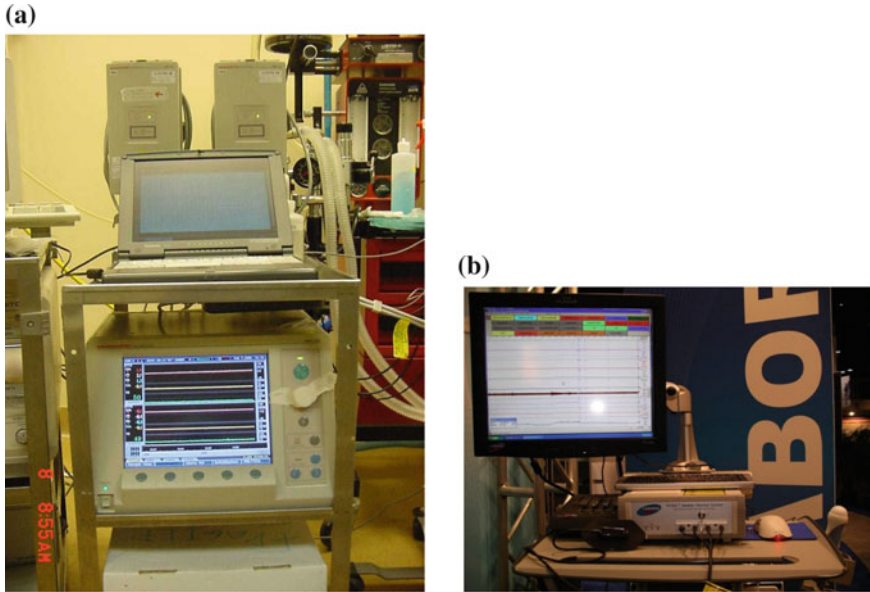


Fig. 20.1 **a** An early spectroscopy instrument with photo multiplier tubes for photon counting, **b** the first commercial iteration of a small laser powered NIRS device with photodiode photon counting designed for bladder studies

fact that fiber-optic cables could be dispensed with, which made a new generation of small self-contained devices possible. The minimal power consumption required by LEDs allowed devices incorporating them to be powered by small batteries, yet, as a LED light source is non-coherent and non-collimated, high light intensity levels were available [10, 12, 14].

Very small self-contained wireless NIRS devices making use of LEDs now enable patients to be monitored more simply than with the original, larger laser powered instruments. It is also possible to monitor subjects with self-contained devices when they are undertaking a range of laboratory based exercise protocols, and when engaged in active physical pursuits including sporting activities. Such devices also make it feasible to study small children, and to monitor ambulant patients over time.

20.2.4 Equations and Algorithms

All spectroscopic instruments require software that addresses the relationship between the absorption of light and the medium it travels through. The convention is to use an equation based on the modified Beer Lambert law. This is derived from a combination of contributions attributed to Lambert and Beer respectively. ‘Each layer of equal thickness of the medium absorbs an equal fraction of the energy

traversing it' and: 'The absorptive capacity of a dissolved substance is directly proportional to its concentration in solution'. Interestingly, Lambert quoted from earlier work by Bouguer which he actually cited, and it was almost a century later that Beer added to his exponential absorption law to include the concentrations of solutions in the absorption coefficient [8, 15].

In the context of NIRS the modified Beer Lambert law is expressed as:

$$A_\lambda = -\log \frac{I}{I_0} = \left(\sum_i \varepsilon_i(\lambda) c_i \right) BL + G \quad (20.1)$$

where A_λ is the light intensity attenuation at wavelength λ , I_0 is the source intensity, I is detected light intensity, $\varepsilon_i(\lambda)$ is the extinction coefficient of chromophore i at wavelength λ , c_i is the chromophore concentration, L is the distance between source and detector (referred to as the inter-optode distance (IOD) in spectroscopy), B is the differential path length factor (DPF), and G is an additive term to address fixed scattering losses.

The inherent assumptions of biomedical NIRS where the modified Beer/Lambert equation is employed dictate that it can only be used to determine absolute changes in concentration relative to the initial baseline at the start of data collection. Consequently, NIRS chromophore data are interpreted based on comparison of the patterns of change, and their magnitude and rate of change.

Individual algorithms differ between NIRS instruments. And an important contribution has been the identification and correction of errors contained in some early versions of incorporated software that compromised the validity of data interpretation reported in early research [16]. Significant efforts have been made to refine and improve the algorithms now incorporated in modern equipment [17, 18].

20.2.5 Spatially Resolved Spectroscopy (SRS)

NIRS devices configured for SRS allow an absolute measure of tissue oxygen saturation to be obtained. SRS requires two or more emitters positioned at different distances from the detector, so NIR light intensity is measured as a function of distance (see Fig. 20.2—bottom). With appropriate software algorithms incorporated the ratio of oxygenated to total tissue hemoglobin can be calculated from which an absolute measurement of tissue oxygen saturation is made [13, 19, 20]. Having an absolute measure of tissue oxygen saturation is clinically useful, although values predominantly reflect venous oxygen saturation, as only the minority of the blood in tissue is in capillaries and arterioles. However, the values derived from different SRS devices are not directly comparable to one another, due to a number of key differences, which include variations in the software algorithms employed [12].

CW SRS NIRS instruments of small size and with telemetric capacity are now available and represent an important advance in biomedical applications of NIRS.

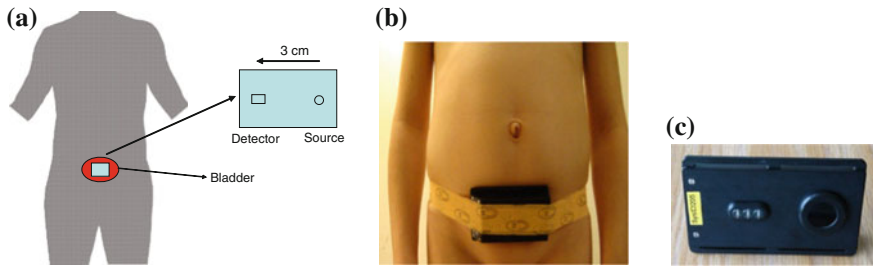


Fig. 20.2 **a** location of NIRS device for transcutaneous bladder monitoring on the abdominal skin, across *midline* superior to the pubis. **b** A self-contained wireless NIRS device with LED's taped over the bladder in a child. **c** Underside of wireless device showing the patient interface, with 3 LED's configured for SRS by each being positioned at a different distance from the photodiode detector to provide different 3 inter optode distances

Such instruments are the ones currently used to monitor changes in bladder hemodynamics, oxygen supply and demand, and to measure absolute tissue oxygen saturation in the anterior wall of the bladder [10, 12, 19, 21]. The evolution of wireless NIRS devices and the related computer aided graphics and algorithms they require demonstrates the innovation and interdisciplinary collaboration between clinicians, researchers, engineers and computer scientists now inherent in computer aided design system concepts [22]. Miniature self-contained devices with LEDs and SRS capacity make non-invasive transcutaneous bladder monitoring straightforward and an attractive option for clinicians and their patients, including children, the elderly and those with special needs (e.g. post spinal cord injury) [23].

Hence, transcutaneous bladder spectroscopy is an advance of relevance in medicine, and current wireless devices offer new options for research and an avenue for progress towards other biomedical applications of NIRS.

20.3 Physiology Relevant to Bladder NIRS

The bladder is uniquely vascularized so that the microcirculation can maintain perfusion and oxygen delivery as the organ expands as it fills with urine and contracts during voiding. Various diseases compromise the ability of the microcirculation to maintain perfusion during voiding, and/or affect the mechanics and physiology of bladder contraction [10].

As the bladder empties, contraction of the detrusor muscle occurs, which requires the provision of additional blood flow via the microcirculation to supply the oxygen and substrates required to meet the metabolic demands of contraction. Individual components of the physiologic processes required to do this and their net effect can be monitored through changes generated in O_2Hb and HHb concentration in the microcirculation of the bladder wall. Variations occurring in the organ's hemodynamics

and oxygen supply and demand can then be inferred from these changes, and this information is not available by other means, including the current ‘gold standard’ test (UDS) used to evaluate the bladder in voiding dysfunction. Because UDS only measures pressure and flow and is also invasive as it requires transurethral and rectal catheterization, non-invasive transcutaneous optical monitoring of the bladder using NIRS is both relevant and clinically desirable [11, 19, 24].

20.4 Data Processing and Analysis

Software able to convert the raw optical data obtained into changes in chromophore concentration is also an essential component incorporated into NIRS instruments. Data are displayed as real time trends and patterns of change in chromophore concentration from baseline [2, 9].

The value of such data during bladder spectroscopy is that clear differences are seen between healthy asymptomatic patients and those with disease. This is particularly true for chromophore data obtained during voiding as the bladder contracts to expel urine, and, importantly in the context of the validity of the physiologic information inferred, the trends and patterns of chromophore change observed ‘match’ NIRS data from other tissues where altered hemodynamics and impaired oxygen supply and demand occur. Hence, bladder NIRS provides novel insights into the underlying causes of voiding dysfunction.

Figure 20.3 illustrates an example of the multiple data streams that must be incorporated into a composite time synchronized graphic display in our application of bladder spectroscopy, where we interrogate the organ simultaneously with non-invasive transcutaneous NIRS during the invasive pressure flow studies that are the current ‘gold standard’ diagnostic procedure. Such composite displays are necessary to allow comparison of the interrelationship of multiple simultaneous events and parameters.

In some instances, we have also found it beneficial to improve the clarity of NIRS-derived data by using software-based smoothing filters which incorporate a “moving Gaussian” or “moving average” process. The resulting data are easier to analyze because any interference resulting from high frequency signal noise present is then significantly reduced. An example of data pre and post application of such software is shown in Fig. 20.4. An additional approach that we find optimizes our ability to interpret the trends and relationships between changes in chromophore concentration is to modify the way the 3 NIRS parameters are displayed in relationship to each other. To do this we bias the signals for each chromophore to zero at the chosen point of interest in the voiding cycle, e.g. ‘permission to void’ or ‘start of uroflow’.

Artifact removal

Another way software based approaches can improve the clarity and quality of the data is through detection and reduction of interference in the signal. Such signal

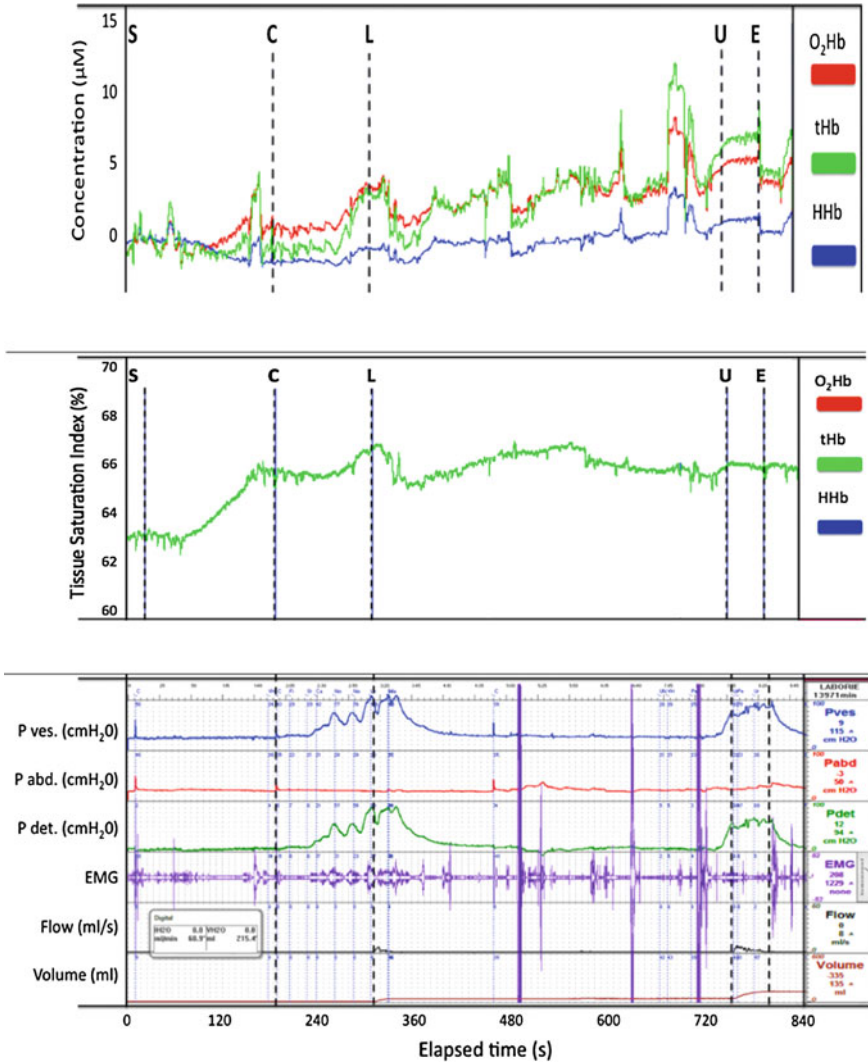


Fig. 20.3 A composite graph of the multiple data streams that software program design must incorporate into a single, real time graphic display when noninvasive NIRS bladder studies are done simultaneously with the current ‘gold standard’ invasive urodynamic pressure flow studies (UDS). Such combined studies monitor patients as the bladder fills and then empties recording S = start of filling, C = cough, L = leakage of urine, U = start of uroflow (voiding), E = end of uroflow. Data streams include: O₂Hb, tHb and HHb (*top graph*) Tissue saturation index (*middle*), and vesical, abdominal and detrusor pressure, perineal electromyogram (EMG), and urine flow and volume (*bottom*)

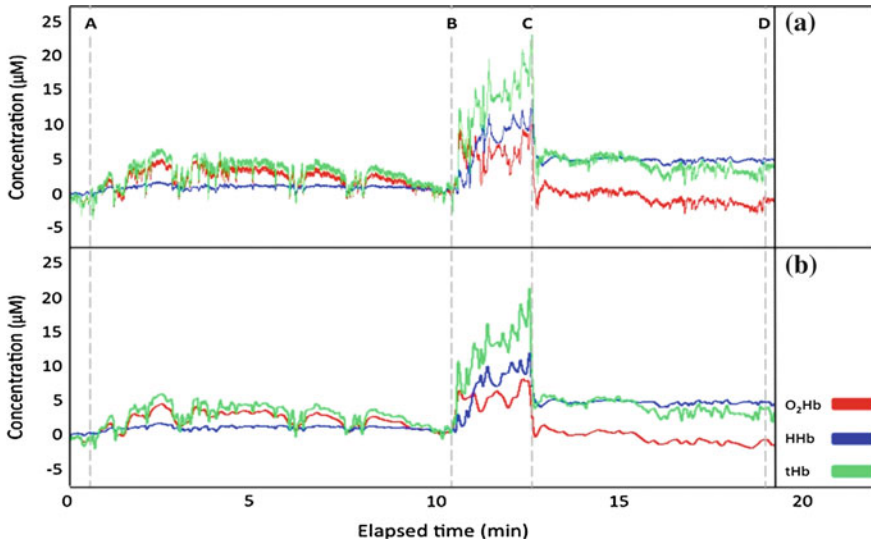


Fig. 20.4 Composite graph of 2 versions of the same NIRS data set from a paraplegic patient during voiding. Graph **a** shows the graphic display of O₂Hb, HHb and tHb following conversion of the raw optical data by proprietary software incorporated in the device. While Graph **b** shows the same data following application in addition of a moving average filter. The abrupt unidirectional changes (>3–4 µM) that occur simultaneously in all 3 chromophore parameters are due to movement caused by involuntary muscle spasms. A = start of bladder filling, B–C = movement as patient repositioned, D = filling ends

‘noise’ has various forms including systemic interference from the cardiac cycle and respiration and motion artifact.

Cardiac pulsation and the systemic effects of respiration may be evident in NIRS data. Although physiologic in nature and able to provide valuable information in some instances, the presence of these parameters in the data stream may compromise desired elements of interpretation at other times. For this reason, software based methods can be employed for reduction of this form of interference e.g. as in fNIRS data [25, 26].

Motion artifacts are non-physiological-changes in the NIRS signal resulting from movement by the subject. Motion artifact can obscure information in the data, although in bladder spectroscopy studies we find the important physiologic trends and magnitude of change are evident clearly enough for this not to be a problem [24]. However, in NIRS muscle studies in actively exercising athletes for instance, the data quality can be improved by reduction/removal of the effects of motion. And, in the past, the effects of movement have proved to be a limiting factor in clinical applications of NIRS in ambulant patients, and also during experimental applications of NIRS monitoring in exercise science and sports medicine.

Hence various signal processing and software based methods have been proposed to identify and remove or reduce the effect of these artifacts from NIRS data [27, 28].

These methods use the temporal or spectral features of the artifacts to identify them. This includes using the negative correlation between O_2Hb and HHb [29], using a moving standard deviation scheme to detect motion and applying spline interpolation [30], and application of Wiener filtering and Kalman filtering with appropriate model on noise and data to reduce motion artifacts [31, 32]. Wavelet decomposition is another promising approach for detection and removal of interference from fNIRS signals. Continuous Wavelet transform can be used to detect blocks contaminated with motion artifact using a hard threshold on the wavelet transform amplitudes [33]. Methods based on the Discrete Wavelet Transform (DWT) have also been proposed to detect and reduce the effect of motion [34, 35]. These methods for reduction of interference can be readily incorporated into NIRS processing software. Open source software, such as HOMER2,¹ developed by MGH-Martinos Center for Biomedical Imaging, are available for processing of NIRS data, and such software incorporates some of the above mentioned methods.

Motion artifact in NIRS data streams can also be reduced using adaptive filtering. In this approach, data from a reference channel which has high correlation with motion is utilized to reduce the motion artifacts. Different forms of reference channels, including an accelerometer [35–37] and an auxiliary optical channel can be used for this purpose [34, 38]. The optical reference channel is usually selected such that it samples from superficial layers of the tissue to reduce its correlation with NIRS signal of interest from deeper tissue layers.

20.4.1 Diagnostic Algorithms

Data obtained from bladder NIRS have made it possible to develop diagnostic algorithms. Such algorithms can identify the presence or absence of bladder outlet obstruction (BOO) due to prostate enlargement. The key elements incorporated reflect NIRS-derived differences in detrusor hemodynamics. The first algorithm developed used NIRS data combined with a measure of peak flow and the post-voiding volume of residual urine volume. The second was a classification and regression tree (CART) algorithm that used NIRS data alone from the entire voiding cycle [39, 40]. Both have comparable discriminant ability, good specificity and sensitivity, and concordance with the diagnostic algorithm used for the current ‘gold standard’ UDS diagnosis of BOO based on pressure flow data [41].

Classification and regression trees are machine-learning methods that construct prediction models by partitioning the data and fitting a simple prediction model within each partition [42]. The resulting model can be represented as a decision tree. Classification trees are designed for dependent variables that take a finite number of unordered values. In contrast, regression trees are for dependent variables that take continuous or ordered discrete variables [43]. The binary prediction trees that result can be further improved by inclusion of multiway splits [44].

¹ www.nmr.mgh.harvard.edu/optics.

Confidence that bladder NIRS measures physiologic changes comes from: the physics of penetration, scattering and chromophore absorption of NIR photons, and the consistent anatomic position of the bladder and ultrasound evidence that the anterior wall maintains its spatial geometry with a NIRS device on the lower abdominal skin during voiding. Also, changes are only detected over the bladder (not from control emitter/detectors elsewhere on the abdomen) and only occur in relation to events in the voiding cycle [10, 19, 24].

20.5 Computer Aided Systems Engineering

Computer systems engineering (CSE) is a process that controls the development of a technical system with the objective of achieving an optimum balance of all the system's elements. CSE transforms the goals of the design into clearly defined system parameters, and allocates and integrates those parameters to the various development disciplines needed to realize the systems, products and processes [22].

The process of computer aided design (CAD) utilizes increasing complex and diverse computer software programs to create, modify, analyze and optimize the design, inter-disciplinary interaction and production of a growing range of devices and products. The disciplines involved commonly include mathematics, computing, engineering and social sciences. Definition of the goals of the design lead to 'problem definition', followed by 'model construction', 'model analysis', 'implementation' and ultimately, 'problem resolution'. Data from subsequent trials of the system generated can be fed back into the model for further modification. Amongst the advantages of CAD programs are the ability to view and rotate the model in 3D and see the mechanical effects of substituting or modifying components of structural materials [45, 46].

Our current development program includes interdisciplinary collaboration across several of these disciplines. We are exploring the use of CAD software programs, but are to date not aware of any NIRS devices that have taken advantage of the CAD process. We have used 3D printing in the production of our prototypes for use in Africa. This has made it possible to encapsulate the emitters and the sensor in an extruded feature that has improved the coupling of the emitter/detector interface with the skin, and also achieves significant reduction in interference by ambient light [47].

20.6 Design: Process, Opportunities, and Challenges

An essential step in the process of developing a new biomedical device is to compare it to the current 'gold standard' technology used for diagnosis, or, where a similar instrument exists to the extant device itself. This step has been applied in developing the technology and software for our NIRS bladder screening concept. Our initial laser powered design [48] was tested against comparable commercial NIRS

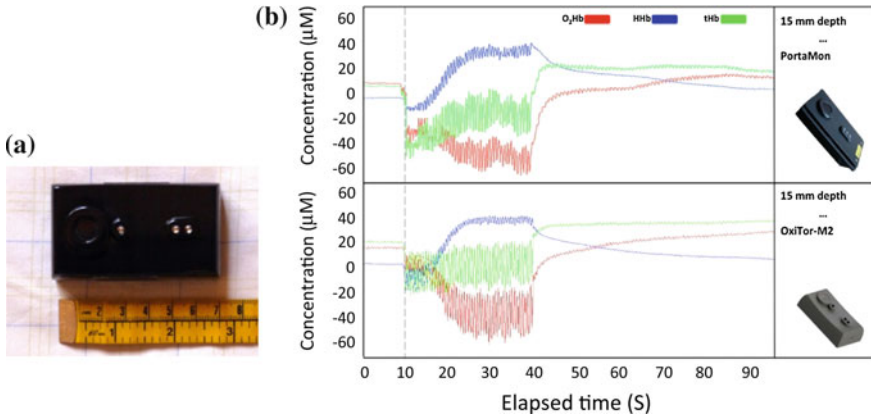


Fig. 20.5 **a** Prototype of NIRS bladder device for use as a non-invasive screening tool for bladder disease in Africa; **b** Paired graphs of data obtained during the design stage when comparing the prototype (OxiTor) to our reference ‘gold standard’ instrument (Portamon). The data sets are from a single patient where the same protocol was applied using each device in turn to monitor the effects of forearm ischemia generated during an isotonic exercise regimen. The patterns and magnitude of change in chromophore data (O₂Hb, HHb and tHb) are comparable between devices. The effects of movement during exercise are visible. No motion artifact removal software has been applied, however the physiologic trends of each data stream remain clear

instruments and then through simultaneous use with current ‘gold standard’ invasive UDS diagnostic equipment [10, 39]. Our current wireless prototype [49] was compared first in studies using a forearm ischemia protocol against the commercial instrument used to date for all our bladder spectroscopy studies, and Fig. 20.5 shows comparison graphs obtained during isotonic exercise for our device and the reference instrument. The data obtained allow the trend and magnitude of chromophore change to be seen clearly, however, as discussed above, the effect of movements inherent during sustained exercise on raw data can be addressed by solutions to improve the clarity of the chromophore data. Further evolution of device design and software for this purpose will undoubtedly be ongoing.

While devices designed to incorporate LED’s as the NIR light source already represent a major advance, in the context of design, elements of choice and compromise are required when using LEDs in self-contained wireless devices. In order to optimize the benefits of LEDs, other components such as detectors, batteries, and telemetric communication components in particular have to be matched to them carefully. For example, the inclusion of ‘Bluetooth’ capacity, while ideal for potential linkage to remote data storage modules, and especially to mobile phones, does use more power than other wireless communication systems, which results in the need to incorporate a battery with higher capacity. However, the additional size and weight of such a battery can be justified where the specifications of the ‘Bluetooth’ component meets specific design goals for the device. In this context a current attraction of using ‘Bluetooth’ is its ability to link several wireless devices simultaneously from a single module.

This illustrates the necessity for the design process to weigh the precise functional goals of a device against the pros and cons of including any given component, so as to ensure that there is synergy between every component ultimately included, and the ultimate design characteristics of the device are delivered as optimally as possible.

Other current challenges related to the design of NIRS devices include many related to the properties of light transmission through tissue, the modification of algorithms based on in-vitro concepts to be applied in an in-vivo context, and the practical challenge of monitoring the anterior wall of the bladder during voiding when the wall thickness and size of the bladder alter.

One issue, for example, relates to the decrease in signal power that occurs due to reduced photon penetration through pigmented skin and secondary to scattering and attenuation of light in the tissue of interest. It would be an obvious advantage to improve the signal to noise ratio at the detector output to address both issues. However, while the desired effect can be achieved by increasing the source optical power, power must be kept within a safe range to prevent any risk of tissue damage from exposure to excessive energy.

Hence, the design process must also address the relative advantages and disadvantages of any modification intended to optimize any element of the design. The potential to improve performance must be weighed against the practicality of achieving the desired change, the need to adhere to the specific performance goals of the device, and the ability of the system as a whole to operate within the required limits for safety. To do this optimally for future iterations of bladder NIRS devices, it is clear from advances made in the design of other biomedical devices that CSE concepts should be comprehensively incorporated into the development process and the use of CAD programs explored.

20.7 Future Directions

As the relative complexity and technical specifications for our prototype NIRS devices increase, we will recreate them using 3D CAD (computer aided design) software. Model validation will then be followed by computer based refinements to the initial design and iterative development of a suitable 'second generation' product for clinical trials. Design flaws identified and potential improvements evident from these trials will be addressed by further solutions and enhancements made to the virtual model. The purpose is to ensure that the product ultimately developed benefits from testing enabled by the CAD simulation package.

One specific goal of our design process is to enhance working prototypes of the simplified NIRS device we have developed for use by rural health workers in low income countries as a screening tool for early detection of bladder outlet obstruction (BOO). This goal adds a number of novel dimensions to NIRS device design, including the need to ensure that inclusion of the most inexpensive components does not compromise the required performance characteristics of the device, and that the

intended users have a product and data collection interface compatible with their experience of other technologies and level of expertise.

Renal failure due to progression of undetected BOO is a major health problem in Africa, associated with significant morbidity and early death in adult life [50]. BOO occurs secondary to conditions that include the common, age-related, enlargement of the prostate gland (prostatic hyperplasia). Undetected BOO progresses, and patients who present in countries like Africa with advanced disease can only usually be offered a palliative treatment option such as bladder drainage through an indwelling suprapubic catheter. The effect of this treatment is limited to relief of symptoms, and having a catheter in place carries a significant morbidity due to infection, as well as negatively impacting a patient's quality of life to a major degree. Detection of the onset of a degree of BOO requiring treatment intervention by using NIRS bladder screening will allow referral to a centre able to provide care. Such care is available to arrest BOO and prevent progression to end stage disease and potential early death, provided the diagnosis is made early enough in the evolution of the disease.

In addition to optimal device design to achieve an inexpensive, robust and user-friendly version of the NIRS device, the CSE process will be investigated as a means optimizing the computer aided transfer of screening data via cellphone linkage to an administrative centre for diagnostic analysis. Population specific diagnostic algorithms will be developed that are able to reliably identify those with and without sufficient BOO to require referral from rural areas to a centre able to offer interventional treatment. Hence in addition to the advantage of the NIRS-based screening system identifying those in need of treatment, the data transfer and diagnostic components will also make it possible to rationalize the use of limited specialist services in low and middle income countries; examples of areas where such benefits would apply include much of sub-Saharan Africa and rural China.

Acknowledgments The authors wish to acknowledge funding support from Grand Challenges Canada, "Stars in Global Health" program for elements of this project. And the Child and Family Research Institute, British Columbia Children's and Women's Hospital, Vancouver, Canada for a special funding award to AJM for development of biomedical applications in photonics.

References

1. Van Beekvelt, M.C., Colier, W.N., Wevers, R.A., Van Engelen, B.V.M.: Performance of near-infrared spectroscopy in measuring local oxygen consumption and blood flow in skeletal muscle. *J. Appl. Physiol.* **90**(2), 511–519 (2001)
2. Ferrari, M., Mottola, L., Quaresima, V.: Principles, techniques and imitations of near infrared spectroscopy. *Can. J. Appl. Physiol.* **29**(4), 463–487 (2004)
3. Wolf, M., Ferrari, M., Quaresima, V.: Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *J. Biomed. Opt.* **12**, 062104 (2007)
4. Hamaoka, T., McCully, K.K., Niwayama, M., Chance, B.: The use of muscle near-infrared spectroscopy in sport, health and medical sciences: recent developments. *Philos. Trans. R. Soc. Lond. A* **369**(1955), 4591–4601 (2011)

5. Boushel, R., Langberg, H., Olesen, J., Gonzales-Alonzo, J., Bulow, J., Kjaer, M.: Monitoring tissue oxygen availability with near infrared spectroscopy (NIRS) in health and disease. *Scand. J. Med. Sci. Sports* **11**(4), 213–222 (2001)
6. Macnab, A.J.: Biomedical applications of near infrared spectroscopy. In: Barth, A., Haris, P.I. (eds.) *Biological and Biomedical Spectroscopy Volume 2 Advances in Biomedical Spectroscopy*, pp. 340–343. IOS Press, Amsterdam (2009)
7. Delpy, D.T., Cope, M.: Quantification in tissue near-infrared spectroscopy. *Philos. Trans. R. Soc. Lond. B* **352**(1354), 649–659 (1997)
8. Ferrari, M., Binzoni, T., Quaresima, V.: Oxidative metabolism in muscle. *Philos. Trans. R. Soc. Lond. B* **352**(1354), 677–683 (1997)
9. Delpy, D.T., Cope, M., van der Zee, P., Arridge, S., Wray, S., Wyatt, J.S.: Estimation of optical path length through tissue from direct time of flight measurements. *Phys. Med. Biol.* **33**(12), 1433–1442 (1988)
10. Macnab, A.J., Shadgan, B., Stothers, L.: The evolution of wireless near infrared spectroscopy: applications in urology and rationale for clinical use. *J. NIRS* **20**(1), 57–73 (2012)
11. Pannek, J.: Editorial comment on: classification of male lower urinary tract symptoms using mathematical modelling and a regression tree algorithm of noninvasive near-infrared spectroscopy parameters. *Eur. Urol.* **57**(2), 332–333 (2010)
12. Macnab, A.J., Shadgan, B.: Biomedical applications of wireless continuous wave near infrared spectroscopy. *Biomed. Spectrosc. Imaging* **1**(3), 205–222 (2012)
13. Suzuki, S., Takasaki, S., Ozaki, T., Kobayashi, Y.: A tissue oxygenation monitor using NIR spatially resolved spectroscopy. *Proc. SPIE* **3597**, 582–592 (1999)
14. Bozkurt, A., Rosen, A., Rosen, H., Onaral, B.: A portable near infrared spectroscopy system for bedside monitoring of newborn brain. *Biomed. Eng. Online* **4**, 29 (2005)
15. Cope, M., Delpy, D.T., Wray, S., Wyatt, J.S., Reynolds, E.O.R.: A CCD spectrometer to quantitate the concentration of chromophores in living tissue utilizing the absorption peak of water at 975 nm. *Adv. Exp. Med. Biol.* **248**, 33–40 (1989)
16. Macnab, A.J., Gagnon, R.E.: Potential sources of discrepancies between living tissue near infrared spectroscopy algorithms. *Anal. Biochem.* **236**(2), 375–377 (1996)
17. Cooper, C.E., Springett, R.: Measurement of cytochrome oxidase and mitochondrial energetics by near-infrared spectroscopy. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **352**(1354), 669–676 (1997)
18. Piantadosi, C.A., Hall, M., Comfort, B.J.: Algorithms for in vivo near-infrared spectroscopy. *Anal. Biochem.* **253**(2), 277–279 (1997)
19. Macnab, A.J., Shadgan, B., Stothers, L., Afshar, K.: Ambulant monitoring of bladder oxygenation and hemodynamics using wireless near-infrared spectroscopy. *Can. Urol. Assoc. J.* **7**((1—2)), E98–E104 (2012)
20. Matcher, S., Kirkpatrick, P., Nahid, K., Cope, M., Delpy, D.T.: Absolute quantification methods in tissue near infrared spectroscopy. *Proc. SPIE* **2389**, 486–495 (1995)
21. Shadgan, B., Afshar, K., Stothers, L., Macnab, A.J.: Near-infrared spectroscopy of the bladder: a new technique for studying lower urinary tract function in health and disease. *Proc. SPIE* **7548**, 754804 (2010)
22. Madsen, D.A.: *Engineering Drawing and Design*, p. 10. Clifton Park, Delmar (2012)
23. Shadgan, B., Macnab, A.J., Stothers, L., Nigro, M.: Monitoring of lower urinary tract function in patients with spinal cord injury using near infrared spectroscopy. *Proc. SPIE* **8027**, 802717 (2012)
24. Macnab, A.J., Shadgan, B., Stothers, L.: Monitoring detrusor oxygenation and hemodynamics non-invasively during dysfunctional voiding. *Adv. Urol.* (2012), Article ID 676303, 8, (2012). doi:[10.1155/2012/676303](https://doi.org/10.1155/2012/676303)
25. Zhang, Y., Brooks, D.H., Franceschini, M.A., Boas, D.A.: Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *J. Biomed. Opt.* **10**(1), 11014 (2005)
26. Jang, K.E., Tak, S., Jung, J., Jang, J., Jeong, Y., Ye, J.C.: Wavelet minimum description length detrending for near-infrared spectroscopy. *J. Biomed. Opt.* **14**(3), 034004 (2009)

27. Cooper, R.J., Selb, J., Gagnon, L., Phillip, D., Schytz, H.W., Iversen, H.K., Ashina, M., Boas, D.A.: A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Front. Neurosci.* **6**, 147. doi:10.3389/fnins.2012.00147
28. Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, P.J., Gagnon, L., Boas, D.A., Cooper, R.J.: Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *NeuroImage* **85**(1), 181–191 (2014)
29. Cui, X., Bray, S., Reiss, A.L.: Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage* **49**(4), 3039–3046 (2010)
30. Scholkmann, F., Spichtig, S., Muehlemann, T., Wolf, M.: How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiol. Meas.* **31**(5), 649–662 (2010)
31. Izzetoglu, M., Devaraj, A., Bunce, S., Onaral, B.: Motion artifact cancellation in NIR spectroscopy using wiener filtering. *IEEE Trans. Biomed. Eng.* **52**(5), 934–938 (2005)
32. Izzetoglu, M., Chitrapu, P., Bunce, S., Onaral, B.: Motion artifact cancellation in NIR spectroscopy using discrete Kalman filtering. *Biomed. Eng. Online* **9**(1), 16 (2010)
33. Sato, S., Tanaka, N., Uchida, M., Hirabayashi, Y., Kanai, M., Ashida, T., Konishi, I., Maki, A.: Wavelet analysis for detecting body-movement artifacts in optical topography signals. *NeuroImage* **33**(2), 580–587 (2006)
34. Robertson, F., Douglas, T., Meintjes, E.: Motion artifact removal for functional near infrared spectroscopy: a comparison of methods. *IEEE Trans. Biomed. Eng.* **57**(6), 1377–1387 (2010)
35. Molavi, B., Dumont, G.A.: Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol. Meas.* **33**(2), 259–70 (2012)
36. Blasi, A., Phillips, D., Lloyd-Fox, S., Koh, P.H., Elwell, C.E.: Automatic detection of motion artifacts in infant functional optical topography studies. *Adv. Exp. Med. Biol.* **662**, 279–284 (2010)
37. Virtanen, J., Noponen, T., Kotilahti, K., Virtanen, J., Ilmoniemi, R.J.: Accelerometer-based method for correcting signal baseline changes caused by motion artifacts in medical near-infrared spectroscopy. *J. Biomed. Opt.* **16**(8), 087005 (2011)
38. Zhang, Q., Strangman, G.E., Ganis, G.: Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: how well and when does it work? *NeuroImage* **45**(3), 788–794 (2009)
39. Stothers, L., Guevara, R., Macnab, A.J.: Classification of male lower urinary tract symptoms using mathematical modeling and a regression tree algorithm of non-invasive near infrared spectroscopy parameters. *Eur. Urol.* **57**(2), 327–333 (2010)
40. Guevara, R., Stothers, L., Macnab, A.J.: Mathematical modeling methodology for generation of a diagnostic algorithm using near-infrared data. *Spectroscopy* **25**(1), 1–11 (2011)
41. Macnab, A.J., Shadgan, B., Stothers, L.: Monitoring physiologic change in the bladder in health and disease. a new biomedical application of near-infrared spectroscopy. *Biomed. Spectrosc. Imaging* **2**, 289–299 (2013)
42. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: *Classification and Regression Trees*. CRC Press, New York (1998)
43. Hastie, T., Tibshirani, R., Friedman, J.H.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, pp. 305–313. Springer, New York (2001)
44. Loh, W.H.: Classification and regression tree methods. In: Ruggeri, F., Kenett, R.S., Faltin, F.W. (eds.) *Encyclopedia of Statistics in Purity and Reliability*, pp. 323–325. Wiley, London (2008)
45. Starly, B., Fang, Z., Sun, W., Shokoufandeh, A., Regli, W.: Three-dimensional reconstruction for medical-CAD modeling. *Comput.-Aided Des. Appl.* **2**(1–4), 431–438 (2005)
46. Wang, Z., Aarya, I., Gueorguieva, M., Liu, D., Luo, H., Manfredi, L., Wang, L., McLean, D., Coleman, S., Brown, S., Cuschieri, A.: Image-based 3D modeling and validation of radiofrequency interstitial tumor ablation using a tissue-mimicking breast phantom. *Int. J. CARS* **7**(6), 941–948 (2012)

47. Molavi, B., Dumont, G., Shadgan, B., Macnab, A.J.: Attenuation of motion artifact in near infrared spectroscopy signals using a wavelet based method. *Proc. SPIE* **7890**, 78900M (2011). doi:[10.1117/12.875741](https://doi.org/10.1117/12.875741)
48. Molavi, B., Shadgan, B., Macnab, A.J., Dumont, G.: Non-invasive optical monitoring of bladder filling to capacity using a wireless NIRS device. *IEEE Trans. Biomed. Circuits Syst.* (2013). doi:[10.1109/TBCAS2272013](https://doi.org/10.1109/TBCAS2272013)
49. Macnab, A.J., Stothers, L.: Development of a near infrared spectroscopy instrument for applications in urology. *Can. J. Urol.* **15**(5), 4233–4240 (2008)
50. Arogundade, F.A., Barsoum, R.S.: CKD prevention in sub-Saharan Africa: a call for governmental, non-governmental and community support. *Am. J. Kidney Dis.* **51**(3), 515–523 (2008)

Chapter 21

Designing and Manufacturing Quartz Crystal Oscillators

Benfano Soewito

Abstract In this decade, the number of wireless devices grow significantly, caused the frequency stability is very important. Almost all wireless devices use the quartz crystal oscillator to generate stable frequency. To get good quality of quartz crystal oscillators, then it needs to be taken carefully from designing up to how to manufacture them. In this paper, we present the characteristic of quartz crystal oscillator and the factors that affect the frequency stability of quartz crystal oscillator. Knowing these factors to design crystal oscillator is essential. The characteristic of crystal oscillator will depend on the material purity, cutting angle, dimension, surface roughness and edges, thickness, material of electrode, and cleanliness. Crystal oscillator manufacturing process is very long, starting from row material angle marking, cutting, shaping, lapping, electrode evaporation, mounting, frequency adjusting, and aging. All these process will discuss in this paper, including the testing.

21.1 Introduction

In this decade, the numbers of devices that use wireless technologies grow significantly. Starting from the most simple such as TV's remote to highly complex equipment such as mobile phones or satellite. All these devices aim to transfer information or signals using radio waves. These radio waves have a frequency between 3 kHz and 300 GHz. Each range of frequency is used for specific purposes. For instance, 470–860 MHz for TV transmission, 1–2 GHz for GPS and mobile GSM phone, 2 GHz for wireless LAN and microwave ovens and so on. Because so many wireless devices, it is clear that the frequency stability is very important; if not there will be interference between the wireless devices. Almost all wireless devices use the quartz crystal oscillator to generate stable frequency. To get good quality of quartz crystal oscillators, then it needs to be taken carefully from designing up to how to manufacture them.

B. Soewito (✉)

Bina Nusantara University, Kebon Jeruk Raya No. 27, Jakarta 11530, Indonesia
e-mail: benfano@gmail.com

Quartz crystal has a unique characteristic that is the piezoelectric effect. The piezoelectric effect is the effect of material when we applied an action to it. In piezoelectric, the material will generate the electricity when we applied mechanical stress or pressure on it. Vice versa, the materials will response when we applied electricity to the material. This effect can be useful for many applications such as sensor, frequency generator, energy alternative, etc. In this paper we present the quartz crystal for generating frequency. In this paper we present designing and manufacturing quartz crystal oscillator starting from growing quartz crystal to final testing of quartz crystal oscillator.

21.2 Background

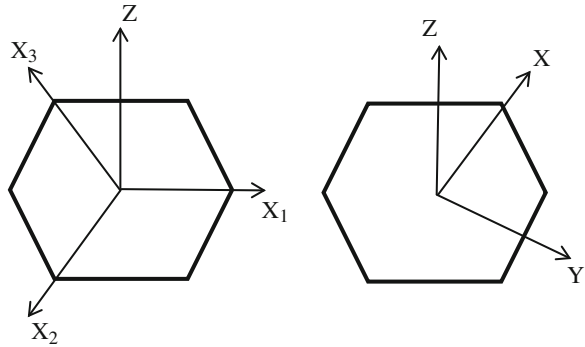
In this section presents the background related to the quartz crystal oscillator. First we will introduce the piezoelectric material and theory then explaining the application of the quartz crystal in many areas.

21.2.1 Piezoelectric

The piezoelectric effect found in some material in which the material will generate the electricity charge when the mechanical stress applied to the material. Vice versa, if the electricity field applied to the material, the material response by mechanical strain. The effect of piezoelectric was introduces at the first time by Coulomb in the year 1815 [1, 2]. He had a hypothesis and theory that some material can generated electricity when pressure applied on that material. Five year after Coulomb introduced his theory, and then Hauy and Becquerel tried to prove the Coulomb's theory. His experiment support the Coulomb's theory, he concluded that the electricity charge could be generated by stretching rubber. In 1880, the brothers Pierre Curie and Jacques Curie proved the piezoelectric effect in crystalline minerals such as tourmaline, quartz, topaz, cane sugar, and Rochelle salt. They made experiment to apply mechanical stress to those crystalline minerals and measured the voltage from electrical charge that resulted by those crystalline minerals.

Since Curie brothers found the effect of piezoelectric then many research had been done in this area including the theory, model and experiment. In the beginning of the 19th century, Langevin started to make transducer from piezoelectric material. Until the frequency stability versus temperature change of quartz crystal oscillator was found by Born in 1933. Born found that the quartz crystal oscillator that made by 35° to the axis Z of quartz crystal had the best frequency stability to the temperature change. He gave the name AT cut for angle cutting 35° . He also introduces the BT cut that had cutting angle of 49° . After this finding, all of the precision frequency control was made by quartz crystal oscillator.

Fig. 21.1 On left is Bravais—Miller coordinate axis system and on the right is Orthogonal coordinate axis system



There are four type of piezoelectric material can be identified since Curie brothers discovered the effect of piezoelectric [3]. These types of piezoelectric materials are single crystal quartz, single crystal Rochelle salt, barium titanate ceramics, and lead zirconate titanate (PZT) ceramics. In this paper we focused on the single crystal quartz. The quartz crystal made by silicon and oxygen or silicon oxide (SiO_2). Its characteristic form is a result of the unit cells by which the crystal grows [4]. These unit cells are identical and consist of atoms arranged in a repetitive geometric pattern. In quartz crystal there are four axes of symmetric because of its geometrical shape. The first axis is optic axis. The other three axes lie in a plane perpendicular to the optic axis and separated by an angle of 120° . These axes are known as electric axes because along the direction in which they lie the maximum piezoelectric effect is observed [5]. The properties of quartz crystal depend on direction of these axes. In order to describe the characteristic of quartz crystal easier, then different axial systems need to be familiarized. In Fig. 21.1, showed two coordinate axis system to explain different properties of quartz crystal. The Bravais—Miller axis system was used to explain atomic system in quartz crystal. The Orthogonal axis was used to explain the piezoelectric and mechanical properties of quartz crystal.

Basically the application of quartz crystal can be divided into several groups: Sensor (microphones, pressure sensor, force sensor, strain gauge), Actuators (loud-speaker, piezoelectric motors, valves), Generator (energy harvesting, cigarette lighter), Transducer (ultrasonic sonar devices), and Frequency standard (oscillator).

21.2.2 Quartz Crystal Oscillator: History and Market

The research of quartz crystal oscillator initiated by A.M. Nicholson of the Bell Telephone Laboratories and W.G. Cady at Wesleyan University in 1917 [6]. Two year later, Cady used a quartz piezoid to control the frequency of an oscillator and in a series of papers during the next three years he described the use of quartz bars and plates as frequency standards and wave filters. It is generally accepted that Cady was the first to use a quartz piezoid to control the frequency of an oscillator circuit [6].

According to Virgil [6], Bell Telephone Laboratories established a quartz laboratory in 1923. And then one of individuals who recognized the potential of the quartz crystal unit was August E. Miller. Miller went to quartz crystal business, making quartz crystal oscillator for amateur radio operators. Quartz crystal was blank very expensive at that time. General Radio Company offered quartz crystal blank for sale at the price of 35–50. This was very large amount of money in 1925 [7]. Finally, in 1926, the AT&T radio station WEAJ in New York City became the first radio station in the United States to control its frequency with a quartz crystal unit [8].

Efforts to develop a practical unit having a low frequency temperature coefficient were successful in 1934 when the AT and BT cuts were discovered independently by Koga from Japan, Bechmann and Straubel from Germany, and Lack, Willard, and Fair from United States [8].

According to Ji Wang [8], in 1943 there was 130 manufactures were engaged in the production of crystal units. They were scattered over 20 states from Oregon to Florida. Most of them were in the same area. There were twenty three factories that produced crystal units in Chicago, twenty in New York area, fifteen in Carlisle area, and fourteen in Kansas City area.

Quartz crystal unit market will be about more than 10 billion USD with growth rate of 7.5 % annually. These because of more devices that work based on the wireless technique. In 2011, there are about 450 companies and suppliers globally that had business in quartz crystal units. 30 % of crystal unit were installed in the telecommunication devices while the remaining 70 % were installed in consumer electronics products. Japan is supplying the high-end products for telecommunications, while china is producing crystal units for the consumer markets with a global market share about 25 % [8].

21.2.3 Quartz Crystal Synthetic

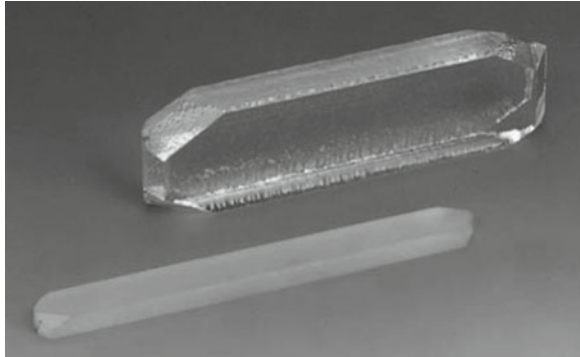
The quartz crystal is formed from silicon and oxygen, become SiO_2 (silicon dioxide) as shown in Fig. 21.2 [10]. At the beginning of producing crystal oscillator in 1920s, the manufacturer used natural quartz crystal. Many disadvantages used natural quartz crystal to make quartz crystal units. The number of natural quartz crystal is very limited and the purity is very low. The purity will have an impact to the impedance of quartz crystal unit and will drain the power resources. Besides that, the determination of axis is very difficult. All these make the price of crystal oscillator unit become very expensive in 1920s, \$30–\$50 for a unit.

The research to make synthetic quartz crystal was started in the mid nineteenth century. The scientist attempted to create quartz crystal in laboratory with create the conditions in which the quartz crystal growth in nature. In 1845, the first person who tried to synthesize quartz crystal was a geologist, Karl Emil. He created microscopic hexagonal quartz crystal in a pressure cooker for eight days. However, the quality and size of the crystals that were produced by these early efforts were poor [9]. Since 1930, all the telecommunications industry used the quartz crystal oscillator

Fig. 21.2 A cluster of natural quartz crystal. Adapted from the evolution of time measurement, part 2: quartz clocks by Michael A. Lombardi. In IEEE instrumentation and measurement magazine, October 2011



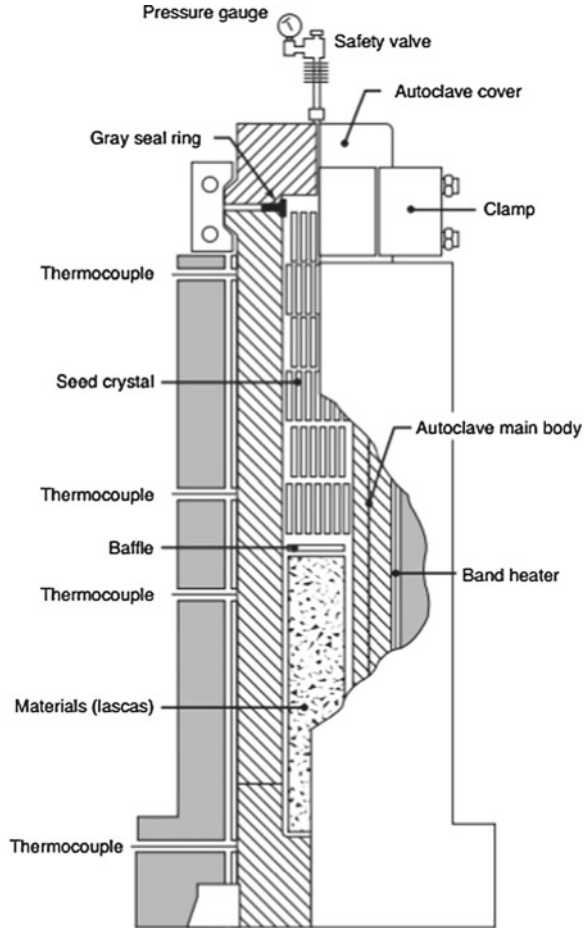
Fig. 21.3 Synthetic quartz crystal. Adapted from synthetic quartz crystal by Nihon Dempa Kogyo Co., LTD



to control the frequency but the supply of quartz crystal is very limited. After the war many researcher and scientist attempted to make the synthetic quartz crystal. Finally in 1950s, synthetic quartz crystal were being produced and sold commercially (Fig. 21.3).

Synthetic quartz crystal as shown in Fig. 21.3 is manufactured in a vertical-type autoclave (high-temperature and high-pressure oven) using the hydrothermal synthesis method. An autoclave is partitioned by a baffle into two compartments: an upper and lower [11]. Seed crystals are placed in the upper compartment (growth zone) and materials (lascas) in the lower (dissolution zone) as shown in Fig. 21.4. A dilute alkaline solution is then poured into the remaining 60–80 % of free space and, after being covered the autoclave is heated with a heater. When the temperatures of the upper and lower compartments of the autoclave reach between 300 to 320 °C and 380 to 400 °C, respectively, the alkaline solution expands and is compressed; the pressure inside reaches 130–145 MPa.

Fig. 21.4 One schematic depiction of synthetic quartz crystal growth. Adapted from synthetic quartz crystal by Nihon Dempa Kogyo Co., LTD



According [11], under these high temperatures and pressures the materials in the lower compartment of the autoclave dissolve in the alkaline solution to become the SiO_2 saturated solution. This saturated solution rises due to the convection caused by the temperature difference between the upper and lower compartments of the autoclave. When the solution reaches the upper compartment of the autoclave, it becomes supersaturated because of the lower temperature of the compartment, and according to the degree of the temperature difference SiO_2 is crystallized on the seed crystal [11]. The solution then returns to the lower compartment of the autoclave and dissolves the materials, thereby becoming the SiO_2 saturated solution, and due to convection it raises and the cycle repeats. The repetition of this process leads to the successive growth of synthetic quartz crystals [11].

Manufacture crystal oscillator unit from the synthetic quartz crystal is easier than natural quartz crystal. Synthetic quartz crystal has high purity or almost 100% pure.

The axes are already determined in the growth process and it is very helpful in cutting process. We can easy to make AT, BT cut or others cutting angle. Synthetic quartz crystal available globally, we can order as much as we need them with variety size.

21.2.4 Thickness Vibration

There are several vibration modes in quartz crystal oscillator as shown in Fig. 21.5. Each angle cut will has its own vibration mode. For example, quartz crystal oscillator with AT and BT cut have thickness shear vibration mode, and CT, DT, and SL cut have face shear vibration mode.

AT cut with thickness shear vibration mode can be vibrate from 0.5 to 300 MHz. The orientation cut angle of AT cut between $35^{\circ}15'$ and $35^{\circ}18'$ and BT cut is $-49^{\circ}08'$ against Z axis. CT, DT, and SL cut with face shear vibration mode can generate frequency 75–900 kHz. The orientation cut angle of CT, DT, and SL cut are 38° , -52° , and -57° consecutively.

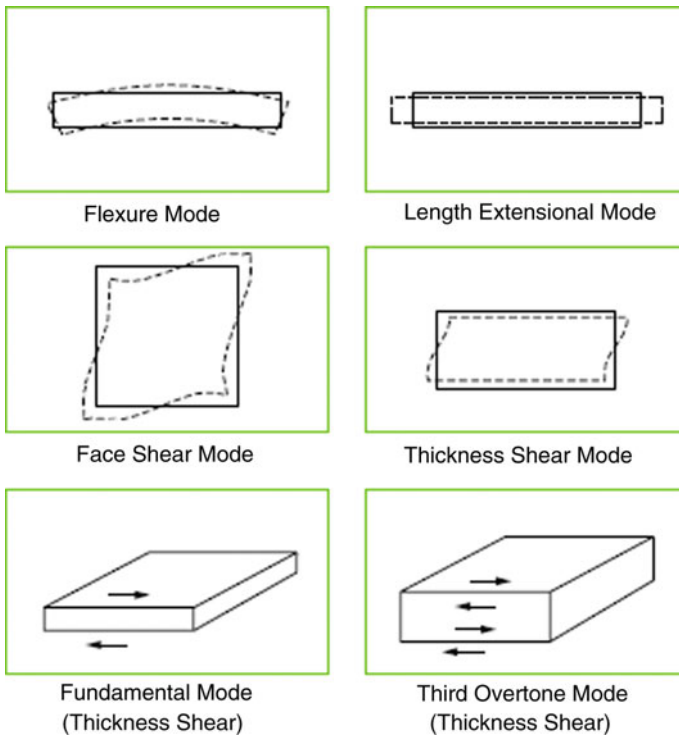


Fig. 21.5 Quartz crystal vibration mode. Adapted from quartz crystal basic theory by token passive components

21.3 Designing Quartz Crystal Oscillator

In designing quartz crystal oscillator, the first step is we have to know several electronic parameters and the environmental condition where the oscillator will be placed. The electronic parameters that we have to know are frequency, maximum resistance, load capacitance, and frequency tolerance. The environmental condition that we have to know is the temperature range where the crystal oscillator will be operated.

Based on those information we can decide the angle cut of quartz crystal follow the graphic shown in Fig. 21.6.

The most popular quartz cut is AT cut, because the most stable frequency to temperature changing is AT cut. The frequency is depending on the thickness of quartz blank. We can follow the Eq. 21.1 to determine the thickness of quartz blank with AT cut. Different cutting has different coefficient cutting.

$$f [kHz] = \frac{1,670}{thickness [mm]} \tag{21.1}$$

For example, to produce quartz crystal oscillator 10 MHz or the thickness 0.167 mm, we must cut the synthetic quartz crystal about 0.24 mm. After cutting process, they will be lapped to the frequency closed to 10 MHz or 0.167 mm.

Next step is to determine the diameter of electrode. The electrode diameter depend on the load capacitance that the required by user. The load capacitance is an external capacitor in order to adjust the frequency to the specific frequency. It can put in parallel or series to the crystal unit (Fig. 21.7).

Fig. 21.6 Frequency versus temperature characteristic of various quartz cutting angle. Adapted from technical terminology by TXC Corporation

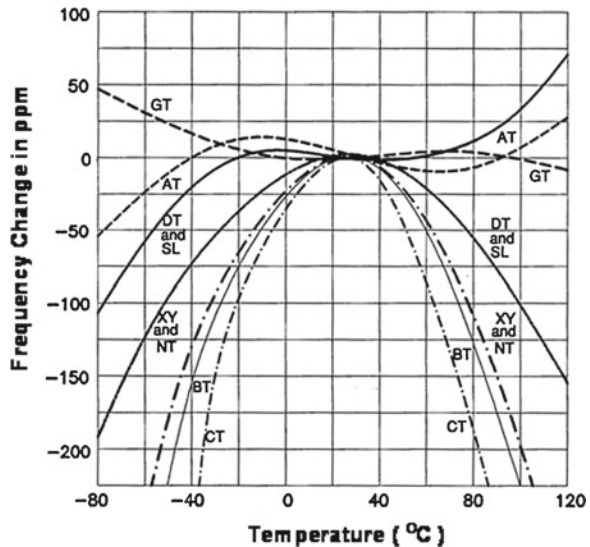


Fig. 21.7 Inside view of crystal unit

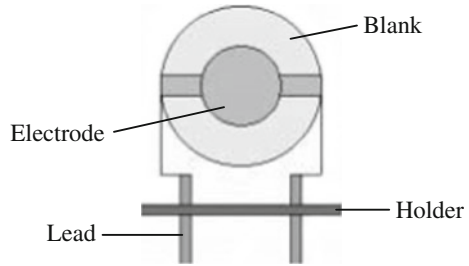
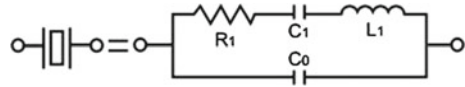


Fig. 21.8 Equivalent circuit of crystal unit



Crystal unit has equivalent circuit as shown in Fig. 21.8. The C_0 is a shunt capacitance represents the capacitance of the crystal electrode including capacitance of the crystal holder and leads. C_1 is motional capacitance represents mechanical elasticity. L_1 is motional inductance represents mechanical inertia. R_1 is motional resistance represents mechanical losses.

The quality factor Q of a crystal unit as shown in Eq. 21.2 is the quality factor of the motional arm at resonance. The maximum stability that can be attained by the crystal unit is directly related to Q , the higher Q the smaller the bandwidth.

$$Q = \frac{2\pi \cdot fr \cdot L_1}{R_1} \tag{21.2}$$

21.4 Manufacturing Quartz Crystal Oscillator

The raw material for manufacturing quartz crystal unit is synthetic quartz crystal. Quartz crystal unit manufacturing process is very long process. The processes start with lapping process or grinding the surface of synthetic quartz crystal to final test.

21.4.1 Lumbering Synthetic Quartz Crystal

The lumbering synthetic quartz crystal is a process to grind the surface of synthetic quartz crystal. This process use abrasive green silicon carbide powder #500. There are two purposes of this process. Firstly, is to make the size of synthetic quartz crystal to be the same. Secondly, is to be easy to propagate the x-ray in the next process. After process lumbering, the synthetic quartz crystal become lumbered bar.

21.4.2 Lumbered Bar Cutting

A steel board was placed on the x-ray machine, and the lumbered bar will be placed on a steel board using UV silicon glue. The x-ray was used to determine the right position of cutting angle lumbered bar. Depending on the size of steel board, one steel board can accommodate 4–6 lumbered bar per layer and it can be three layers maximum. After the lumbered bar placed in the right position then the steel board and lumbered bar will be placed in the UV lamp conveyor to dry the UV glue. Then the steel board and lumbered board placed on cutting machine. Normally, cutting machine uses blade to cut the lumbered bar. It will take about 10–12h to cut three layer of lumbered bar.

21.4.3 Wafer Lapping

After process cutting the lumbered bar become a piece of quartz crystal with the thickness about 0.20–0.24 mm (use 0.25 mm spacer in cutting process), and size about 10×27 mm. The piece of quartz crystal was called the wafer. After cutting process, the thickness of wafers did not the same. The goal of lapping wafers is to make the wafer thickness be the same or in the range 0.19–0.21 mm.

21.4.4 Dimensioning or Shaping

The next process is the process of dimensioning or shaping. In this process, wafers are shaped into a blank with the size and shape in accordance with the order. In this process wafers were stacked of one another and then glued together to form a wafer stick.

Then the wafer stick is made round or cut according to the size of the order. After this process, the wafers quartz crystal become the blank quartz crystal. It can be round or rectangular as shown in Fig. 21.9.

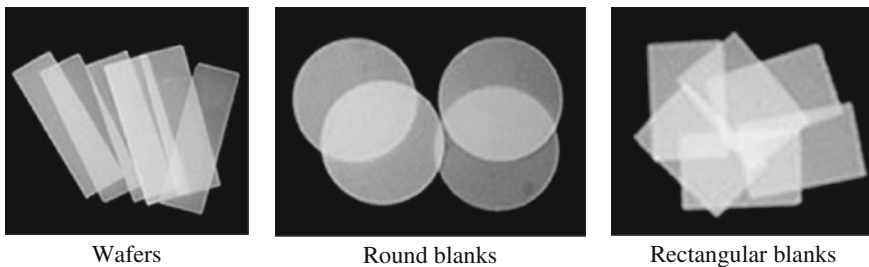


Fig. 21.9 Wafers and blanks

21.4.5 Blank Lapping

In the blank lapping process, the thickness is no longer the target, but a frequency is a target, where the target frequency is in accordance with the order from user. In this case the thinner blank then the higher the frequency. The formula that connects between the frequency and the thickness is shown in Eq. 21.1. Frequency [kHz] = $1,670/\text{thickness [mm]}$. In the blank lapping process used green silicon carbon powder. If we want the crystal oscillator work on the overtone frequency, then after lapping used green silicon carbon, the blank have to be polished.

21.4.6 Etching

It may have particles of green silicon and oil still stick on the surface of blank. The purpose of etching process is to clean the surface of the blank from the particles and smoothing the blank surface. The etching process is done by immersing the blank in fluorite acid or HF acid solution for a few seconds.

21.4.7 Cleaning

In this process, crystal blanks are washed in the ultrasonic washing machine with chemical liquid soap. The purpose of this cleaning process is to clean the surface of crystal blank from the particle and oil lapping that still attached to its surface. Cleanliness of the surface of the crystal blank will determine the quality of the crystal oscillator. In this process, crystal blank is inserted into the basket and then the basket placed into a stainless steel ultrasonic tub that was filled with solution of liquid detergent. Crystal blank was washed by ultrasonic machine approximately one to three minutes, depending on the thickness of the crystal blank to be washed and the number of blanks in the basket. When finished, the blank crystal soaked with alcohol and put them back to the ultrasonic machine in the basket with alcohol. After that, immerse the crystal blank with alcohol to absorb water. Once it's dry so no water spots on the surface of the crystal blank. After that, the crystal blank is inserted directly to the oven with a temperature between 80 and 110 °C for 1 h. Once out of the oven then the crystal blank is ready to the next process, which is base plating process. If the blank do not clean, it will increase the motional resistance.

21.4.8 Base Plating

In this process, blank is given electrode on both sides by using silver or gold. When using gold it will be more resistant to corrosion or rust. But generally use silver.

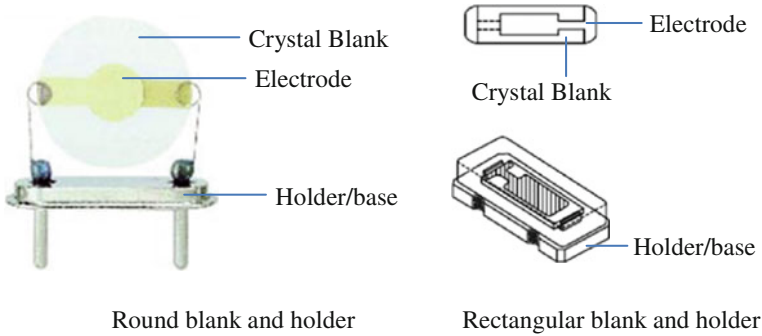


Fig. 21.10 Crystal blank on the holder

Blank will be placed on the mask, and the mask is inserted into the machine base coating. The process of base plating uses an evaporation technique at low pressure or vacuum. Actually, this low pressure just to lower the melting point of silver or gold, beside to pull out the oxygen gas trapped in the base plating machine.

21.4.9 Mounting

Mounting is the process to place the blank on the holder as shown in Fig. 21.10. The holders depend on the type of crystal. Some holder for round blank and some for rectangular blank.

21.4.10 Frequency Adjustment

In this process frequency is adjusted so close to the target by adding silver to the electrode. Crystals are placed on a ring and covered by a mask in accordance with electrode diameter and is inserted into the vacuum machine. If the frequency more than 50 PPM from the target then silver will be sprayed on the electrode until the frequency range to 5 PPM.

21.4.11 Sealing

Sealing is process to cover crystal blank and holder. There are several techniques to seal between cover and holder. The most popular are resistance welding technique. Resistance welding is a process that use electrically to generated heat and force

is applied to surface of material to bonds between holder and cover. The welding process is done in a chamber, fill in by nitrogen gas. The purpose of nitrogen is to pull out the oxygen gas, because the oxygen gas can make the electrode oxidation.

21.4.12 Final Test

The last process is final test. The quartz crystal unit is tested use the automated final test machine. Normally, the parameter which was measured as follow:

1. Frequency range.
2. C_0 (shunt capacitance), represents the capacitance of the crystal electrode including capacitance of the crystal holder and leads.
3. C_1 (motional capacitance), represents mechanical elasticity.
4. L_1 (motional inductance), represents mechanical inertia.
5. R_1 (motional resistance), represents mechanical losses.

21.5 Conclusion

Stability frequency on wireless devices became very important, because more and more the wireless devices used in daily life. From simple equipment such as TV remote, AC remote to complex equipment such as cell phone, access point, etc. These devices will not be able to function properly when the frequency interfering between the devices. Crystal oscillator unit is an electronic component that is responsible for frequency stability. Therefore all steps and processes in designing and manufacturing crystal oscillator must be considered carefully.

Orientation of cutting angle is very critical to the stability of frequency to change of temperature. Quartz crystal unit with the AT cut has the most stable frequency compare to others cutting angle.

The motional resistance R_1 will be higher if the surface of quartz blank is not clean or there are some particles on the surface. The higher motional resistance will need more energy to vibrate. In other word, it will drain the battery more if installed in the wireless devices.

References

1. Ballato, A.: Piezoelectricity: old effects and new applications. IEEE Ultrason. Trans. Ferroelectr. Freq. Control **42**, 916–926 (1995)
2. Ballato, A.: Piezoelectricity: history and new thrusts. IEEE Ultrason. Symp. **1**, 575–583 (1996)
3. Fujishima, S.: The history of ceramic filters. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**(1), 1–7 (2000)

4. Hewlett Packar: Fundamentals of quartz oscillators. Application Note 200-2, Electronic Counters Series, USA (1997)
5. Worthen, C.E.: Piezo-electric quartz plates. *Gen. Radio Exp.* **4**(9) (1930)
6. Bottom, V.E.: A history of the quartz crystal industry in the USA. In: Proceedings of the 35th Annual Frequency Control Symposium, pp. 3–12 (1981)
7. Brown, P.R.J.: The influence of amateur radio on the development of the commercial market for quartz piezoelectric resonators in the United States. In: Proceedings of the IEEE International Frequency Control Symposium, pp. 58–65 (1996)
8. Yong, Y.-K., Wang, J.: Theory and analysis of quartz crystal resonators. In: A Tutorial at the Joint Conference of the IEEE International Frequency Control Symposium and the European Frequency & Time Forum (2011)
9. Moriya, K., Ogawa, T.: Growth history of a synthetic quartz crystal. *J. Cryst. Growth* **58**(1), 115–121 (1982)
10. Lombardi, M.A.: The evolution of time measurement, Part 2: quartz clocks. *IEEE Instrum. Meas. Mag.* 41–48 (2011)
11. Nihon Dempa Kogyo Co. LTD: Synthetic quartz crystal. http://www.ndk.com/catalog/AN-SQC_GG_e.pdf (2014). Accessed April 2014

Chapter 22

Evaluation of Cache Coherence Mechanisms for Multicore Processors

Malik Al-Manasia and Zenon Chaczko

Abstract Multiple core designs have become commonplace in the processor marketplace, and are therefore a major focus in modern computer architecture research. Thus, for both product development and research, multiple core processor performance evaluation is a mandatory step in marketplace. Multicore computing has presented many challenges for system designers; one of which is data consistency between a shared cache or memory and the local caches of the chip. This is also known as cache coherency. The cache coherence mechanisms are a key component in the direction of accomplishing the goal of continuing exponential performance growth through widespread thread-level parallelism. In the scope of this research, we have studied the available efficient methods and protocols used to achieve cache coherence in multicore architectures. These protocols were further modeled and evaluated utilizing Simics simulator for multicore architectures. We also explored the weaknesses and strengths of different protocols and discussed the way of improving them.

22.1 Introduction

Enhancement of microprocessors is guided greatly by Moores Law, which forecasts that the number of transistors per silicon area is doubled every eighteen months [38]. Computer architects are embarking on a fundamental shift in how the transistor bounty is used to increase performance, while Moores Law is expected to continue at least into the next decade [36].

There is a great correlation between power and frequency. When the frequency is enhanced, the power will also be enhanced and after that the temperatures will also increase [41]. Multicore processors take advantage of this relationship by combining

M. Al-Manasia (✉) · Z. Chaczko
Faculty of Engineering, University of Technology Sydney, Broadway,
NSW 2007, Australia
e-mail: malik.a.al-Manasia@student.uts.edu.au

Z. Chaczko
e-mail: zenon@eng.uts.edu.au

multiple cores. Each core is able to run at a lower frequency, by splitting up the power provided to a single core normally between all cores [18]. The performance will enhance whereas the power and temperatures are still under control.

Numerous new applications are becoming multithreaded and computer architecture is turning its focus towards parallelism. This is because it's very hard to enhance the performance of single core processors by increasing the clock frequencies, let alone the difficulties probable such as heating or speed of light if the frequency exceeds certain ranges, the design and verification needs a large team as well.

Multicore design is faced with many challenges like any new design. These challenges require identification and understanding. One of the main challenges facing Chip Multiprocessors (CMPs) is the competition for shared resources, this challenge forms a restriction bottleneck [11, 23, 24]. Some of the shared resources are: main memory bandwidth and capacity, cache bandwidth and capacity, memory subsystem interconnection bandwidth and system power. Memory system scalability is the big issue the multicore future will face and challenge.

The first on-chip multiprocessor for the computing market of the general purpose was presented in 2000 by IBM. Followed in 2005 by AMD that presented the two processor version for the server market, also in the end of 2007 and 2008 kickoff quad-core and triple-core processors were introduced, whereas an 8-core chip for computer-farm application was produced by Sun in 2006. Tileria however, on the 20th of August 2007 gave off its 64-core processor. Intel also released its two-processor versions in 2005 for the server market; its quad core processor was presented on Dec 13, 2006. Within the coming years Intel is expected to release its 80-core processor prototype each running at 3.16GHz.

Cache memory is defined as a specific memory subsystem whereas data of persistent usage is stored for fast access. Caching is used to increase the speed of large amount of slow, affordable memory by using a small amount of expensive memory. Regarding multicore processors, cache coherency stands for the credibility of data stored in each cores cache. Multicore processors may have distributed and shared caches on the chip, so we should account for coherence protocols to assure that when a core reads from memory, it reads the current piece of data and not a value that has been updated by a different core.

The cache coherence problem is illustrated in Fig. 22.1. Figure 22.1 illustrates a shared L2 cache by four cores with private caches via a bus. The cores access location X sequentially. At the beginning core 1, brings a copy to its cache by reading X from L2 cache. After that, core 4 brings a copy to its cache by reading X from L2 cache. Thereafter core 4 changes X's location value from 8 to 5. With a write-through cache, causing the L2 cache location to be updated, but in (action 4) when core 1 read location X again, it will read the old value 8 from its own cache rather than reading its correct value 5 from L2 cache.

The writeback caches complicate the situation even further. Core 4's write would not update L2 cache right away; instead it would barely set the dirty (or modified) bit concerned with the cache block holding location X. Contents of cache block would be written back to L2 cache solely when this cache block is subsequently replaced from the cache of core 4. The reading of the old value is not inclusive to core 1;

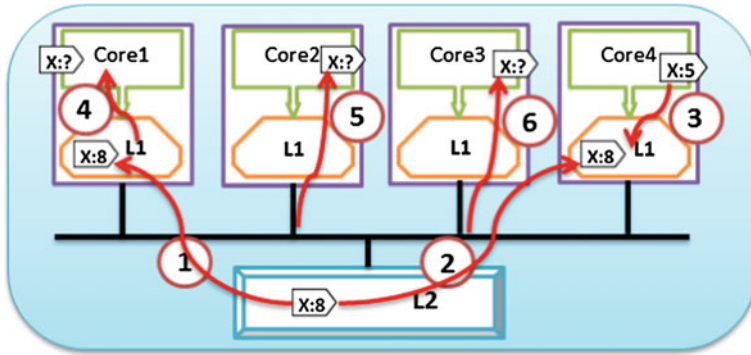


Fig. 22.1 Cache coherence problem

furthermore core 2 and core 3 will miss in their caches when reading location X (actions 5 and 6) by reading the old value 8 instead of 5 from L2 cache. Last but not least, if more than one core write different values in their write-back caches to location X, The end value that will reach the L2 cache will not be related to the sequence in which the writes to X occurred instead it will be determined by the sequence in which the cache blocks that contain X are replaced.

Coherence between the caches has to be enforced in order for correct execution. This process is affected by two main factors: Performance and implementation cost.

In general in hardware-based, there are two methods for cache coherence, a snooping protocol and a directory-based protocol. The Snoopy cache-coherence methods require sending information to all of cache controllers. However, if the number of cores increased the cache messages will increase, also then the required bus which connect the caches and all messages pass through it bandwidth will be bigger than available one, then total saturation of bus bandwidth occurs. These techniques can be used in small-scale systems due to this limitation. However, the Directory-based protocol, scales to larger numbers of processors or cores than snoopy-based coherence protocol, since it enables multiple coherence actions to take place at the same time.

There are three critical attributes that have an impact on the performance of any cache coherence protocol, thus being:

Low-latency Cache-to-Cache Misses: A cache-to-cache miss is a miss frequently caused by accessing shared data that requires another processors cache to provide the data. To decrease the latency of cache-to-cache misses, a coherence protocol should ideally support direct cache-to-cache misses [33]. To efficiently support the frequent communication and synchronization in these workloads, systems are required to optimize the latency of cache-to-cache misses [6].

Bandwidth Efficiency: A cache coherence protocol should conserve bandwidth to decrease the cost and avoid interconnect contention (since contention reduces performance).

Scalability Challenges: Hardware resources must scale efficiently, as the future is multithreaded. Memory hierarchy is the main hardware scalability dilemma. CMPs depend on large, multi-level cache hierarchies to diminish the high cost, limited

bandwidth and high latency of main memory. Caches generally use nearly the half of chip area and a substantial fraction of system energy. The main problems limit cache hierarchies scalability are cache associativity, partitioning and coherency. Caches currently spend significant energy and latency to implement associative lookups, making them inefficient and have a considerable impact on system performance. The previously proposed partitioning techniques suffer from two main weaknesses: they are not scalable and limited to a small number of coarse-grain partitions, and partitioning often reduces performance. Also, traditional coherence schemes do not scale beyond a few tens of cores.

The objective of this paper is to explain the relative behavior of different coherence protocols (traditional protocols and Token Coherence based protocols). We aim not to (1) Generate ultimate execution times or throughput rates for our simulated systems or (2) Evaluating such protocols on all of the future's system configurations. We use an approximation of a chip multiprocessor system in order to accurately obtain relative comparisons and evaluations instead. Full system simulation and modeling of the first-order timing effects for approximating an aggressive multicore system operating commercial loads are the means to reach such an aim. Our goal is to get the first-order effects, although similar to most architectural simulations—Capturing all system's aspects in precise detail is not what we try to do.

22.2 Background and Related Work

In this section, we describe the traditional protocols (Snooping based protocols and Directory based protocols) in addition to the Token coherence protocol: how they are working, what the advantages and disadvantages of each of them and how the improvement of these protocols can be done.

22.2.1 Snooping Protocols

In Snooping protocols the processor cores snoop every bus transaction and respond with appropriate state changes for the corresponding cache lines depending on two elements: the cache line status and bus transaction type. Two primary policies classify the snoop-based coherence protocols: the invalidation-based protocols (e.g., the write-once [21], the Synapse [16], the Berkeley [25] and the Illinois [13]) and the update-based protocols (e.g., the Firefly [45] and the Dragon [37]). Hence invalidation-based coherence has been favored over update-based coherence protocols in most up-to-date systems (e.g., [9, 12, 13, 22, 29, 39, 44]), this paper considers only invalidation-based cache coherence protocols.

The main present advantage of snoop-based multiprocessors is the low average miss latency, especially for cache-to-cache misses. The responder knows fast that it has to send a response because a request is sent directly to all the other processors and memory modules in the system. Low cache-to-cache miss latency is of great

importance for workloads with considerable amounts of data sharing. Replying with data from processor caches when possible can reduce the average miss latency if cache-to-cache misses have lower latency than fetching data from memory (i.e., a memory-to-cache miss). Low-latency memory access is a result of the tightly-coupled nature of these systems [33].

Formerly, snooping has had two additional advantages. First, shared-wire buses used to be cost-effective interconnects for numerous systems and bus-based coherence offered a complexity-effective approach to applying cache coherence. Second, bus-based snooping protocols were comparatively simple. This advantage that was of great importance in the past is now much less importance; New snooping protocols that use virtual buses are often as complex or more complex than alternative approaches to coherence.

The first primary disadvantage of snooping is that even though system designers have evolved beyond shared-wire buses snooping designers are still bound to choosing interconnects that can provide virtual-bus behavior (i.e., a total order of requests) when they choose interconnect. These virtual-bus interconnects could be more expensive (e.g., by requiring switch chips), may obtain lower bandwidth (e.g., due to a bottleneck at the root), or might acquire higher latency (since all requests need to reach the root). On the contrary, an unordered interconnect (such as a directly connected grid or torus of processors) might have more attractive latency, bandwidth and cost attributes [33].

The second primary disadvantage is that snooping protocols are still naturally broadcast-based protocols; i.e., protocols whose bandwidth requirements increase with the number of processors. This broadcast requirement limits system scalability even after eliminating the bottleneck of a shared-wire bus or virtual bus. To control this limitation, recent proposals [8, 34, 42] aim at reducing the bandwidth requirements of snooping by using destination-set prediction (also known as predictive multicast) instead of broadcasting all requests. These proposals suffer from snooping's other disadvantage: They rely on a totally-ordered interconnect, Even though they reduce request traffic.

Some of the best techniques that can be used to improve the Snoopy Cache Coherency protocols:

1. Three wired OR signals: In this technique, when any other cache has a copy of block besides the requester the first signal is asserted, and when any cache has exclusive copy of block the second signal given. The third signal is asserted when all snoop actions are finished on the bus [18]. When the third signal is asserted, the other two signals are safely examined by the requesting L1 and the L2. Performance can be improved by implementing these signals using low-latency L-Wires since all of them are on the critical path.
2. Voting wires is another technique used to enhance snoopy based protocol with low latencies. Generally cache to cache transfers occur from the data in the modified state, whereas there is a single supplier [17]. Although, a block can be retrieved from other cache rather than memory in MESI protocol. Multiple caches share copy voting mechanism is generally used to provide data therefore voting mechanism works with low latencies and enhances processor performance.

22.2.1.1 Directory-Based Protocols

Memory is distributed among different processors in directory based protocols and directory is maintained for each such memory. Currently, several Chip Multiprocessors, as Piranha [7] also use directory protocols in order to maintain cache coherency. L1 cache misses are sent to L2 caches and a directory which store the status of block is maintained across each L2 caches. The request goes to home node where the original data is stored to check whether it has. When request comes from requester node from another cache, if it is not available the request goes to remote node by home node and first fetches data from remote node and sends it later on to requester node. Also write-invalidate-direct based protocol is employed in one of the most common chip multiprocessors technology we are using, which is core2duo.

Directory protocols target the avoidance of the scalability and interconnection limitations of snooping protocols. Directory protocols predate snooping protocols for a fact, with Censier and Feautrier [10] and Tang [43] performing early work on directory protocols in the late 1970s. Systems that use these protocols also known as distributed shared memory (DSM) or cache-coherent non-uniform memory access (CC-NUMA) systems are preferred when scalability (in the number of processors or cores) is a first-order design constraint. These protocols often sacrifice fast cache-to-cache misses in exchange for this scalability, even though these protocols are significantly more scalable than snooping protocols, Examples of systems that use directory protocols include Stanfords DASH [27, 28] and FLASH [26], MITs Alewife [2], SGIs Origin [29], the AlphaServer GS320 [20] and GS1280 [15], Sequents NUMA-Q [14], Crays X1 [1], and Piranha [7].

Directory protocols better scalability than snooping protocols and avoidance of snoopings virtual bus interconnect are the two primary advantages of directory protocols. The most discussed and studied advantage is perhaps the significantly improved scalability of directory protocols. By only contacting those processors that might have copies of a cache block (or a small number of additional processors when using an approximate directory implementation), the traffic in the system grows linearly with the number of processors. In contrast, the endpoint traffic of broadcasts used in snooping protocols grows quadratically [33]. Combined with a scalable interconnect (one whose bandwidth grows linearly with the number of processors), Using directory protocol the system is permitted to scale to hundreds or thousands of processors.

Two scalability dilemmas are encountered when using large system sizes:

First, the amount of directory state required becomes great concern.

Second, interconnect of reasonable is not truly scalable.

A deep study of these two problems have been applied extensively, and actual systems supporting hundreds of processors exist (e.g., the SGI Origin 2000 [29]).

The ability to exploit arbitrary point-to-point interconnects is the second and maybe the more important advantage of directory protocols. The point-to-point interconnects are generally have high-bandwidth and low-latency [33].

Directory protocols are of two main disadvantages. First, the extra interconnect traversal and directory access is on the critical path of cache-to-cache misses. Hence the memory lookup is normally performed simultaneously with the directory

lookup memory-to-cache, misses do not incur a penalty. Directory lookup latency is similar to that of main memory DRAM in numerous systems, and thus locating this lookup on the critical path of cache-to-cache misses increases cache-to-cache miss latency considerably. Although the directory latency can be decreased by using fast SRAM in order to hold or cache directory information, the extra latency presented by the additional interconnect traversal is harder to mitigate. A combination of these two latencies enhances cache-to-cache miss latency significantly. With the prevalence of cache-to-cache misses in many important commercial workloads, these higher-latency cache-to-cache misses might have a dramatical impact on system performance.

The storage and manipulation of directory state could be considered the second disadvantage of directory protocols. This disadvantage was present on earlier systems, ones that used dedicated directory storage (SRAM or DRAM) thus adding to the total system cost. On the other hand, to save directory state while eliminating additional storage capacity overhead; numerous modern directory protocols have used the main system DRAM and reinterpretation of bits used for error correction codes (ECC) (e.g., the S3mp [40], Alpha 21364 [39], UltraSparc III [22], and Piranha [7, 20]) by increasing the number of memory reads and writes [19] storing these bits in main memory enhances the memory traffic [33].

Some of the best techniques that can be used to improve the Directory based Cache Coherency protocols:

1. Exclusive Read Request for a block in a shared state.

The L2 cache is a clean copy and upon receiving a request from the L1 cache the L2 cache is going to invalidate every L1 cache in this approach. Before sending the data to the processor, The L1 cache will receive an invalidate acknowledgement from the other L1 caches. Normally the L2 cache requires a hop for the reply and where as it requires 2 hops for an acknowledgment. So the latencies of the Hop and the latencys of reply and acknowledgement messages should be of the same value [4]. In this approach both the acknowledgment and reply messages are sent at the same time via the corresponding low latency L-wires and low power PW-wires. This approach enhances the performance and lowers the consumption of power.

2. Read request for block in exclusive state.

The exclusive owner will receive a read request for the L2 cache sends as the requesting L1 receives a copy of data from the L2 cache. The exclusive owner will send a reply message to the requesting L1 cache pointing out that the data sent by the L2 cache is valid if the exclusive copy is a clean one. If the requesting node requests for data the exclusive owner will send a copy of data to the requesting L1 cache while updating the data in the L2 cache. The requesting cache will not go on until it receives a message from the exclusive owner. Simultaneously the data the L2 cache sends the data through slow PW wires [4].

An acknowledgement message should be sent to the requester from the exclusive owner through low bandwidth L-wires if the owner copy is a clean copy, whereas if the owner copy is a dirty copy then the message should be sent through the B-wires

and the write back to the L2 is done via PW-wires. This approach mainly follows the ways of improving the performance by sending the prioritized data through the L-wires and the least prioritized one through PW-wires.

3. The methods of which the Proximity Aware coherence protocols enhances the performance of a multicore thus by lowering the unnecessary access to the off-chip memory are described below.

22.2.1.2 Token-Based Protocols

The constraint of directory indirection are removed without sacrificing either decoupling of the interconnect from the coherence protocol or decoupling of coherence from consistency thus is done by using the recently-proposed token coherence protocol [31, 33, 34]. Token coherence take on token counting to resolve races without the need to require home node or an ordered interconnect. Token coherence contains even further levels of decoupling by separating the correctness substrate from the systems performance policy. The correctness substrate is decoupled further on invoking safety and avoiding starvation.

Token counting does not ensure that a request is satisfied in the end even though it ensures safety. Thus, the correctness substrate gives persistent requests to prevent starvation. The processor initiates a persistent request when it detects possible starvation. Using a fair arbitration mechanism, the substrate then activates at most one persistent request per block. Each system node remembers all activated persistent requests (for example, in a table at each node) and forwards all tokens for the block those tokens are present at the time being and received in the future to the request initiator. Finally, the initiator performs a memory operation (a load or store instruction) and deactivates its persistent request when it has the necessary tokens.

Token coherence performance policies have been developed [33] to approximate an unordered broadcast-based protocol (inspired by snooping protocols), a bandwidth-efficient performance policy that emulates a directory protocol, and a predictive hybrid protocol that uses destination-set prediction [34].

TokenB protocol focuses on both avoiding indirection latency for cache-to-cache misses (like snooping protocols) and not requiring any interconnect ordering (like directory protocols). One seemingly clear approach is to directly send broadcasts on an unordered interconnect.

22.3 Evaluation Methodology

The objective of evaluation is to explain the relative behavior of different coherence protocols (traditional protocols and Token Coherence based protocols). The Simics full-system multiprocessor simulator (WindRiver) extended with the Wisconsin GEMS simulation environment [35] is used in order to perform the analysis [5].

Table 22.1 Simulation parameters

Private L1 caches	64 KB, 4-way set associative, split D/I, 1 ns latency (2-cycle latency)
L2 cache	Unified, 4 MB, 6 ns latency (12 cycles)
Main memory	4 GB, 80 ns (160 cycles)
Coherence protocol	MOESI protocol
Interconnect link	4 GB/second or unbounded bandwidth 15 ns latency (30 cycles)

Throughout this research, we demonstrate how the overall performance of different protocols by the runtime i.e. measuring the time necessary to complete certain amount of work. The metric instructions-per-cycle has been used by other works instead of runtime in judging performance improvements. However; system timing effects of multiprocessor workloads may alter the number of instructions executed therefore, Instruction Per Cycle (IPC) is not a suitable metric for evaluating the coherence protocols and systems.

The measurement is started at the parallel phase so that we avoid measuring thread forking. Until now, a full system checkpoint (to provide a well-defined starting point) is used to initialize the system state and to simulate the execution until the end of the parallel phase. The number of cycles is recorded and referred to as application time in order to complete the parallel phase.

Endpoint traffic (in messages per miss) and interconnect traffic (in terms of bytes on interconnect links per miss) are other ways besides reporting runtime that measure and report the traffic [4]. The endpoint traffic shows the amount of controller bandwidth needed for handling incoming messages. The amount of link bandwidth used by the messages are indicated in the interconnect traffic as they traverse the interconnection.

We used three multi-threaded commercial workloads from the Wisconsin Commercial Workload Suite [3]: an online transaction processing workload (OLTP), a Java middleware workload (SPECjbb), and a static web serving workload (Apache). The previously mentioned workloads operate on a simulated 16-core SPARC processor that runs Solaris 9. The simulated system has 4 GBs of main memory.

Following are details of how the main components of the system are modeled (Table 22.1).

22.4 Analysis and Evaluations

In order to evaluate the demand system; full-system simulation is used. Using full system simulation allows for evaluating the proposed systems when running realistic scientific applications on top of actual operating systems. As well as capturing the subtle timing effect that can't be captured with trace-based evaluation. We use cycles per transaction to be our only metric of performance. Figure 22.2 shows normalized runtime (smaller is better) that TokenB is faster than Snooping with unlimited bandwidth links by (21–34).

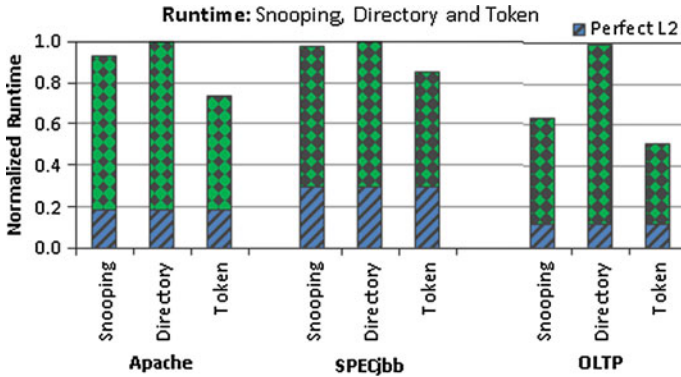


Fig. 22.2 Runtime of Snooping, Directory, and TokenB. The runtime of Snooping, Directory, and TokenB with unbounded link bandwidth

TokenB's endpoint traffic is similar to or less than Snooping while has more interconnect traffic than Snooping. Figure 22.3 illustrates the endpoint traffic (in normalized messages per miss received at each endpoint coherence controller), and Fig. 22.4 exhibits the interconnect traffic (in normalized bytes per miss). When examining only data and non-reissued request traffic, TokenB and Snooping are practically identical. TokenB adds some additional traffic overhead (comes from reissued and persistent requests), but the overhead is small for all three of our workloads. Snooping and TokenB both use extra traffic for writeback control messages, but because of the detailed implementation decisions in Snooping involving writeback acknowledgment messages, Snooping uses more traffic for writebacks than TokenB. Snooping sends a writeback request on the ordered interconnect to both the memory and to itself as a marker message. If it is still the owner of the block, it receives the marker message and sends the data back to the memory. Ignoring this precise implementation overhead leads us to conclude that these protocols generate the same amounts of traffic.

Fig. 22.3 Endpoint Traffic of Snooping, Directory, and TokenB. The endpoint traffic (in normalized messages per miss) of Snooping, Directory, and TokenB

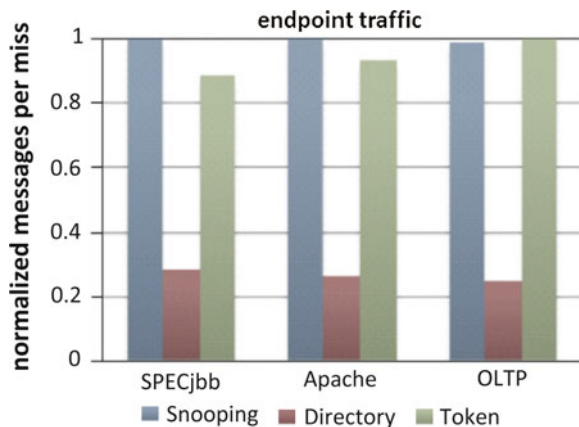
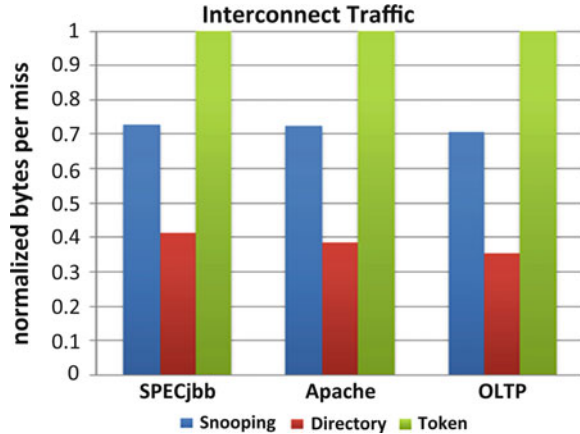


Fig. 22.4 Interconnect Traffic of Snooping, Directory, and TokenB. The interconnect traffic (in normalized bytes per miss) of Snooping, Directory, and TokenB



TokenB generates bigger endpoint traffic and interconnect traffic than Directory. These figures show that TokenB generates more traffic than Directory about three times more endpoint traffic and interconnect traffic). Thus results in, incredibly higher bandwidth coherence controllers are required by TokenB.

TokenB depends on broadcast, which limits its scalability. TokenB is less scalable than Directory, hence DIRECTORY avoids broadcast. However, as the number of processors increases, TokenB endpoint bandwidth improves linearly. The interconnect traffic difference between TokenB and Directory enhances slowly. Thus, TokenB can operate well for almost up to 64 processors if bandwidth is rich (by using high-bandwidth links and high-throughput coherence controllers). On the other hand, TokenB is not a good choice for larger or more bandwidth-limited systems.

22.5 Conclusions and Recommendations for Future Work

The cache coherence mechanisms are a key element aiming at accomplishing the goal of proceeding exponential performance growth through widespread thread-level parallelism. The available efficient methods and protocols were studied in this paper, which were used to achieve cache coherent in multicore architectures. These protocols (Snooping-based protocols, Directory-based protocols and TokenB-based protocols) modeled and evaluated on the Simics/GEMS simulator. The weaknesses and strengths of each protocol were demonstrated and we discussed how the improvement of them can be done.

TokenB is both (1) better than Snooping and (2) faster than Directory when bandwidth is plentiful. TokenB is better than SNOOPING because it uses the same amounts of traffic and can outperform Snooping by exploiting a faster, unordered interconnect. TokenB is of a higher speed than Directory in bandwidth-rich situations by avoiding placing directory lookup latency and a third interconnect traversal on the

critical path of common cache-to-cache misses. On the other hand, TokenB uses a moderate amount of additional interconnect traffic and considerable more endpoint message bandwidth than Directory for small systems. Thus, Directory performs better than TokenB in a bandwidth-constrained situation. Although TokenB is a message-intensive protocol, it is only one of many possible high performance policies.

The choice of coherence protocol is as complicated and subtle today as it has ever been. CMPs will enable even more cost-effective multicore processors by reducing the number of discrete components in the system. For the future work, we recommend further modeling for each of the snooping, Directory, and Token protocols by testing additional benchmarks rather than the three benchmarks that we use in this paper, in the same architecture for more accuracy results.

References

1. Abts, D., Scott, S., Lilja, D.J.: So many states, so little time: verifying memory coherence in the Cray X1. In: IEEE Proceedings of International Parallel and Distributed Processing Symposium, p. 10(2003)
2. Agarwal, A., Bianchini, R., Chaiken, D., Johnson, K.L., Kranz, D., Kubiawicz, J., Lim, B.-H., Mackenzie, K., Yeung, D.: The MIT Alewife machine: architecture and performance. In: IEEE Proceedings of 22nd Annual International Symposium on Computer Architecture, pp. 2–13 (1995)
3. Alameldeen, A.R., Martin, M.M., Mauer, C.J., Moore, K.E., Xu, M., Hill, M.D., Wood, D.A., Sorin, D.J.: Simulating a 2M Commercial Server on a 2K PC. *Computer* **36**(2), 50–57 (2003)
4. Al-Manasia, M., Al-Omari, F., Al-Jarrah, M.: Modeling and evaluation of cache coherence mechanisms for multicore processors. Masters Thesis, Yarmouk University (2011). <http://repository.yu.edu.jo/handle/123456789/1505>
5. Al-Manasia, M., Chaczko, Z.: A survey of computer system architecture simulators, case study: sniper. In: Proceedings of the 2nd Asia-Pacific Conference on Computer-Aided System Engineering, APCASE 2014, 10–12 February 2014, South Kuta, Indonesia, pp. 14–15 (2014). ISBN: 978-0-9924518-0-6
6. Barroso, L.A., Gharachorloo, K., Bugnion, E.: Memory system characterization of commercial workloads. In: ACM SIGARCH Computer Architecture News **26**(3), 3–14 (1998)
7. Barroso, L.A., Gharachorloo, K., McNamara, R., Nowatzky, A., Qadeer, S., Sano, B., Smith, S., Stets, R., Verghese, B.: Piranha: a scalable architecture based on single-chip multiprocessing. In: ACM SIGARCH Computer Architecture News **28**(2), 93–282 (2000)
8. Bilir, E.E., Dickson, R.M., Hu, Y., Plakal, M., Sorin, D.J., Hill, M.D., Wood, D.A.: Multicast snooping: a new coherence method using a multicast address network. In: IEEE Proceedings of the 26th International Symposium on Computer Architecture, pp. 294–304 (1999)
9. Borkenhagen, J.M., Hoover, R.D., Valk, K.M.: EXA cache/scalability controllers. In: IBM Enterprise X-Architecture Technology: Reaching the Summit, pp. 37–50 (2002)
10. Censier, L.M., Feautrier, P.: A new solution to coherence problems in multicache systems. *IEEE Trans. Comput.* **100**(12), 1112–1118 (1978)
11. Chandra, D., Guo, F., Kim, S., Solihin, Y.: Predicting inter-thread cache contention on a chip multi-processor architecture. In: IEEE 11th International Symposium on High-Performance Computer Architecture, pp. 340–351 (2005)
12. Charlesworth, A.: Starfire: extending the SMP envelope. *IEEE Micro* **18**(1), 39–49 (1998)
13. Charlesworth, A.: The sun fireplane interconnect. *IEEE Micro* **22**(1), 36–45 (2002)
14. Clapp, R., Lovett, T.: STiNG: A CC-NUMA computer system for the commercial marketplace. In: IEEE 23rd Annual International Symposium on Computer Architecture (1996)

15. Cvetanovic, Z.: Performance analysis of the alpha 21364-based HP GS1280 multiprocessor. In: IEEE Proceedings of 30th Annual International Symposium on Computer Architecture, pp. 218–228 (2003)
16. Frank, S.J.: Tightly coupled multiprocessor system speeds memory-access times, vol. 1. Electronics, United States (1984)
17. Galles, M., Williams, E.: Performance optimizations, implementation, and verification of the SGI challenge multiprocessor. In: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, vol. 1, pp. 134–143. IEEE (1994)
18. Geer, D.: Chip makers turn to multicore processors. *Computer* **38**(5), 11–13 (2005)
19. Gharachorloo, K., Barroso, L.A., Nowatzky, A.: Efficient ECC-based directory implementations for scalable multiprocessors. In: Proceedings of the 12th Symposium on Computer Architecture and High-Performance Computing (SBAC-PAD 2000) (2000)
20. Gharachorloo, K., Sharma, M., Steely, S., Van Doren, S.: Architecture and design of AlphaServer GS320. In: ACM SIGARCH Computer Architecture News **28**, 13–24. ACM (2000)
21. Goodman, J.R.: Using cache memory to reduce processor-memory traffic. In: 25 Years of the International Symposia on Computer Architecture (selected papers), pp. 255–262. ACM (1998)
22. Horel, T., Lauterbach, G.: UltraSPARC-III: designing third-generation 64-bit performance. *IEEE Micro* **19**(3), 73–85 (1999)
23. Hsu, L.R., Reinhardt, S.K., Iyer, R., Makineni, S.: Communist, utilitarian, and capitalist cache policies on CMPs: caches as a shared resource. In: Proceedings of the 15th International Conference on Parallel Architectures and Compilation Techniques, pp. 13–22. ACM (2006)
24. Iyer, R.: CQoS: a framework for enabling QoS in shared caches of CMP platforms. In: Proceedings of the 18th Annual International Conference on Supercomputing, pp. 257–266. ACM (2004)
25. Katz, R.H., Eggers, S.J., Wood, D.A., Perkins, C., Sheldon, R.G.: Implementing a cache consistency protocol, vol. 13. IEEE Computer Society Press (1985)
26. Kuskin, J., Ofelt, D., Heinrich, M., Heinlein, J., Simoni, R., Gharachorloo, K., Chapin, J., Nakahira, D., Baxter, J., Horowitz, M.: The stanford flash multiprocessor. In: IEEE Proceedings the 21st Annual International Symposium on Computer Architecture, pp. 302–313 (1994)
27. Lenoski, D., Laudon, J., Gharachorloo, K., Gupta, A., Hennessy, J.: The directory-based cache coherence protocol for the DASH multiprocessor, vol. 18. ACM (1990)
28. Lenoski, D., Laudon, J., Gharachorloo, K., Weber, W.-D., Gupta, A., Hennessy, J., Horowitz, M., Lam, M.S.: The stanford dash multiprocessor. *Computer* **25**(3), 63–79 (1992)
29. Loudon, J., Lenoski, D.: The SGI origin: a ccNUMA highly scalable server. In: Proceedings of the 24th International Symposium on Computer Architecture. Silicon Graphics Inc. (1997)
30. Martin, M.M., Harper, P.J., Sorin, D.J., Hill, M.D., Wood, D.A.: Using destination-set prediction to improve the latency/bandwidth tradeoff in shared-memory multiprocessors. In: IEEE Proceedings of the 30th Annual International Symposium on Computer Architecture, pp. 206–217 (2003)
31. Martin, M.M., Hill, M.D., Wood, D.A.: Token coherence: decoupling performance and correctness. In: IEEE Proceedings of the 30th Annual International Symposium on Computer Architecture, pp. 182–193 (2003)
32. Martin, M.M., Sorin, D.J., Ailamaki, A., Alameldeen, A.R., Dickson, R.M., Mauer, C.J., Moore, K.E., Plakal, M., Hill, M.D., Wood, D.A.: Timestamp snooping: an approach for extending SMPs. In: ACM SIGARCH Computer Architecture News. **28**, pp. 25–36. ACM (2000)
33. Martin, M.M.: Token Coherence, University of Wisconsin (2003)
34. Martin, M.M., Hill, M.D., Wood, D.A.: Token coherence: a new framework for shared-memory multiprocessors. *IEEE Micro* **23**(6), 108–116 (2003)
35. Martin, M.M., Sorin, D.J., Beckmann, B.M., Marty, M.R., Xu, M., Alameldeen, A.R., Moore, K.E., Hill, M.D., Wood, D.A.: Multifacet’s general execution-driven multiprocessor simulator (GEMS) toolset. In: ACM SIGARCH Computer Architecture News **33**(4), 92–99 (2005)
36. Marty, M.R.: Cache coherence techniques for multicore processors. PhD thesis, University of Wisconsin (2008)

37. McCreight, E.M.: The dragon computer system. In: *Microarchitecture of VLSI Computers*, pp. 83–101. Springer (1985)
38. Moore, G.E.: Cramming more components onto integrated circuits. Reprinted from *Electronics* **38**(8), 114 (1965) *IEEE Solid-State Circuits Newslett.* **11**(5), 33–35 (2006)
39. Mukherjee, S.S., Bannon, P., Lang, S., Spink, A., Webb, D.: The alpha 21364 network architecture. *IEEE Hot Interconnects* **9**, 113–117 (2001)
40. Nowatzky, A., Aybay, G., Browne, M., Kelly, E., Lee, D., Parkin, M.: The S3. mp scalable shared memory multiprocessor. In: *IEEE Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, vol. 1, pp. 144–153 (1994)
41. Rao, W.: Multi processors, their memory organizations and implementations by Intel and AMD (2009) <http://ece.uic.edu/~wenjing/courses/fa08ECE569/ECE569/w21.pdf>
42. Sorin, D.J., Plakal, M., Condon, A.E., Hill, M.D., Martin, M.M.K., Wood, D.A.: Specifying and verifying a broadcast and a multicast snooping cache coherence protocol. *IEEE Trans. Parallel Distrib. Syst.* **13**(6), 556–578 (2002)
43. Tang, C.: Cache system design in the tightly coupled multiprocessor system. In: *Proceedings of the National Computer Conference and Exposition*, 7–10 June 1976, pp. 749–753. ACM (1976)
44. Tendler, J.M., Dodson, J.S., Fields, J., Le, H., Sinharoy, B.: POWER4 system microarchitecture. *IBM J. Res. Dev.* **46**(1), 5–25 (2002)
45. Thacker, C.P., Stewart, L.C., Satterthwaite Jr, E.H.: Firefly: a multiprocessor workstation. *IEEE Trans. Comput.* **37**(8), 909–920 (1988)
46. WindRiver, Wind River Simics “Full System Simulation”. www.windriver.com/products/simics/

Chapter 23

Aggregator Operation for EVs in Korean Smart Grids Testbed

Wang-Cheol Song and Zhong Ming Huang

Abstract As interest on the Smart Grid technology has recently increase, research and development of the electric vehicle (EV) as well as studies on the renewable energy, distributed generation and energy storage have been actively conducted. The EV is an entity to be able to buy and sell the electricity in the grid and play important roles to interact with various other entities based on the smart grid technology. In this paper, it is described why the aggregator is essential in Korean electricity environment and how it operates in the smart grid to buy and sell electricity on the behalf of EVs and negotiate electricity prices in the market.

Keywords Aggregator · EV · Smart grids · KEPCO · Jeju smart grid testbed

23.1 Introduction

As interest on the Smart Grid technology has recently increased, development of the electric vehicle (EV) as well as study on the smart grid technology have been actively conducted [1]. As the smart grid technology is also got much interested in Korea, Korean government has planned to establish a Smart Grid Testbed in Jeju Island, and the testbed has been constructed from 2009 to 2013 [2]. As the results of the Testbed Project, proliferation of EVs as well as establishment of the infrastructure for EVs have been well progressed. The Jeju province has a plan to make Jeju island as a representative place for the EV with the vision of Carbon free island Jeju by 2030. So, several smart-grid technology based businesses are expected to be tried in Jeju Island in the near future.

W.-C. Song (✉) · Z.M. Huang
Department of Computer Engineering, Jeju National University, Jeju, Korea
e-mail: kingiron@gmail.com

Z.M. Huang
e-mail: huangzhongming0302@gmail.com

As another considerable point, an important law about the electricity business the Electricity Enterprises Act has been revised in Korea in April 2014. It is to permit private enterprises to participate in operation of the response management market that has been operated only by Korea Electric Power Corporation (KEPCO) and Korea Power Exchange (KPX), so that it is now thought business environments for various electric power suppliers to act in the Korean electricity market are well established.

When the existing vehicles are replaced with EVs, it is expected that the more EVs requires the more electric power. Kempton and Tomic introduce the mechanical power of the U.S. light vehicles exceeds the electric power generation of the country by a factor of 24 [3]. However, by the vehicle-to-grid (V2G) concept, the EV charger can bi-directionally deliver the electricity to the grid as well as charge the battery, and the electric networks can utilize electric power from a parked EV [4].

Therefore, in the smart grids framework, the EV can have a role to be able to buy and sell the electricity and to interact with various other entities in the power grids. EVs can carry out transactions of electricity based on the real-time market prices of electricity using the V2G service of the smart grids. However, as the electricity capacity from only an EV is not enough to participate in the market, a new entity in the smart grids to gather individual electricity of EVs should be introduced in order for EVs to act in the market. It is called an Aggregator which is introduced in 2001 by Kempton et al. [5]. The aggregator serves as a middleman between EVs owner and electrical utilities or market for the electricity business. It collects and distributes the electric power bi-directionally for a certain group of EVs between grids and EVs. It can supply properly electricity to EV users based on individual contracts and can collect small-scale electricity from individual EVs in order to transact electricity in the market in large scale on behalf of its associated groups of EVs. So, the aggregator creates a “virtual power plant” where the number of vehicles plugged at any given time of the day and the amount of available electrical energy and power are properly estimated. Its concept is already introduced in many publications [3, 6–8]. In this paper, we propose how the aggregator can operate in the market as a broker of EVs in Korean Smart Grid environment.

23.2 Smart Grids and EVs in Korea

In order to address climate change, the Korean government has recognized the necessity of rolling out the Smart Grids as the infrastructure for the low carbon and green industry, and has pursued the Smart Grids initiative as a national policy to implement the vision of Low carbon, Green growth. Korean government has thought the smart grids technology as one of the most important future technologies from 2008 and announced a long term plan in three stages. As the first stage, the Smart Grid Testbed had been constructed in Jeju Island until 2010 and is being operated to test relevant technologies. The testbed is currently used as an infrastructure for Korean anchor enterprises to examine their products and technologies. The second stage is to expand the testbed into metropolitan areas to add intelligence on the part of consumers from

2011 to 2013. In the third stage, it is planned to complete the nation-wide Smart Grid project for all of the intelligent grid networks from 2012 to 2030.

Upon this plan, the Smart Grid project has been conducted in five areas: Smart Power Grid, Smart Place, Smart Transportation, Smart Renewable, Smart Electricity Service, and the testbed as the first action plan had been deployed in Jeju Island for those technologies until 2013. The testbed is located in northern east area in Jeju Island, and about 6,000 houses are included in the testbed. Twelve consortia of 168 companies joined the testbed project.

As results of the testbed project, the infrastructure for EVs including many electricity chargers has been deployed: until Dec 2013, sixty seven fast chargers and four hundreds ninety seven slow chargers had been deployed as shown in Fig. 23.1. It is because Jeju province wants to make Jeju island as one of representative places for the EV with the vision of Carbon free island Jeju by 2030. Therefore Jeju province plans to substitute 10 % of total vehicles in Jeju Island with EVs by mainly changing the rent-a-cars and vehicles belonging to public institutes, and deploy 29,000 EVs and 29,700 electricity chargers until 2017. Also, 30 % of total vehicles are planned to be changed to the EVs by mainly changing buses and rent-a-cars, so that 94,000 EVs and 94,900 chargers are to proliferate until 2020. Until 2030, Jeju Province wants to replace 100 % of the vehicles with EVs by 371,000 EVs and 225,000 chargers. Therefore, it is expected that various businesses based on Smart Grids will be developed using EVs in Korean environments.

Besides, a law related to the electricity transactions—the Electricity Enterprises Act has been revised in Korea in 2014. The existing electricity transaction model was to apply only the reliability based demand response (DR) program [9] that reduces demands on duty from the low demand resources in the bidding prices for the allocated demand reduction quantities already determined through demand forecasting on the previous date. The electricity has been transacted only for providing the electric power reliably in KEPCO’s monopolized way. But by integrating the economic-based DR program to the reliability DR after the mid 2014, and integrating the incentive-based DR program additionally from 2015, private enterprises are permitted to participate in operations of the response management market that had been operated only by Korea Electric Power Corporation (KEPCO) and Korea Power Exchange (KPX), so that business environments for various electric power suppliers to act in the Korean electricity market have been newly created.

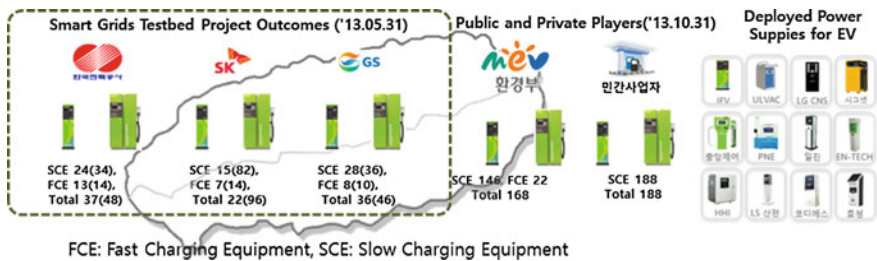


Fig. 23.1 Several kinds of EV chargers in Jeju smart grids testbed

In this changing environment, it is expected that EV aggregators can participate in the electricity transaction market as a main subject by collecting sufficient quantity of the electric power from a certain group of EVs according to the V2G concept in Korea in the near future.

23.3 Aggregator Operations

23.3.1 The Aggregator in the Smart Grids

The following entities in Fig. 23.2 can be considered as main subjects in the Smart Grids: Electric power Generator, Transmission System Operator (TSO), Distribution System Operator (DSO), Electricity Supplier, Electricity Market, Aggregator, EV, Electricity consumer and so on. The electric power companies generate and deliver electricity in sufficient quantity to areas that need electric power through grid connections. Electric power generation is conducted by central or distributed power stations. TSO is an operator that transmits electrical power from generation plants over the electrical grid to regional or local electricity distribution operators in electrical power business. DSO in the grids distributes electrical energy from the transmission system to consumers. Korea Electric Power Corporation (KEPCO) is traditionally a corporation responsible for generate, transmit and distribute the electric power in Korea, but

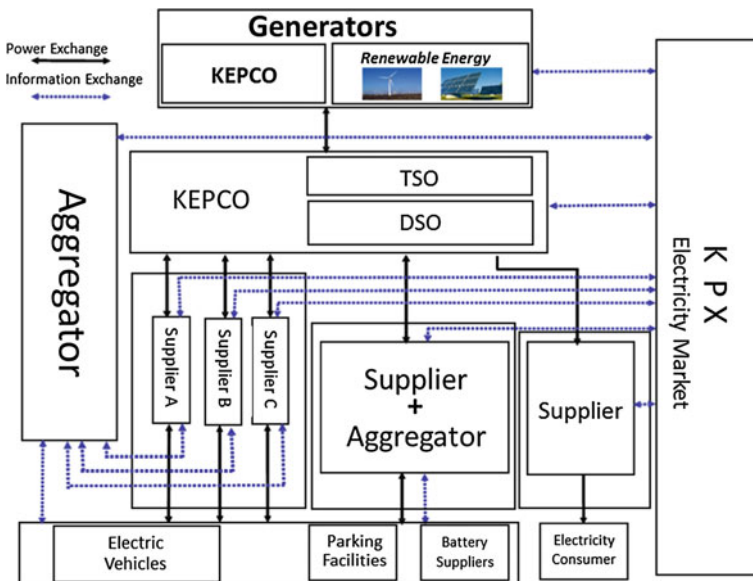


Fig. 23.2 A framework for aggregator operations in Korea

as several kinds of renewable energies are recently developed to generate electricity and sell it to KEPCO, various renewable power sources can contribute to generating electricity now. Also, until now the Electricity Supplier is monopolized by KEPCO in Korea. But, as private companies such as SK Energy and GS Caltex have participate in the Korean Smart Grid Testbed Project and it is expected that many companies including them should have some kinds of roles to supply the electricity to customers with development of the smart grids technology in the near future, we expect private companies will join the electric power supplier.

The electricity transactions can be made in the market through participation of those entities in the smart grids. Electric power can be directly delivered to EV users by suppliers via TSO and DSO from Generators, or an aggregator can mediate to supply it from suppliers to EV users. No matter how to deliver the electric power, however, all of transaction information for EVs buying and/or selling the electric power may be transferred through the aggregator acting as a broker. It is because an individual EV itself cannot participate in the electricity market nor have any transactions with electrical utilities due to its small power capacity. Actually Quinn et al. have compared the electricity transactions through the aggregator to the direct transactions of the EVs [10]. The direct transactions have shown low availability during a day, but transactions through the aggregator have resulted in 100% availability in the most of time. It is because the EV owner's individual behaviors cannot be managed but the aggregator can control available capacity of user groups by proper contracts according to its own business model.

Figure 23.3 shows how the aggregator can exchange the electricity with power grids. EVs are connected to the grids for power and to the Internet for communications through their aggregator and other grid operators. The grid operator continuously monitors grid operations and decides to send dispatch command to control the

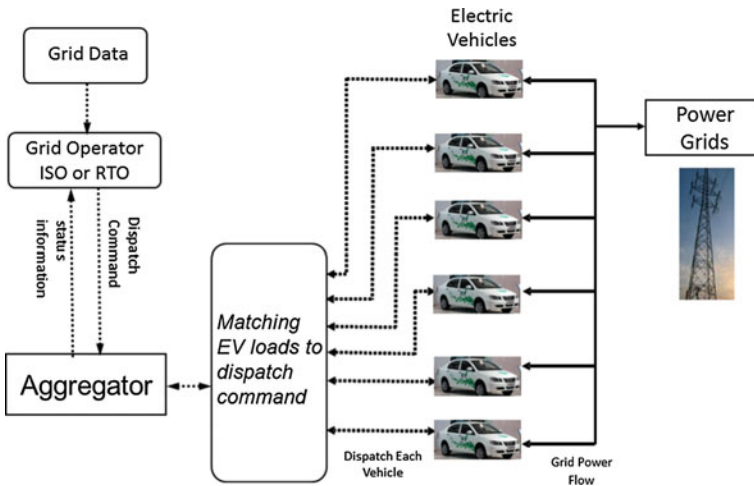


Fig. 23.3 Aggregator operation

aggregate load to the aggregator and/or actors in smart grids. When a command is received by an aggregator, it chooses some EVs to perform actions assigned by that particular command.

The aggregator as a business entity can be realized as several types of corporations. Kempton et al. have mentioned that the aggregator can be local distribution companies, Energy Service Companies, vehicle manufacturers, cell phone operators, electricity retailer or etc. [5]. As another example, an aggregator that establishes an infrastructure in a parking lot in the area of houses, shopping centers or business buildings and holds contracts with a certain groups of EV owners can be considered. If it makes contracts with certain memberships with EV owners and various price policies are provided to them, it can possibly limit moving patterns of EVs or control charging level according to the contracts. Bessa et al. introduces the following four types of groups of EV users that the aggregator may handle [11]: uncontrollable load, partially-controllable load, controllable load and controllable resource. As the EV owners in cases of the first and the second types of users may simply want their EVs to be appropriately charged to amount they want without further services, the aggregator cannot control over charging rate. But for cases of the controllable load and the controllable resource, the aggregator can identify charging location, charging time and control the charging rate according to their contracts. In later cases, the aggregator can suggest some kinds of options to EV owners and properly utilize the charged electric power of its EVs, and then it can join the electricity market and exchange the electricity with power grids.

23.3.2 Role of the Aggregator

Ahmed et al. [12] describes the followings are required for the aggregator.

- Users update to generate charging policy
- Transfer of charging policy (computed by the optimization engine) to the charger
- Monitoring of the charging/V2G process
- Monitoring of net energy level from all vehicles
- Monitoring of connected vehicles to calculate net storage capacity and quantity

The charging policy is to determine when the aggregator can get the electric power from the grids to charge the EV and when the electricity collected from EVs can be delivered to the power grids. As explained in Fig. 23.3, delivery of the electric power between EVs and power grids is performed by dispatch command from the grid operator based on the electricity prices and time slots determined in the electricity market with referencing the aggregator status. Therefore, the aggregator should manage several things such as the charging policy, energy levels for charging and discharging, the number of connected EVs and etc.

It is needed to clarify how the aggregator exchanges what information with actors in the smart grids. Figure 23.4 shows the role of the aggregator between the chargers and public and private suppliers to charge the EVs.

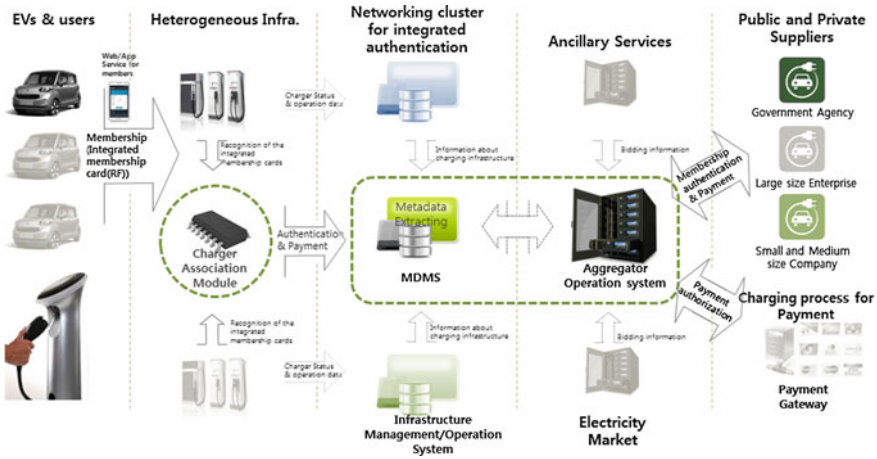


Fig. 23.4 Aggregator operation for charging EVs

1. EV & Users: As in [13, 14] authentication and payment information should be exchanged between the EV user and the charger before and after the charging, respectively. These authentication and payment can be performed using various devices such as RFID, the credit card and etc., and usually using the membership based on the contracts between the aggregator and EV owners. Also, users can get information about the charging status as well as the payment process through the web or smart phone application services.
2. Heterogeneous Infrastructure: As introduced in Fig. 23.1, there are various chargers and many of them are incompatible each other. As the aggregator should handle information from the chargers, the charger association module (ACM) can integrate the information from various chargers and transfer it to the aggregator. For this collection, each charger gets information about membership of EV owner. All of information about chargers status and operation data, authentication data, payment data and EV owners data can be transferred to the aggregator.
3. Meta-Data Management System: Based on received data through ACM, this system stores meta-data for all data that the aggregator needs by interacting with Infrastructure Management/Operation System and Networking cluster for integrated authentication. Functional meat-data for interacting with EV chargers and Aggregator Operation System is to handle data about the charger, charging information and membership authentication as shown in Fig. 23.5.
4. Aggregator Operation System: It plays a subject role to sell and buy the electricity in the electricity market based on the charging status of a certain group of EVs with which it has contracts. Also, it can use the spinning and regulating reserve through the ancillary services. It can use its own optimization engine to participate in the electricity market by selling or buying the electric power in proper prices, and after transaction the electric power can be transmitted or received to/from the power grids during allocated time slots.

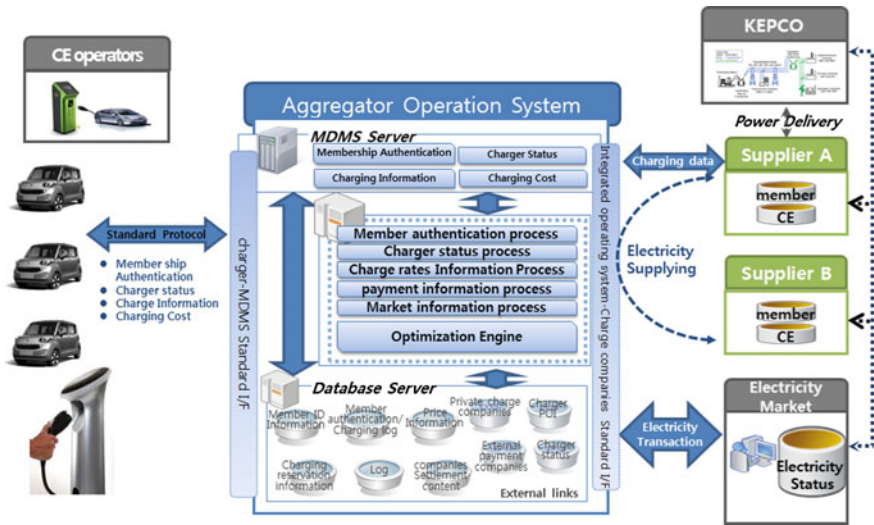


Fig. 23.5 Aggregator operation system

Figure 23.5 shows an architecture that an aggregator interacts with other systems in the smart grids. Actually one of the main roles of the aggregator is to buy or sell the electricity in the market by properly estimating electricity prices based on information gotten from the electricity market. Then, it can provide the electricity to its EVs, and also be a virtual generator using the electricity collected from its EVs to power grids at the best prices. As this role is important, but very complex, this transaction should be managed by the human beings. Therefore, the aggregator operation system just run the optimization engine using data gotten from the electricity market and provide the best prices and time slots to the manager, then it can participate in bidding at the market on the behalf of the manager using the market information process module. Besides, the aggregator can handle EVs and the chargers using several modules such as member authentication process, charger status process, charger rate information process and payment information process.

The electricity market consists of two kinds of markets in Korea: Day-ahead market and Hour-ahead market. The Day-ahead market opens at 15 o'clock and continues until 16 o'clock on the previous date. And, the decided reserves and reduction quantity is notified to entities participating in the market at 17 o'clock. The Hour-ahead market opens 3 h ago and closes 2 h ago before actual electric power delivery. Also, the decided reduction amount is notified 1 h ago to the correspondents. Korean electricity market deals the electricity in a unit of 30 min. Figure 23.6 shows a time sequence of the electricity exchange in the Korean market, and contains some variables considered for the optimization computation.

In this paper we have explained operations of the aggregator in the Smart Grids framework considering Korean current situation. As results of Korean Smart Grids testbed, current situation about EVs and charging infrastructures deployed in Jeju

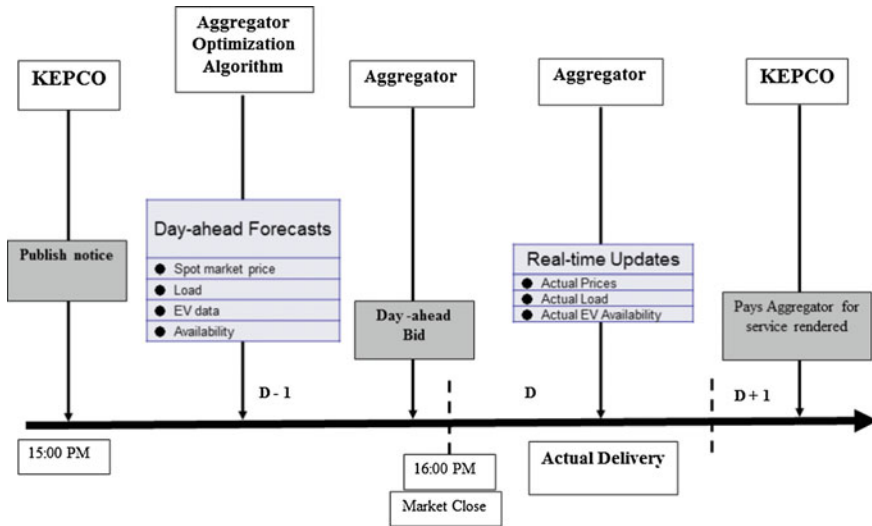


Fig. 23.6 Korean market operation

Island is described. As a recently revised law—the Electricity Enterprises Act is introduced, technical and business effects are considered for the aggregator in the system. Also, our aggregator and its operations are described in those environments. In the future, we expect many kinds of businesses can be created using the aggregator in Korea based on the Smart Grid Technology.

23.4 Concluding Remarks

In the paper we have proposed a framework for applying EVs to the smart grids in Korean situation. Interactions with several entities of the smart grids have been considered. Korean government has established a Smart Grid testbed that can be used to examine products and services in the power grids. As results from establishing the testbed, many EVs and related infrastructure have been built in Jeju Island. Also, the aggregator concept has been introduced and its architecture and operations has been described. The aggregator can be a broker to buy or sell the electric power in the market, and also with proper contracts, can be a virtual power plant in the smart grids. It is postulated that the proposed framework would enable EV users to buy and sell its electricity in the market, and the aggregator can have one of the most important roles to build the smart grid based business.

Acknowledgments This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE), Korea Institute for Advancement of Technology (KIAT) through the Inter-ER Cooperation Projects.

References

1. Vale, Z., et al.: Computational intelligence applications for future power systems. In: Madureira, A., Ferreira, J., Vale, Z. (eds.) *Computational Intelligence for Engineering Systems*, vol. 46, pp. 176–193. Springer, Dordrecht (2011)
2. Jeju Smart Grid Testbed. Retrieved from <http://smartgrid.jeju.go.kr/eng/>
3. Kempton, W., Tomic, J.: Vehicle-to-grid power fundamentals: calculating capacity and net revenue. *J. Power Sources* **144**(1), 268–279 (2005)
4. Kempton, W., Letendre, S.E.: Electric vehicles as a new powersource for electric utilities. *Transp. Res. Part D: Transp. Environ.* **2**(3), 157–175 (1997)
5. Kempton, W., Tomic, J., Letendre, S., Brooks, A., Lipman, T.: Vehicle to grid power: battery, hybrid, and fuel cell vehicles as resources for distributed electric power in California. Technical Report ECD-ITS-RR-OI-03, UC Davis Institute for Transportation Studies June 2001
6. Brooks, A.: Vehicle-to-grid demonstration project: grid regulation ancillary service with a battery electric vehicle. Technical Report, AC Propulsion, Inc. December 2002
7. Kempton, W., Udo, V., Huber, K., Komara, K., Letendre, S., Brunner, S.D., Pearre, N.: A test of vehicle-to-grid (V2G) for energy storage and frequency regulation in the PJM system. Technical Report, MAGIC Consortium, January 2009
8. Guille, C., Gross, G.: A conceptual framework for the vehicle-to-grid (V2G) implementation. *Energy Policy* **37**(11), 4379–4390 (2009)
9. Yoo, T.H., Kwon, H.-G., Lee, H.C., Rhee, C.-H., Yoon, Y.-T., Par J.-K.: Development of reliability based demand response program in Korea. In: 2011 IEEE PES Innovative Smart Grid Technologies (ISGT), 17–19 January 2011, pp. 1–6 (2011)
10. Quinn, C., Zimmerle, D., Bradley, T.H.: The effect of communication architecture on the availability, reliability, and economics of plug-in hybrid electric vehicle-to-grid ancillary services. *J. Power Sources* **195**(5), 1500–1509 (2010)
11. Bessa, R.J., Matos, M.A.: The role of an aggregator agent for EV in the electricity market. In: 7th Mediterranean Conference and Exhibition on Power Generation, Transmission, Distribution and Energy Conversion MedPower 2010, 7–10 November 2010, pp. 1–9 (2010)
12. Ahmed, F.S., Laghrouche, S., Lassabe, F.: Aggregator based ICT architecture for electric vehicle Fleet operators. In: 2013 International Renewable and Sustainable Energy Conference (IRSEC), 7–9 March 2013, pp. 542–547 (2013)
13. Song, W.-C.: Authentication system for electrical charging of electrical vehicles in the housing development. In: *Proceedings of Security-Enriched Urban Computing and Smart Grid. Communications in Computer and Information Science*, vol. 78, pp. 261–266 (2010)
14. Song, W.-C. et al.: Authentication for electrical charging of EVs in the private area considering E-roaming. 2013 IEEE ICCE-China Workshop (ICCE-China), 11–13 April 2013, pp. 41–42 (2013)

Part IV
Data-Oriented and Software-Intensive
Systems

Chapter 24

Bio-informatics with Genetic Steganography Technique

Raniyah Wazirali, Zenon Chaczko and Lucia Carrion

Abstract Biological systems have been a rich source of stimulation for computer security specialists. A wide number of approaches have been proposed over the last decade for encoding data using deoxyribo nucleic acid (DNA), giving rise to the developing area of DNA data hiding. In this work, a new data hiding technique based upon DNA characteristic have been developed. DNA matrix has been used to represent the secret message. After that DNA matrix converted to Quick Response (QR) representation that offers a broad scope of practical usage. In addition, the paper provides an idea of choosing the optimal locations of the QR in order to obtain rightmost position. A new system based on the genetic algorithm has been developed. Finally, we embed the QR codes into the most appropriate location by applying the Haar Wavelet technique on the resulting DNA signals and LSB with assist of the GA in order to reduce the error between the cover image and the stego image. Experimental results have presented a high PSNR which indicate a high level of quality in stego-image with high capacity.

24.1 Introduction

For many years, computer security professionals have been intrigued by biological systems. Many Biological systems in the past have been a great source of inspiration for many computer professionals all over the world [13]. Therefore, over the previous decade, a number of strategies have been proposed concerning the use of deoxyribo nucleic acid (DNA) in encoding information. This has led to an emerging

R. Wazirali (✉) · Z. Chaczko · L. Carrion
Faculty of Engineering and Information Technology, University of Technology,
Sydney, Australia
e-mail: Raniyah.A.Wazirali@student.uts.edu.au

Z. Chaczko
e-mail: zenon.chaczko@uts.edu.au

L. Carrion
e-mail: Lucia.Carrion@uts.edu.au

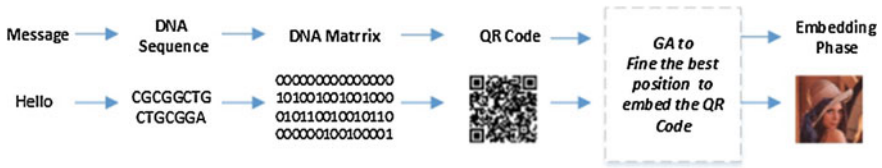


Fig. 24.1 Flowchart of the proposed method of DNA with QR code

field known as DNA data embedding [8]. In this field, data hiding is based upon DNA characteristics of a particular individual. Therefore, the secret message is represented by the DNA matrix. Furthermore, conversion of the DNA matrix into Quick Response (QR) opens the door to a variety of practical applications. When biologists discovered the structure of DNA, they started to change their perspective regarding information science. This is because it became apparent that information processing could be used not only to process, but also to represent and understand structures, activities and distribution of living organisms [16]. This has also led to many changes in the field of computer science. Some of these changes include DNA-based cryptography, molecular DNA computing and DNA Steganography. Steganography has been described as the science of developing hidden messages within harmless messages or carriers [14]. Carriers could be videos, sound tracks, images and DNA. Data hiding within the DNA is based on its apparent randomness. Furthermore, its tremendous storage capability as well as the ability to synthesize sequences in whatever length makes it the best medium of hiding data. This paper discusses some of the approaches used in DNA Steganography. Furthermore, the paper proposes to find the most appropriate positions on the spatial domain and frequency domain that provide slightest statistic features disruption on the spatial domain and frequency domain. Based on Genetic Algorithm (GA), GA affords a technique for hiding QR in the rightmost location which assurance minimum disturbance on the cover images. Figure 24.1 shows the flowchart for the proposed method.

24.2 Related Works

The first successful DNA based hiding technique utilizes microdots similar to the ones used during the World War II. Furthermore, an encryption algorithm has already been developed. This algorithm uses a plaintext binary encoded message and mixes it with dummy strands in order to achieve information hiding founded on DNA binary sequences [12]. The other approach is a reversible scheme for information hiding that utilizes reversible contrast technology. The approach that will be discussed in detail, in this paper is the DNA_QR technique. This data hiding method utilizes a steganography algorithm that hides DNA encrypted secret messages within QR codes that are then randomized and finally embedded within a common image [10] as shown in Fig. 24.1. QR codes are two dimensional matrix barcodes useful in encoding

Table 24.1 Triplet character encoding rule

Character = Triplet			
1 = ACC	A = CGA	K = AGG	U = CTG
2 = TAG	B = CCA	L = TGC	V = CCT
3 = GCA	C = GTT	M = TCC	W = CCG
4 = GAG	D = TTG	N = TCT	X = CTA
5 = AGA	E = GGC	O = GGA	Y = AAA
6 = TTA	F = GGT	P = GTG	Z = CTT
7 = ACA	G = TTT	Q = ACC	= ATA
8 = AGG	H = CGC	R = TCA	, = TCG
9 = GCG	I = ATG	S = ACG	. = GAT
0 = ACT	J = AGT	T = TTC	:= GCT

information. Compared to regular barcodes, QR codes are more common because they offer a higher storage capacity. DNA_QR combines a strong DNA encryption algorithm with a two-stage data hiding technique making the whole process very hard to break [19].

24.2.1 DNA Encryption

DNA encryption is an information hiding techniques inspired by the micro-dots technique used during World War II. Through this technique, scientists are able to produce artificial DNA strands containing secret messages [1, 15]. Table 24.1 displays a triplet that is used in encoding individual characters or numbers. DNA encryption utilizes a simple substitution algorithm that encodes individual characters into DNA sequences through Eq. 24.1 where the decoding through Eq. 24.2.

$$Encoding : X \longrightarrow Y \tag{24.1}$$

$$X \in [A, B, \dots, Z, 1, 2, \dots, 0, " .", " ,", " ;", " :"]$$

$$Y \in [xyz :: x, y, z \in [A, C, G, T]]$$

The function used for decoding is:

$$Decoding : Y \longrightarrow X \tag{24.2}$$

24.3 Polynomial Representation of DNA Matrix

In the following sections, DNA sequences are represented as a two-dimensional signal images using matrices. After that, a sparse polynomial representation of the DNA is considered. Lastly, these formulations will be discussed as QR formats.

Firstly, various scholars note that the best way of storing information is through the use of QR images instead of DNA sequences, this is because the resulting database is not only compact but is also easier to traverse when performing searches. However, apart from QR code, DNA sequence can also be represented in the form of a matrix consisting of four rows. Each row represents the presence of a DNA base, be it T, G, C or A. For instance, a DNA sequence with five rows and a base of TCCGATAACGACT is represented as follows in Eq. 24.3:

$$\begin{matrix}
 A & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
 C & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
 G & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 T & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{matrix} \tag{24.3}$$

In the above matrix, the first row is a representation of occurrence of molecule A in the DNA sequence. The other rows represent molecules C, G, and T, therefore, the matrix can be visually represented using dots signifying occurrence of digit one and blanks signifying occurrence of digit zero. Secondly, DNA encryption involves the development of sparse polynomials [6]. Actually, the DNA matrix represented above is a sparse matrix made up of four sparse vectors in each one of the bases (base A, base C, base G and base T). Therefore, row vectors can be utilized to create four sparse polynomials using formula 24.4:

$$p_{9x0} = \sum_{i=1}^n a_i x^i \tag{24.4}$$

In the above case, a_i represents $\{0, 1\}$, whereas n represents Genome length. Therefore, the previously documented matrix results in Eqs. 24.5–24.8 sparse polynomials:

$$P_A = X^5 + X^7 + X^8 + X^{11} \tag{24.5}$$

$$P_C = X^2 + X^3 + X^9 + X^{12} \tag{24.6}$$

$$P_G = X^4 + X^{10} \tag{24.7}$$

$$P_T = X^1 + X^6 + X^{13} \tag{24.8}$$

Therefore, the power of each of the terms can be stored within each polynomial instead of storing the whole sequence. The following are the resulting Eqs. 24.9–24.12:

$$P_A = (5, 7, 8, 11) \tag{24.9}$$

$$P_C = (2, 3, 9, 12) \tag{24.10}$$

$$P_G = (4, 10) \tag{24.11}$$

$$P_T = (1, 6, 13) \tag{24.12}$$

DNA sequences have millions of bases. Therefore, trying to represent each one of the bases becomes impossible. What DNA encryption does is that it only stores

sparse polynomial numbers made up of very big numbers. The number of storage bits required is reduced by storing only the difference between powers. This is a common practice in lossless data compression . Therefore, when this technique is applied to the discussed DNA string, the following equations are obtained:

$$P_{dA} = (5, 2, 1, 3) \quad (24.13)$$

$$P_{dC} = (2, 1, 6, 3) \quad (24.14)$$

$$P_{dG} = (4, 6) \quad (24.15)$$

$$P_{dT} = (1, 5, 7) \quad (24.16)$$

When multiples of the lengths of the DNA string are augmented with the four polynomials above forming one polynomial, the following is the result show in 24.17.

$$P = (5, 12, 20, 31, 33, 36, 45, 57, 61, 71, 72, 78, 91) \quad (24.17)$$

Finally, the difference sparse polynomial is given by the following Eq. 24.18:

$$p_d = (5, 2, 1, 3, 2, 1, 6, 3, 4, 6, 1, 5, 7) \quad (24.18)$$

This last vector occupies only 52 bits compared to the original vectors that occupied 104 bits. This is a compression ratio of a half. Therefore, representing DNA sequences with sparse polynomials reduces the amount of time consumed traversing the database. This is because each sequence in the database will be stored in the form of four sparse polynomials. These are PdT, PdG, PdC and PdA. At the same time, query sequences will also be represented as four sparse polynomials. These will include PdqT, PdqG, PdqC and PdqA. Hamming distance between all adjacent vectors (hT, hG, hC, hA) are determined with the following equation:

$$H = h_A + h_C + h_G + h_T \quad (24.19)$$

This reveals that the vector having the least hamming distance of (H) will be the closest to the desired query sequence.

24.4 QR Format Design

A QR code is a two dimensional code or a visual matrix code that can be read by Smartphone cameras or QR code readers. These codes are made up of black modules arranged in such a way that they assume a square pattern upon a white background. The encoded information could be a URL, embedded text or any other type of data. Figure 24.2 is an example of a commercial QR code.

Fig. 24.2 QR code



24.5 GA to Find the Rightmost Position

The genetic algorithm (GA) is a search and optimization approach founded on the doctrines of natural selection and genetics and. It enables population involving many people to evolve under itemized selection rubrics to a state which makes the most of the “fitness” while minimizing the cost function. The technique was advanced by John Holland in 1975 [7]. The genetic algorithm begins with no information of the exact solution and based totally on replies from its progress operators such as reproduction, crossover and mutation to get the most suitable solution. By beginning at some independent ideas and examining in parallel, the GA approach prevents local minima and meets to achieve optimal solutions. Therefore, GAs have been used to be accomplished of finding high performance ranges in multifarious domains without suffering the challenges related with high dimensionality, as may happen with rise decent techniques or approaches that trust on imitative information [7, 18].

24.5.1 Process of Concealing the QR

After producing the QR code, the GA will search for the rightmost position to embed this QR code with minimum distortion. The GA will help in reduce the error between the cover image and the stego-image. The process of finding the best place to embed the QR code as follow [17].

1. Creating a random population of chromosomes; each chromosomes with length equal to the size of the produced QR code
2. Assessing the objective (fitness) function; The evaluation of the given function hinges on the PSNR criterion, which is supposed to reach its minimal value for the process to have any meaningful results. PSNR, or Peak Signal to Noise Ratio, being the criterion as the foundation for the fitness function, it is traditionally defined with the help of the functions (24.20) and (24.21):

$$PSNR = 10 \times \log_{10} \frac{Max^2}{MSE} \quad (24.20)$$

$$MSE = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n (A_{ij} - B_{ij})^2 \quad (24.21)$$

The Peak Signal to Noise Ratio PSNR value is considered to be the fitness function. Any value under 30 dB of PSNR values indicate low quality (i.e., distortion caused by embedding is high). A high and acceptable quality stego-image should strive PSNR value of 40 dB, or greater [5]. The higher PSNR value means the minimum changes and the higher quality. The score matrix will evaluate the changes of the cover image for each block.

3. Repeating the steps a–c until the new population is made:
 - a. Choosing a pair of chromosomes (probability increasing together with the function of fitness);
 - b. Forming two new strings with a crossover of chromosomes;
 - c. Mutating the newly obtained chromosomes and plant the new strings in-to the population.
4. Swapping the new and the previous population;
5. As long as the optimum solution can be provided by correcting the value of error with the amount of generations or the maximum amount of generations is attained before it ceased to grow at the point where it serves as the location of the best chromosome, the experiment can be considered successful.

24.6 Embedding DNA_QR in Images Using DWT and LSB

24.6.1 Least Significant Bit Substitution

Two different methods can be used in the creation of QR codes and embedding information within them. The first method is known as LSB. According to various scholars, Least Significant Bit (LSB) substitution is the most commonly used Steganography technique. Presented with any image, LSB can replace least significant bits of every byte with bits from the hidden message. For a gray-level image, every pixel is made up of 8 bits. Since one pixel is capable of displaying 28 variations, these translate to 256 variations. The main idea behind LSB substitution is embedding confidential data starting from the rightmost bits so that the procedure of embedding the data does not greatly affect original pixel value. Rightmost bits have the smallest weight. LSB can be represented mathematically using the Eq. 24.22

$$x'_i = x_i - x_i \bmod 2^k + m_i \quad (24.22)$$

In the Eq. (22), x'_i stands for the stego-image's i th pixel value, x_i stands for the i th pixel value of the original image, whereas m_i stands for the decimal value belonging to the i th block in the confidential data. K stands for the number of LSB that will be substituted. In the extraction process, k -rightmost bits are copied directly. The extraction message can be represented using the Eq. 24.23. Figure 24.3 shows an example of the LSB method.

$$m_i = x'_i \bmod 2^k \quad (24.23)$$

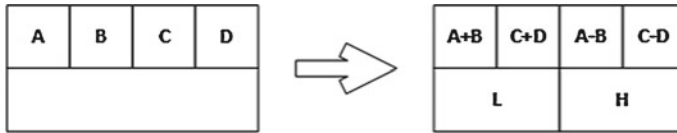
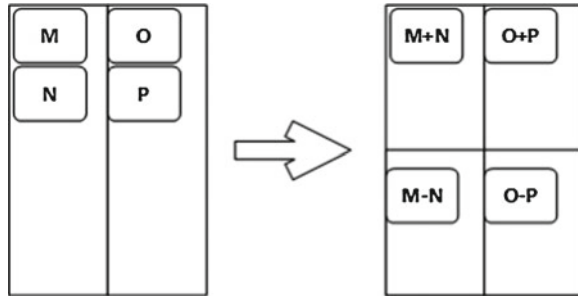


Fig. 24.3 Vertical operation

Fig. 24.4 Horizontal Operation



24.6.2 Discrete Wavelet Transform

The second method of embedding secret message is DWT. This paper utilizes the Haar-DWT frequency domain transform technique [4, 9]. Any 2-dimensional Haar-DWT is made up of two operations. These are the horizontal and vertical operations. This section discusses procedures involved in manipulation of 2-dimensional Haar-DWT. Firstly; pixels are scanned from left to right along a horizontal axis. After this, subtraction and addition operations are performed on neighboring pixels. Sums are stored to the left, whereas differences are stored to the right [11]. This is represented in Fig. 24.3. This is repeated until all rows are completed. Pixel sums stand for low frequency sections of the initial image. They are symbolized by ‘L’. Pixel differences symbolize high frequency parts on the initial image. They are represented by ‘H’.

Secondly, pixels are scanned from top to bottom along a vertical axis. Addition and subtraction operations are performed on neighboring pixels with sums being stored on top and differences at the bottom as shown in Fig. 24.4. This process is repeated until all columns are completed. Lastly, four sub-bands denoted as HH, LH, HL, and LL will be obtained. The LL sub-band represents low frequency portion. It looks quite similar to the initial image.

24.7 Performance Analysis

24.7.1 Hiding Capacity

The payload delivered by a steganographic technique illustrates maximum hiding capability of this algorithm. This means that the total number of bits that is capable of being embedded in a cover media is put into consideration. Since this paper is

dealing with DNA media, the hiding capacity of the steganographic technique is measured in terms of bits per nucleotide (bpn) [3]. Assuming that total length of the DNA sequence referenced in this paper is $|S|$ reflecting all nucleotides making up its sequence, then the DNA encryption algorithm substitutes each pair of message bases with another pair of bases. This means that any DNA sequence can hide a secret message as long as the length of its sequence [3]. This is because two bits of binary message (M) can be carried by a single nucleotide. Therefore, the algorithm's full payload can be expressed in terms of message size in bits. This fact is represented by the following equation:

$$Capacity = \frac{\beta}{\gamma} \quad (24.24)$$

where β is the size of message in bits and γ is the size of cover in bases.

24.7.2 Visual Quality of the Stego-Image

The fastest way of determining visual quality of a digital image is using perception of the human eye. Even though this criterion is effective, results usually differ from one individual to the other. An objective analysis of a digital image can be obtained using a parameter known as Peak Signal to Noise Ratio (PSNR). It is defined using Eqs. (24.20) and (24.21).

In order to calculate PSNR, the dB value must be adopted facilitating quality judgment. The larger value of PSNR the greater the image quality. However, this means that there is little difference between the stego-image and the cover image. Conversely, a smaller PSNR value suggests that the level of distortion between the stego-image and the cover image is high. Due to the use of the GA technique, GA supports to find optimal location to embed the resulted QR code which aim to ensure minimum changes in the cover images.

24.7.3 Robustness Against Attacks

What makes DNA Steganography a robust method of data encoding is that four elements are required in order to crack or determine the secret message. These include the embedding technique, DNA matrix, DNA sequence and break QR code [2]. The Genetic Algorithm provide another layer of security due to the embedding the QR code in unexpected place. The QR code will be embedded in the noisiest location which will make it very strong against the visual attack.

24.8 Experimental Results

Experimental results of the proposed method are presented and discussed in this section. The program was written in Matlab with ZXing library for QR encoding. The images verified in our experiment are all 8-bit images with size 512×512 .



Fig. 24.5 The original Lenna image



Fig. 24.6 **a** DNA and QR code with LSB, **b** DNA and QR code with DWT, **c** DNA encoding without QR in LSB, **d** DNA encoding without QR in DWT

Figure 24.5 shows the original image of the Lenna.jpg. Figure 24.7 shown the stereo image in different situations. Figure 24.6a show the image that have been encoded by DNA and QR code and embedded by LSB. Figure 24.6b show the image that have been encoded by DNA and QR code and embedded by DWT. Evaluation has been

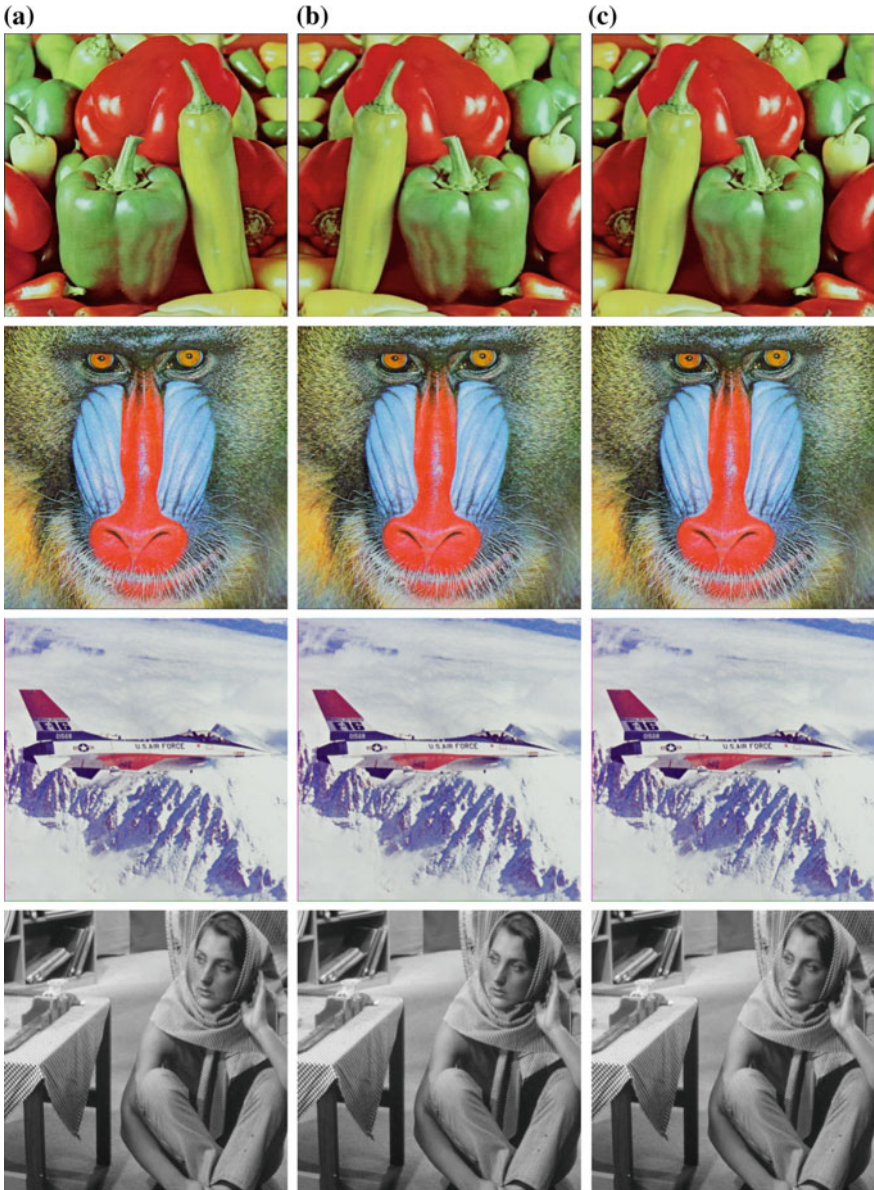


Fig. 24.7 Different variety of the proposed method in LSB and DWT: **a** the original image, **b** the result of DNA with QR coding in LSB method, and **c** the result of DNA with QR code in DWT method

Table 24.2 The result of PSNR and MSE for Lenna

	DNA with QR method		DNA without QR method	
	LSB	DWT	LSB	DWT
PSNR	66.639	701.8532	57.7745	60.9113
MSE	0.0141	0.0053	0.1086	00527

Table 24.3 The result of PSNR for the method in different images

DNA with QR method		
	PSNR (LSB)	PSNR(DWT)
Lenna	66.639	70.853
Peppers	65.854	69.325
Baboon	65.584	68.554
Airplane	65.425	68.521
Barbara	64.985	66.344

done to compare the result of the DNA encryption without QR code to evaluate the quality and the have been shown in Fig. 24.6c, d. The Performance analysis has been explained in Sect. 24.7 (Tables 24.2 and 24.3).

24.9 Conclusion

In conclusion, this paper presents a method that guarantees safety of information as it traverses from one person to another. The main idea behind Steganography is safe and secure data. Therefore, hiding DNA sequences within QR codes offers more practical advantages compared to simply hiding the sequences. Furthermore, the Steganography algorithm is flexible and can be done through DWT of LSB. DNA-QR can be applied on e-banking. Other methods like cryptography can be used together with DNA-QR in order to add another layer of security to secure information exchange. For instance, when cryptography is incorporated, information is first coded before being taken through DNA Steganography. At the same time, other forms of media like music or videos can be used together with Steganography in order to offer well-secured data. Therefore, Steganography is a valid method that offers a credible and well-secured data.

Experiments have been done with assist of the GA to find the optimal locations of the metadata in order to obtain rightmost position. A new system based on the genetic algorithm has been developed. A few images have been tested using various size of text to be concealed. The stego images do not have any noticeable distortion on it that can be realized by the naked eyes. The PSNR value is consider as a fitness function for the proposed method. The proposed method has a high value of the PSNR value which indicates high quality of the stego images.

References

1. Adleman, L.: Computing with DNA: the manipulation of DNA to solve mathematical problems is redefining what is meant by computation. *Sci. Am.* (1998)
2. Cheddad, A., Condell, J., Curran, K., Mc Kevitt, P.: Digital image steganography: survey and analysis of current methods. *Signal Process.* **90**(3), 727–752 (2010)
3. Cui, G., Qin, L., Wang, Y., Zhang, X.: An encryption scheme using DNA technology. In: 3rd International Conference on Bio-Inspired Computing: Theories and Applications, September 2008. 37–42
4. Dan Boneh, J.S., Dunworth, C., Lipton, R.J.: On the computational power of DNA. *Discret. Appl. Math.* **71**, 79–94 (1995)
5. Fridrich, J., Goljan, M., Du, R.: Detecting LSB steganography in scale images. *IEEE Multimed.* **8**, 22–28 (2001)
6. Guo, C., Chang, C.-C., Wang, Z.-H.: A new data hiding scheme based on DNA sequence. *Int. J. Innov. Comput. Inf. Control* **8**(1), 139–149 (2012)
7. Holland, J.: *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence* [Paperback]. MIT Press, Cambridge (1992)
8. Kari, L.: From Micro-soft to Bio-soft: computing with DNA. *World Sci.* **1642**, 1–20 (1997)
9. Kumar D., Singh S.: Secret data writing using DNA sequences. In: 2011 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), April 2011, pp. 402–405 (2011)
10. Leier, A., Richter, C., Banzhaf, W., Rauhe, H.: Cryptography with DNA binary strands. *Biosystems* **57**(1), 13–22 (2000)
11. Magdy, A., Saeb, M., Mohamed, A.B., Khadrage, A.: Haar wavelet transform of the signal representation of DNA sequences. *Haar Wavelet Transform of the Signal Representation of DNA Sequences*, p. 1 (2011)
12. Mousa, H., Moustafa, K., Abdel-wahed, W., Hadhoud, M.: Data hiding based on contrast mapping using DNA medium, **8**(2), (2011)
13. Schyndel, R.G.V., Tirkel, A.Z., Osborne, C.F.: A digital watermark (1994)
14. Shiu, H., Ng, K., Fang, J., Lee, R., Huang, C.: Data hiding methods based upon DNA sequences. *Inf. Sci.* **180**(11), 2196–2208 (2010)
15. Wang, X., Zhang, Q.: DNA computing-based cryptography. In: 2009 Fourth International on Conference on Bio-Inspired Computing, October 2009, pp. 1–3 (2009)
16. Wayner, P.: *Disappearing Cryptography: Information Hiding: Steganography and Watermarking* (2002)
17. Wazirali, R.A., Chaczko, Z., Kale, A.: Digital Multimedia Archiving Based on Optimization Steganography System. *IEEE* (2014)
18. Wu, Y.-T., Shih, F.Y.: Genetic algorithm based methodology for breaking the steganalytic systems. *IEEE Trans. Syst., Man, Cybern. Part B, Cybern.: Publ. IEEE Syst., Man, Cybern. Soc.*, 36(1):24–31 Febuary 2006
19. Xiao, G., Lu, M., Qin, L., Lai, X.: New field of cryptography: DNA cryptography. *Chin. Sci. Bull.* **51**(12), 1413–1420 (2006)

Chapter 25

Augmented Reality and the Adapted of Smart Grid Monitoring for Educational Enhancement

Zenon Chaczko, Wael Alenazy and Amy Tran

Abstract Augmented Reality (AR) technology is increasingly being adopted within the education sector. Technological advancements are changes of users attitudes toward the pedagogical and psychological objectives and goals. This literature review aims to show how the AR technology has evolved to suit specific demands of the education sector. AR technology allows flexible, convenient and effective solution that can enhance the learning outcomes using its various forms and techniques. However, there is a demand to tailor the technology and align it with the ever changing requirements and capabilities of various users in order to improve interactive effectiveness. In this chapter, a case study on AR and Smart-Grid monitoring system show how could be achieved through the video stream and website detection.

25.1 Introduction

Despite being around for over three decades now, Augmented Reality (AR) has not elicited so much attention within the corridors of education sector as it is now. The educators are increasingly finding AR suitable for deployment in education sector. Much of these interests have been down to the increasing scope of its benefits that can be derived by both the educators and the students from. However, the increase in use has primarily been generated by the increasing adoption of information communication technologies which have to dominate every aspects of human life. Technological inventions especially in the computing and mobile telephony

Z. Chaczko · A. Tran
Faculty of Engineering and Information Technology,
University of Technology, Sydney, Australia
e-mail: Zenon.Chaczko@uts.edu.au

A. Tran
e-mail: Amy.Tran@student.uts.edu.au

W. Alenazy (✉)
Preparatory Year Deanship (College), King Saud University,
Riyadh, Kingdom of Saudi Arabia
e-mail: Walenazy@ksu.edu.sa; Wael.Alenazy@student.uts.edu.au

© Springer International Publishing Switzerland 2015
G. Borowik et al. (eds.), *Computational Intelligence and Efficiency in Engineering Systems*, Studies in Computational Intelligence 595,
DOI 10.1007/978-3-319-15720-7_25

industries has revolutionized the use of technology to attain learning outcomes and other pedagogical objectives and goals such as development of social and interpersonal skills which are imperative for the overall growth and development of students. Achieving these objectives and goals requires learning beyond the traditional classroom setting. It requires interactive learning where students can fully utilise both their objective and subjective thinking.

It is these demands of the pedagogical objectives and goals which have seen more and more technology such as AR being incorporated in the learning and teaching processes. Learning requires the use of instruments and approaches that the learners can easily identify with; approaches and instruments that they can easily identify with and learns from without many obstacles. Today's world is dominated by technology especially computing technology. The need to incorporate these technologies into the learning process has seen AR play a vital role in increasing the visualisation of reality contents in learning and teaching process. It is a merger between technology and reality. AR is essential a marriage between the virtual world and the real environment with the aim of making learning all-round, fulfilling, rewarding and fun. It works by generating a composite vision for the user in real time by superimposing the virtual reality onto the real environment and offering a platform for customizing learning. It is the combination of real time scene and virtual scene generated by the

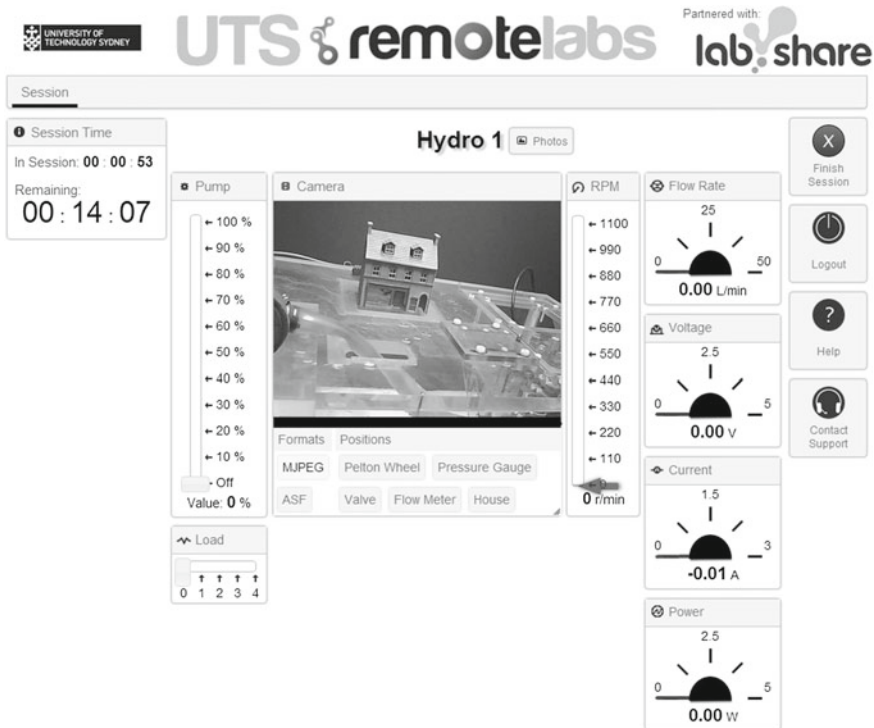


Fig. 25.1 UTS remotelab facility [14]

computer that terminates to new additional information or scene to the user. The augmented reality system allows the students to locate the students to discover and explore the virtual materials as they are in real through the use of overlaid scenes [8]. For example, Fig. 25.1 shows the remotelab environment interface, which is kind of deploying technology into the education sector.

Ideally, augmented reality shifts the meaning of converting the virtual materials into genuine classroom environment by using spectacular technological techniques. It is the technology that schemes digital pedagogical materials onto real objects. The use of AR and its associated technologies in education permits both instructors and students to experience the virtual interaction in real time [10]. Real time learning, sharing of information and pedagogical interactions is a new concept that is in line with the level of civilization in the world today. It is flexible can use the already available technological platforms and devices such as personal computers, laptops, tablets and mobile phones. These are devices commonly used by students and educators alike. Moreover, they are easy to use and therefore suitable for suitable for students from various age groups and education level. Figure 25.2 shows the functionality of using virtual-real lab controlled remotely. Billinghamurst (2002) points out that a smooth transmission between reality and virtuality will enhance and create a new experience. However, the integration into any system needs to be through an intermediate

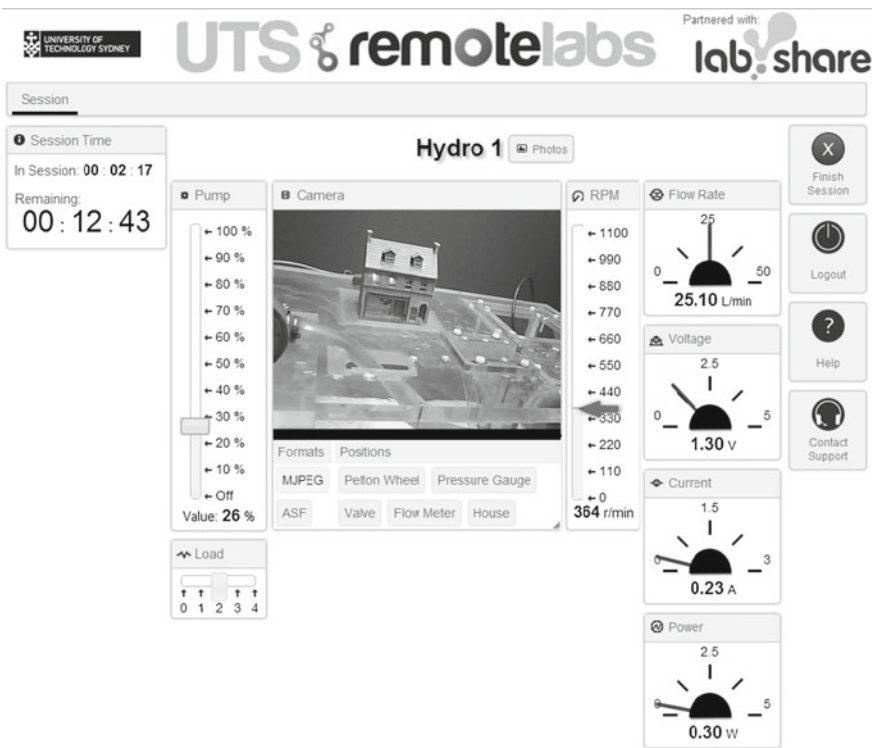


Fig. 25.2 UTS remotelab facility [14]

interface to allow for collaboration. Hence, many of AR and other tangible user interface system have been deployed in classes equipped specially for use with AR. In contrast, in ordinarily equipped classrooms with no special instruments, instructors are responsible for conducting lectures and activities by placing course books in digital form and using computers facilities. Learners follow instructors guidelines while receiving prompts through slides shows, handouts or whiteboard. However, the equipped AR classrooms are relatively expensive making it difficult for their integration into the high collaborative environment that they must to be. Pre-existing teaching using both virtual and real time materials or replanned activities in class is still fairly structured and limited for use with AR [2]. In the education sector, many educators attempt to teach and help their students to acquire the curriculum in an enthusiasm style. By using advanced interactive technology in education, the classrooms are likely to be the favourite place for teaching and learning process. Augmented reality has shaped a new interaction method for many sectors other than education. Consequently, a new interaction level has been rapidly increased by introducing AR. The new advanced technologies within the classroom facilities play a vital role in elevating the motivation levels amongst student while they are taught which is vital when it comes to attaining learning outcomes and other pedagogical outcomes objectives. Figure 25.3 shows the obtained result during the virtual-real time lab session.

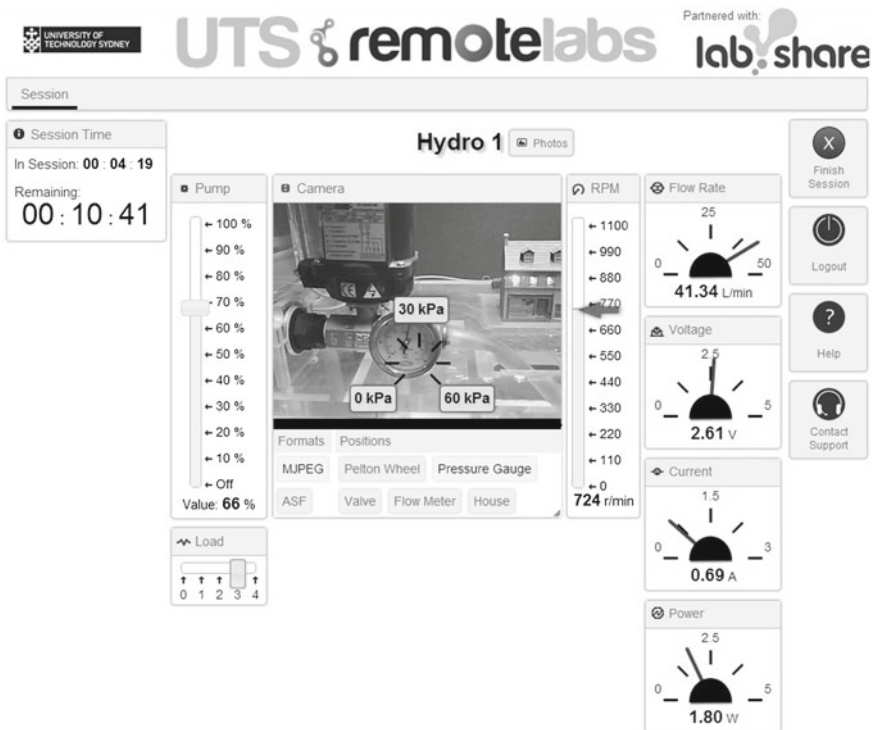


Fig. 25.3 UTS remotelab facility [14]

25.2 Background and AR History

AR gives users a view of reality which has been tweaked or enhanced. The technology augments scenes derived from reality by superimposing virtual reality onto the real context. The distortion is achieved by mixing technology with the actual world familiar to the user. Such simulations of the real world can be applied to any situation or environment and hence its suitability in the education environment where students are supposed to learn and interact with a wide array of situations, ideas and topics in the real world. Therefore, using embedded smart devices with AR enables students and educators alike to explore, in an innovative way, to explore the delivered content or information in today's classrooms according to their desires and the learning outcomes and expectations.

AR was developed in 1960 by computer graphics pioneer Ivan Sutherland and his students at Harvard University [12]. However, the advancements made to the technology are a result of several studies a small number of researchers who studied AR in various institutions such as the U.S. Air Forces Armstrong Laboratory, the NASA Ames research Center and the university of Carolina at Chapel Hill between 1970 and 1980 [12]. In early 1990, Boeing scientists developed an experimental AR system for training purpose [13]. In 1996, a group of researchers at Columbia University developed Touring compact AR system or machine known as MITHrill. Recently, a research is currently ongoing at Naval Research Laboratory in Washington D.C and is aimed at developing Battlefield Augmented Reality Systems (BARS) to supplement real action. This new technology is aimed at developing a system that can integrate virtual and reality to produce tangible reality action that can empower and enhance the learning process. The new technology will provide an intelligent way for instructors and learners to access the content easily and in an exciting way.

25.3 Development Methodology and Techniques of AR System

During the early phase of the integration AR system into the smart-lab or smart classroom, there is a need to involve an interface systems to deal and control AR prompts by the users. One such interface system is called Tangible user Interfaces (TUIs) such as the DigitalDesk which was created by Wellner in 1991. It serves to bridge the divide between digital and paper documents [10]. The main purpose of deploying AR within the tangible user interfaces (TUIs) as an assistance tools is to increase and encourage the collaboration inside the 21st century classrooms. However, testing the implementation setting ignores the complexities of labs or classes. Therefore, the benefits of the additional elements in the classrooms can lead to diminished achievement of the classroom objectives. Hence, specified the evaluation of AR and TUI in the classrooms or labs should be carried out by creating a specific prototype that can be estimated frequently. As a result, several studies have pointed out some aspects that must take into account in order to achieve a high level of acceptances

and deployments. These aspects are: physical size and props that encourage the collaboration by serving interaction visible, using multi-interface to access more than one content, the changes on superficial can have effects that are hard to expect, ensure that using particular technologies that will not interfere with AR and its TUI [10].

25.3.1 Constraints and Requirements

A study by Kerawalla, Luckin, and Woolard (2006) reported that there are diverse constraints when it comes to the immersion AR into education system especially with regards to the attainment of pedagogical goals. These include time, classroom equipment, discipline, and curriculum. Therefore, designing learning and teaching environment needs to consider four principles for succeeded implementation system which are integration, enabling, flexibility and simplicity in deployment.

Due to physical mechanism feature, TUI plays a crucial role in education system. According to Jermann and Dillenbourg (2012), gesturing, physical movement and embodiment can add an instant value to the learning and teaching progress. Gesture plays a role in problem solving and learning by providing the external representation of object or problem. Moreover, it helps the learners to create interpretations by the aspect of freeing up cognitive load. The wisdom of presence in AR also gives the strong cognitive knowledge among the learners because AR provides a balance of the virtualisation and reality.

The combination of expressive-movement-centered view and the space-centered view of tangible interactions contribute to create the data center view which causes the tangible interaction [10]. The expressive-movement-centered view emphasizes on the interaction itself than the physical-digital mapping while the space-centered view focuses on the position of the user in space.

25.3.2 The Constructional Implementation Design

In order to integrate any the innovation system in the modern classes, some consideration must take into account for successful implementation. Learning outcomes are as a result of learning activities in classes, thus subject contents-specificity, learners characteristics and the norm of education consciousness must be considered and how the content can be learned. Furthermore, the individual constraints involve previous users experience and cognitive load. The external and internal constraints need to be satisfied by the users to achieve the objectives of integrating AR in classrooms. Moreover, Kerawalla et al. (2011), suggests that successful deployment of AR should consider: (1) system flexibility to adapt to the users (instructors or student) desires, (2) lecture duration must be short, and (3) contents visualisation [10].

Augmented reality uses as a calculated field position and camera angle to impose a layer of virtual objects over the real background scenes [2]. Due to the mixture topology, users at this stage can interact with the virtual objects and access relevant information easily. Despite the limitations of AR applications that could limit information

availability, studies show that the use of AR and its applications in the near future will significantly become common and widely accepted. So, AR systems can be designed and deployed to provide users in education sector personalised scaffolding and sustenance to create their own personal knowledge as they perceive and experience the real world context. Furthermore, through a proposed system, known as the interactive Multimedia Augmented Reality interface for E-learning by Liarakapis et al. (2002) to explore AR in instruction by overlie Virtual Multimedia Content (VMC) information, the AR techniques can serve as a shared medium resulting in collaboration learning environment. Consequently, these will effectively increase learning motivation, engagement and retention among learners because images are more memorable compared to text. As a result, by considering classrooms constraint, the usability of the deployed system within the class will be increased.

25.4 AR Forms and Principles for AR Learning

The embedded digital information within physical environment takes within the Augmented Reality platform. AR is an emerging technology that utilise smart devices, characterised by context-aware devices that enable participants to interact with real-virtual objects in specific substances. In order to utilise and develop AR system for learning progress, according to Dunleavy (2014), three strategies are important for deploying efficient AR learning tools and include enable and challenges aspect, learning styles and the conversion of unobserved objects into real objects. These strategies are designed to leverage AR affordance or address the medium limitations [5].

25.4.1 AR Forms

Currently, AR is presented in two forms by the instruction designers which are AR based on location and AR based on vision. In general, AR based on location serves to leverage enable GPS into smart devices to present digital media to the users while they are outdoors. On the other hand, AR based on vision displays the pointing object through the camera as a digital media to the users [5]. These two AR forms for both students and instructors have positive impacts because of their immersive sensitive context experiences that they provide by the virtual and physical environments reactions. Furthermore, this also can provide a novel and potentially transformative instrument for education process.

25.4.2 AR Strategies

25.4.2.1 Enable and Challenges aspects

Cognitive overload has been reported as one of the challenges students experience frequently in some typical AR classes. AR comes with the designing strategies that

enable students to manipulate the content by accessing and processing and therefore challenge themselves critically. However, the content processing could be high-level and complex leading also to cognitive-overload amongst students. To resolve that, it should take into account the content materials analysis by creating simplified experience structure, raising the complexity as the experience progresses and frames each experience at every step to reach the chosen learning behaviour. The last phase is to replace the result context with multimedia contexts. Consequently, based on specific learning objects, the enable and challenges aspects could be used in AR experience [5].

25.4.2.2 Learning Style

The second strategy directs the students to learn through the game stories or narratives. By creating interactive stories that are distributed and embedded within physical environments AR allows students to collect and synthesise different information in learning process. So, the AR instructional designers at this stage consider two different learning styles in the classes which involve historical fiction and playing games [5].

25.4.2.3 The Conversion of Unobserved Objects Intoreal Objects

The third strategy enables students to visualise the invisible objects. This can be done by the combination of two AR forms techniques to provide the visibility of unseen things more realistic. Creating interactive augmented reality that leverages both location and vision based AR enable learners to present or turn trigger 3D models in smart devices. As a result, the combination techniques shift the ambiguous objects to real-visual objects that are animated and interactive to learners.

25.5 AR Techniques and Its Demand Among Participants

New technological innovations have been implemented and recommended by government and private sectors to enable communication between people. Recently, technologies have increasingly been adopted in teaching because of their advantages. In particular, human-computer interaction principles have numerous uses and potential for revolutionising the education field. Thus, the meaning of virtual information has been presented in an interactive and convincing way. Moreover, it has been indicated that the use of tangible AR interface, such as marker cards, as well as immersive AR environment that is based on user application interface and hardware devices contribute to collaboration between participants [3].

Although current teaching and learning methods seem to be successful, still there is a need for improving learning experience within the globalisations and revolutions provided for by the advent of the Information Age. It is also important to increase and match the level of understanding among students. AR has the potential to improve

such understanding. For example, many of universities have deployed the web-based Virtual Learning Environments (VLE) to aid the teaching progress. A study showed that the use of virtual application in learning process as a tool allows student to learn speedily and in a satisfactory manner [3]. Multimedia combines text, images, video, animations and sound that utilises teaching material and present them in different styles.

25.6 Visualisation Enhancement and Real-Time Interaction Introduced by AR in Classes

Visualisation system opens new dimensions of human computer interaction by providing accessibility and user interface to all areas. AR provides interaction between the devices and users that leads to a decrease tangible interaction. On the other hand, the traditional interactions, in classroom, have limitations and allow less functionality in the workplace to be accomplished. However, technologies that provide and advanced technique, such as gesture and speech interaction, help our daily activities to interact with real world objects in more natural way [11]. Thus, the functionalities of reality-based-interface provide the next generation interface to interact simultaneously with real and augmented environment [11]. To address all obstacles in traditional classes, the augmented reality technique plays a crucial role in enriching and boosting learners to achieve their goals efficiently. Visualising things and creation of a big loophole in understanding are caused by teaching situation and demands for many concept, such as physics, chemistry or biology [11]. AR tries to solve these issues by combining real and virtual information and visualising it in a proper way to understand with the interaction function.

Over few years, augmented reality has become one of the most utilised technologies. It has been deployed in many aspects for different purposes. This technique has the capability to be installed or set-up in various systems situation for performance improvement. However, there is still need to improve the interaction method in a way that will provide the reality of interaction meaning with objects [15]. Supporting the AR technique need to be proved by applications and tangible novel devices. Most of the applications do not support gesture-base direct manipulating of the augmented scene that is responsible for user interaction with objects for more real interaction efficiency [15]. Therefore, researchers have come up with solutions that enable AR interaction which includes new methodology, deploying users hands and fingers for both virtually and the possibility of physical interaction with objects that will be shown to the end user.

According to Dunleavy (2009), a simulation study showed that technology-mediation within AR technique helps with interaction and collaboration within the highly engaging environment in which teachers and students operate. Although the AR simulation system provided prospective added value to the learning and teaching process, it showed some technical, manageability and cognitive challenges [4]. It is expected that AR technology will continue to progress to deliver high quality multimedia-interaction towards more powerful for mix-reality. Three complementary

technological interfaces have shaped students learning and include: the familiar world to the desktop, emerging multi-user virtual environment(MUVE) and augmented reality [1].

Through the network media, the familiar world to the desktop interface provides distributed knowledge and expertise accessibility across space and time. The interface delivers the models that lie behind most of applications, tools and media in education in particular in K-12 education level. Emerging multi-user virtual environment (EMVE) interface provides student with a graphical virtual context which actively engages them. It is an interface that provides rich graphical interaction environments with digital objects and tools. Augmented reality interface enables ubiquitous computing models [4]. Augmenting students experiences and interaction are infused to the real world by the digital resources of AR interface. Multiple implications and the immersion in virtual environments and augmented realities have dramatically influenced and shaped participants learning styles, strengths and preferences in a new way.

A study showed that the integrated AR system into smart-environment has an essential impact to support the system with different type interactions, such as selection, transformation and controlling. Recently, AR applications have been deployed in laboratories for educational purpose. However, the implementation has an enormous set-up demand of computers, gadgets, sensors and large displays. These requirements cause major obstacles to wide AR acceptance. Nevertheless, most of embedded smart devices have arrived with AR feature enabling users to experience AR applications with the inessential requirements anytime and anywhere [15] which leads to increased acceptance and the enthusiasm level.

During earlier stages of AR, many of applications had not been diffused extensively on smart phones through games, navigation and references [15]. Many of these applications did not give much considerations to the interactions and concentrated on formation display on the top of real interaction [9]. Chun and Hiller (2013) came up with the idea of demonstration of direct free-hand gestures in AR and also enabling a direct interaction in a special lens of the smart device that based AR. Designing different interaction types allows users to move and control the objects accuracy. The model works efficiently and is engaging.

25.7 AR in Classrooms Setting

AR allows for the establishment of collaborative environments through teachers and students interaction within virtual objects leading to the creation of various interactive scenarios in the classroom. Virtual reality is the combination of various display and devices interface which results from immersive interactive 3D computer-generated environment. The Mixed Reality refers to the merger of virtual objects with the real three dimensional scene [16]. There are two models that have been developed for mixed reality (MR) in classroom to support classroom teaching and self-conducted learning. The MR for learning underlines the participants intention which is reflected directly by the perceived usefulness and indirectly through perceived ease of use. As a result, it has been indicated the importance of involving an advanced technology

in education is an essentially tool [16]. It allows for participants to engage in group work, carry out experiments and interact with the superimposed virtual information in real time. However, according to Liarokapis and Anderson (2010) support of AR in classes can be more effective by determining the level of collaboration required based on users awareness.

AR can be used to teach in physical models. In chemistry, a study showed that the capability of manipulating AR models using markers and rotating for virtual objects observation brought enjoyable experience to students. Consequently, AR encourages experimentation and improves spatial skills for students, particularly, to those who need to construct 3D geometric construction in geometry education. Moreover, AR supports dynamic presentation of information. Duarte et al. 2005 noted that AR educational application extends to physics by estimating information that varies in real-time, such as acceleration and velocity [8]. With a blend of hardware and software application, the use of AR in education will prove a key component of the future in the education environment.

25.7.1 The Use of AR and VR Systems

Designing a learning system for specific educational stage or level needs to consider the subject contents on one side and pedagogical and psychological matters on other. So, AR system as an extension of VR systems has a potential to impact by presenting the information to users and dealing with multiple tasks at the same time. This attribute allows the learning process to improve users performances while the lesson is conducted using real objects. However, it must be taken into account that individual users have different learning styles and communication ways, which are very important when designing and implementing mixed reality approach [3]. The idea of multi-visualisation in AR allows learners to switch interactively between different presented content [3]. In addition, through the capability of student to navigate and explore multi-dimensional augmentation of the material in detail by students, the concentration level will increase during the learning process.

Using AR gives additional aspects to the teaching and learning progress. With the AR audio-visual content, students and educators have an alternative method for improving their learning outcomes. The virtual multimedia content can enable a student to see real life 3-D material as well as interact with them naturally. Moreover, it is also possible to support AR systems and applications with the use of compatible sensor devices or computer vision techniques to address the most important issues caused by the registration between real and virtual information [3].

25.7.2 The Use of AR and RFID

In the education sector, many of information technologies have been merged into learning environments in order to overcome the drawbacks associated with old-style classes. Some of these technologies, such as WSN, have been introduced in

smart classroom to improve learning outcomes and increase the motivation level. The advanced technologies lead creation of meaning of ubiquitous learning by deploying several gadgets in the class. This aspect has a potential to impact students attitude and behaviour. However, there is demand within the technology revolution to leverage the learning environment with a value system to far more effective teaching and learning process. Introducing augmented reality to learning process provides rich information between learners.

There are plans to develop a new smart environment which integrates different advanced technologies. For example, the Radio Frequency Identification (RFID) allows for the improvement of any environment when deployed because of its mobility features as a wireless sensor device. The integration, according to Hwang et al. (2011), with the integration of e-maps, ubiquitous learning application and RFID technology, enable learners to assess their cognitive knowledge because of the activities during the lesson [2]. Moreover, the approach also helps the instructors and students to deal with different kinds of knowledge or subjects anytime and anywhere depending on learning demand. As a result, the contribution of advanced technology leads to a new innovation teaching model.

25.8 Case Study

The case study utilises augmented reality (AR) elements to monitor captured of a scene for studying purposes, such as a video stream from an iPad camera, within the novel idea a smart grid layout. The process of the experiment that has taken is to develop application and investigate existing AR toolkits to leverage functionalities. Moreover, the significant objective is to investigate the capabilities AR frameworks and Software Development Kits (SDK). The main focus of toolkits is to provide a simple way to overlay an image or animation over a video stream. So, after a number of prototypes utilising different AR toolkits, ultimately the project has optimised an open source image-processing framework OpenCV. The case study has been experimented with the use of the Metaio AR SDK, Wikitude SDK and Total Immersions DFusion SDK. As a result, the monitoring functionality of this application required image-processing tools and because of that OpenCV seems to be sufficient due to its characteristics.

25.8.1 System Design and Architecture

25.8.1.1 System Resources

It has been chosen iPad running iOS 7.1.1 for testing environment for the iOS application. In addition, Xcode 5.1 which is the Integrated Development Environment (IDE)

required developing on the iOS platform. The IDE is provided free from the App Store by Apple and comes with all the essential framework and libraries required to start developing applications. Having one central IDE makes it easy for third party frameworks, such as AR Software Development Kits (SDK), to import to iOS projects. Moreover, Objective-C, Objective-C++ and C++ have also been selected for programming languages that used to develop AR Smart Grid because Objective-C is the default language in the iOS framework. Objective-C++ combines C++ elements with Objective-C to extend the language allowing the use of C++ features. C++ has been utilised for the complex image processing functionality because of the better performance and also allows easy interfacing with the OpenCV SDK, which is written in C++ [7].

OpenCV is an open source library for computational image processing written in optimized C/C++ [6]. The library has a very active support community and extensive documentation. The source code is freely available on GitHub and supports Windows, Linux/Mac, Android and iOS.

25.8.1.2 System Architecture

Figure 25.4 shows the high level architecture of the AR Smart Grid application. The Input Handler takes either a video stream feed or a website as the scene to be displayed. The Smart Grid provides an overlay of the smart Cells of the grid. If a Cell is selected, the Motion Detector monitors the Cell. The Motion Detector utilises the image processing functionalities of the OpenCV SDK. If motion is detected, the Cell generates an Event, which is displayed on the screen. Setting and event logs are persisted by the local application database.

25.8.2 *Functionality and Screenshots Result*

The main functional of the AR Smart Grid application is to monitor selected grid cells for movements or deviations. The smart grid works to create an overlay scene for monitoring; so the application supports a video stream from the iPad devices camera or any online sources. The application interface involves number cells that can be configured or adjusted through user-interface menu by setting the number of rows and columns for monitoring and detections. The motion detector has a threshold index that can also be adjusted in the menu. This value determines the sensitivity of the motion detector, which is useful for particular calibration. Once the motion is detected in the selected cells, an alert popover will overlay the screen and a log of events will be registered.

Figure 25.5 shows that the AR Smart Grid menu is open. This is where the scene input can be changed from camera to a website and vice versa. Adjusting the number of rows and columns with the slider can change the number of smart grid cells. Both

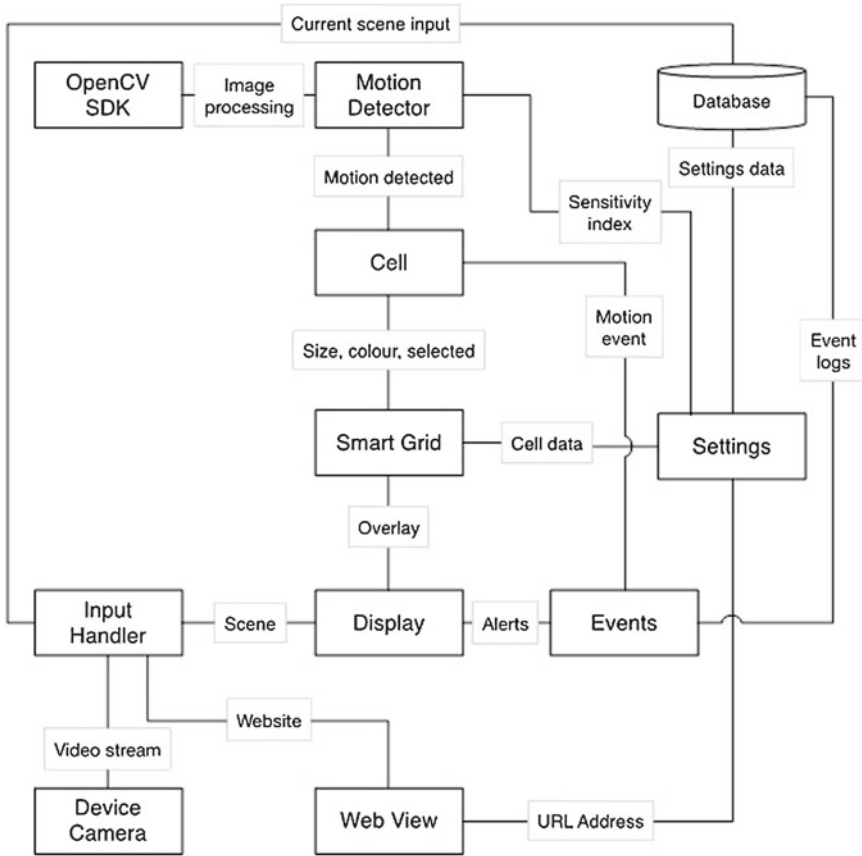


Fig. 25.4 High level architecture of AR Smart Grid application

rows and columns can range from 1 to 10. The Motion Detector Threshold slider controls the sensitivity of the motion detector. This range from 1 to 99, where 1 is the least sensitive and 99 is the most sensitive.

Figure 25.6 shows that the ‘smart grid’ overlay has been adjusted to display six (6) cells, made up of two (2) rows and three (3) columns. This is useful for customising the area of observation.

Figure 25.7 shows that the input source has been changed to a website. The menu has the same configuration options with the addition of a text box that allows the URL source to be changed. This text box only appears for this input source.

Figure 25.8 shows that the cell located at row 2, column 1 has been selected for monitoring. When the graph moved, the alert popover indicating motion has been detected is displayed.



Fig. 25.5 Menu to configure the application



Fig. 25.6 Adjusting the smart grid to six cells

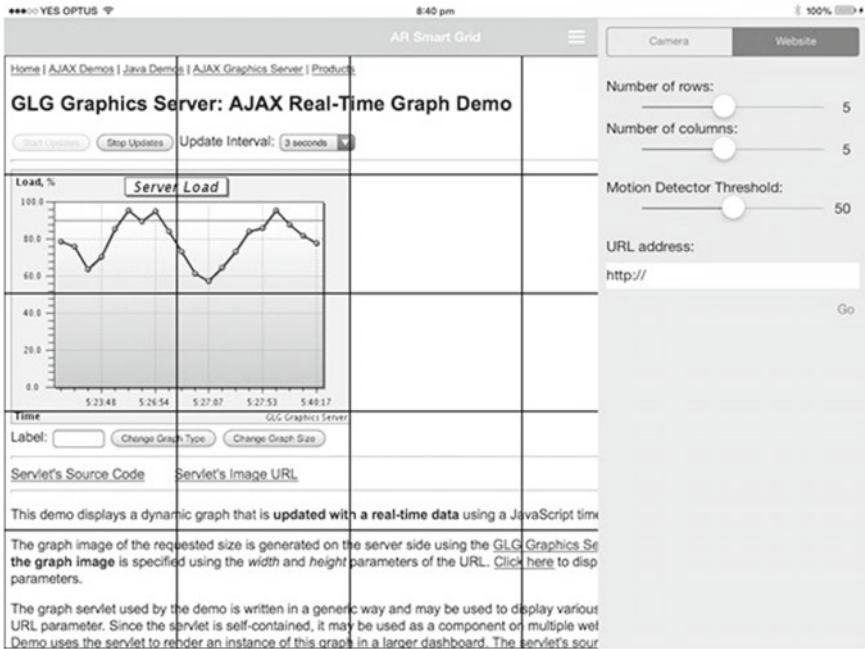


Fig. 25.7 Enter a different URL address in the menu to change source

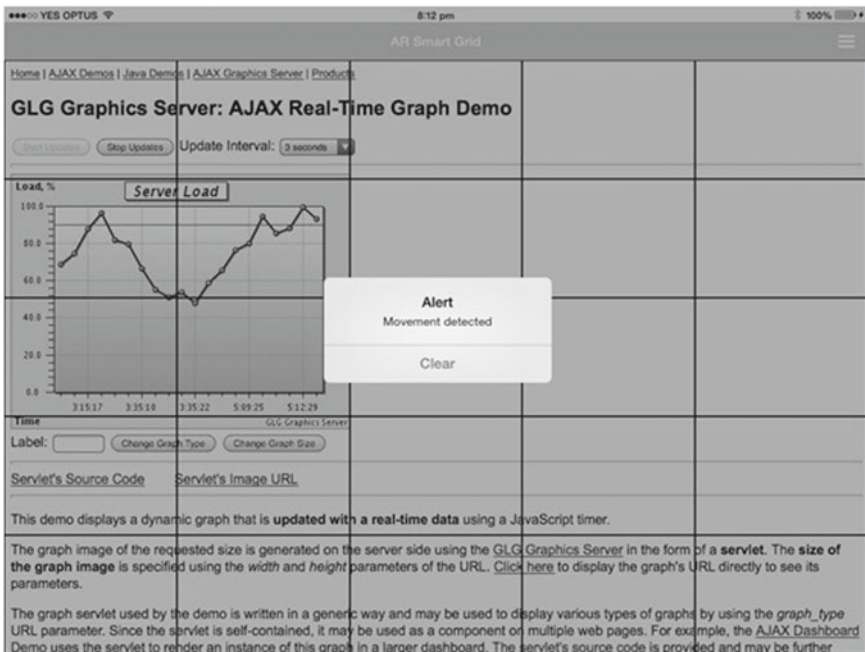


Fig. 25.8 Motion detected alert popover

25.9 Conclusion and Further Study

In the past decades, technologies have increasingly been used within the education sector to increase learning outcomes. Smart devices such augmented reality offer assistance tools have been increasingly integrated within the learning environment to support learning process and create the meaningful learning through an interactive platform. They allow for associative learning which is important in improving knowledge acquisition and retention. AR which is increasingly be deployed in the education environment because of the various improvements made to it, plays a key role in raising motivation levels by presenting them with a more interactive, exciting and familiar environment for learning. It elevates learning to a higher level and puts it on a pedestal that enables customised learning. AR gives the learners a taste of both worlds: it mixes real world experience with the virtual reality with the latter enhancing or distorting the former. Such enhancements give much more details on real world elements compared to traditional classroom learning approach. It uses a platform, information communication technology, which modern day learners and educators are familiar with. With such familiarity, it is easily adopted within the learning environment. It is also very flexible as it can be used for different topics and subjects while the portability of devices using AR adds the much needed convenience in the modern world.

However, there are inherent challenges that come with using technologies such as AR within the learning environment. An integration method into smart classroom is relatively needed to be explored. Moreover, a seamless integration of the technology within the traditional learning environment is challenging because of the need to take into account the various needs within the learning environment. This includes the various learning needs and personalities of the users. Therefore, it is imperative that further studies should be geared towards developing relatively affordable and more versatile AR systems which will take into account the various special needs of learners and users.

References

1. Dede, C.: Immersive interfaces for engagement and learning. *Science* **323**, 66–69 (2009)
2. Chen, D., Chen, M., Huang, T., Hsu, W.: Developing a mobile learning system in augmented reality context. *Int. J. Distrib. Sens. Netw.* (2013)
3. Liarokapis, F., Anderson, E.: Using augmented reality as a medium to assist teaching in higher education (2010)
4. Dunleavy, M., Dede, C., Mitchell, R.: Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *J. Sci. Educ. Technol.* **18**, 7–22 (2009)
5. Dunleavy, M.: Design principles for augmented reality learning. *TechTrends* **58**, 28–34 (2014)
6. OpenCV: Open Source Computer Vision. [online]. Available: <http://www.opencv.org> (2014)
7. Sholtz, P.: How to Make an Augmented Reality Target Shooter Game With OpenCV: Part 34. In: *Raywenderlich Tutorials For Developers & Gamers* (2014)
8. Freitas, R., Campos, P.: SMART: a system of augmented reality for teaching 2nd grade students. In: *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, , vol. 2*. British Computer Society (2008)

9. Azuma, R., Baillet, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. *IEEE Comput. Graph. Appl.* **21**, 34–47 (2001)
10. Cuendet, S., Bonnard, Q., Do-Lenh, S., Dillenbourg, P.: Designing augmented reality for the classroom. *Comput. Educ.* **68**, 557–569 (2013)
11. Prasad, S., Peddoju, S., Ghosh, D.: Mobile augmented reality based interactive teaching & learning system with low computation approach. In: *Computational Intelligence in Control and Automation (CICA), Symposium IEEE* (2013)
12. Höllerer, T., Feiner, S.: *Mobile augmented reality. Telegeoinformatics: Location-Based Computing and Services*. Taylor and Francis Books Ltd, London (2004)
13. Neumann, U., Majoros, A.: Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. In: *Proceedings of the IEEE Virtual Reality Annual International Symposium* (1998)
14. University of Technology, Sydney: UTSremotelabs. [online]. Available: <http://remotelabs.uts.edu.au> (2013)
15. W. Chun, T. Höllerer: Real-time hand interaction for augmented reality on mobile phones. In: *Proceedings of the 2013 International Conference on Intelligent User Interfaces*. ACM (2013)
16. Pan, Z., Cheok, A., Yang, H., Zhu, J., Shi, J.: Virtual reality and mixed reality for virtual learning environments. *Comput. Graph.* **30**, 20–28 (2006)

Chapter 26

Analysis of E-Commerce User Behavior of Indonesian Students: A Preliminary Study of Adaptive E-Commerce

Rianto, Lukito Edi Nugroho and P. Insap Santosa

Abstract Website design is one of several critical factors for the acceptance and success of an e-commerce website. A good website design influences users satisfaction and intention to use the website. This study provides guidance to develop personalized e-commerce for Indonesian students with a method recommended by users and to investigate the characteristics of Indonesian students in using e-commerce. This study used a qualitative method to analyze three components of web design i.e. information, navigation, and visual designs. Data collection was performed using a deep interview technique on Indonesians students who have been familiar with online transactions. The interpretation of qualitative data showed that Indonesian students preferred navigation and information designs to visual design. In particular, Indonesian students preferred a simple layout, easy to use navigation, and trustworthy information that promotes consumers trust.

26.1 Introduction

The purpose of World Wide Web (website) is to help scientists and researchers at CERN (Conseil pour la Recherche Nuclaire Europeene) to complete their tasks. Website is an electronic document publishing that can be accessed by a browser [30]. Currently, website has many changes in both appearance and content. The first generation of website used only text to present information. However, recently the contents of website also filled with images, audio, video, and executable files that allow users to run applications on their browser [4].

Rianto (✉) · L.E. Nugroho · P.I. Santosa
Department of Electrical Engineering and Information Technology,
Universitas Gadjah Mada, Jl. Grafika No. 2, Yogyakarta 55281, Indonesia
e-mail: rianto@mti.ugm.ac.id

L.E. Nugroho
e-mail: lukito@ugm.ac.id

P.I. Santosa
e-mail: insap@jteti.gadjahmada.edu

Furthermore, website has many changes in the user function. In the past, website has only one direct communication; hence, users cannot give a response to the website. Currently website has two direct communications that allow direct interactions between website owners and users by using live chat, messenger or request form. This condition makes a possibility of a website to facilitate online business interactions such as product searching, ordering and buying.

A good website should enable its users to operate it easily, comfortably, and enjoyably. Web designer should focus on the user needs in order to make users ability to use the website efficiently and with satisfaction [31]. E-commerce differs from other website categories i.e. information or entertainment because e-commerce is not technology driven but is user driven [14].

Generally, website design has three components namely information design (ID), navigation design (ND) and visual design (VD). Accuracy of the information on the website is a component of information design. Navigation design is navigational scheme that help users while using the website. Overall graphical look such as balance, aesthetic, emotional appeal and uniformity are among the components of visual design [7].

Based on the function, website is divided into four categories namely entertainment, news, communication, and commerce [21]. You Tube is an example of a website of the entertainment category and BBC is of the news category. On the communication category, Facebook is a good example, whereas Amazon is one of the commerce category. Among the four categories, commerce shows a dramatic increase in term of business which known as e-commerce [26]. E-commerce is defined by product selling, buying, and service through computer network and internet [34].

However, the rapid growth of e-commerce was not comparable with services and features for users. In fact, is about 50% of e-commerce users cancelled their transactions because they found difficulties in using the e-commerce [3, 26]. The difficulties may relate to the poor of the website navigation, not clear information, and slow transaction process. This indicates that e-commerce may have a low level of usability [10].

Usability means how a product can achieve the goals effectively and efficiently [1]. In the e-commerce, usability relates with buying intentions [24], trust [27], and perceived security [18]. Trust and perceived security are important factors that must be provided to increase transaction's intention. In facts, usability website particularly e-commerce was the important factors and therefore improved research in usability's area were needed.

Users have differences characteristic when using an e-commerce. Users may use e-commerce to look for goods in online shops, to compare product prices, and to buy or sell products. Therefore, suitable facilities for e-commerce users are an important factor. Suitable facilities between users and e-commerce are defined by personalization [13]. The purposes of personalization are to provide a customized website which is able to fulfill the requirements of different customers [15] and can highly impact on the perceived usefulness [22]. In the other hand interfaces design can affect online

shoppers trust in the merchant, therefore, will ease making purchases [33]. In the e-commerce, personalization can increase intention of use and related with buying or selling intention. Personalization can be realized with a suitable menu, product search, and payment and delivery method.

Garrigos et al. [13] reported that the personalization is not limited to individuals but can also be a group. Personalization for individual has a disadvantage, because it requires a large and dynamic data [32]. Moreover, this personalization makes difficult to a provider to provide services. This may lead to, for example, television stations to divide the audiences by groups such as age, sex, and program themes. In the e-commerce, users could be grouped by age, sex, and profession such as student, employee, or business owner.

In the other hand, current personalization is done by users such as save setting their themes and a favourite color. Personalization can be done by website engine using representative knowledge to recognize users. This method will create an adaptive website [28] that can recognize the users and adjust website structure to suitable with the users need. Therefore, study about user characteristics to develop user's model is an important factor.

26.2 Related Work

Some research has been done on personalization of the website. Fang and Salvendy [9] investigated web users to provide a good user interface using User Center Design (UCD). Researchers aimed to find the easiness and difficulties of users when using the website. Data collecting was done by interviewing users at Chicago. The results found some rules to provide effective website design. Similar research was done in North America, Korea, China, and England with concerned about (1) user's need; (2) user's objectives; and (3) usability testing [19]. The results of the research were almost similar and also implemented usability testing.

Kuster and Vila [20] used UCD method to explore user's idea using a focus group discussion in small and medium enterprises. The study found that a transaction on a website had three factors i.e. (1) security; (2) informative price; (3) variety in product images. UCD method with focus group discussion was a good method to find website design which was suitable with users.

Personalization can foster user attitudes and maintain a feeling of the website [2]. Whilst, in decision making, personalization can help a user to make a quick decision when there are many confusing choices [3, 25]. The limitations were that some personalization studies have done on visual design, whereas, users required to be facilitated in navigation and content designs.

Personalization research in visual design has been done by Singh and Pereira [29] and found that website design that suitable for users in a certain country was more comfortable than not suitable one. For examples, website design for Arabic users was more comfortable if the scroll bar was at the left side because arabic users start writing from the right side. Moreover, each country has different website designs.

Japan, Germany, USA [6], Egypt [8], Taiwan, Australia [16], and China [23] have their own characteristic designs with regard to personalization using colours, symbols or metaphors that are characteristics in their country.

Based on those reviews, research in the field of personalization particularly in the content and navigation designs was important. The first step was facilitating e-commerce users with suitable visual, content, and navigation by identifying user behaviors for examples, the mostly clicked menus and product purchased, the maximum budget, orientation, and payment method.

One of approach to find users behaviors is web usage mining which consists of the following stages [12]:

1. Detection of user preferences;
2. Identification of users having similar preferences;
3. Construction of a set of web pages highly cohesive from the usage and content points of view;
4. Computation and ranking of interesting recommendations for the active user.

The present study is different from a previous study because the present study was not to improve an existing website but to build a website with appropriate facilities based on Indonesian characteristics. In this study, user profiling was determined by interviewing user to find their characteristics. The characteristics, than were used to develop a model that has ability to identify the users.

26.3 Research Method

The present study was a preliminary study in personalization website, which obtained user's requirements for e-commerce transactions. Using a qualitative method to observe the phenomena of users when interact with an e-commerce website, data collection was performed using a deep interview technique on Indonesian students who have been used online transactions. Investigations were about how features should be presented in e-commerce such as product display, payment, delivery, symbols and text, and colour combinations. Respondents of the current study were ten students from a university in Yogyakarta. Respondents were distinguished by academic level and background, gender, and socioeconomic.

The interview process was divided into three series, including. exploring the students behavior when using e-commerce, exploring students experiences, and collecting the students ideas about suitable features in e-commerce. The purposes of these series were:

1. To find the student characteristics in using e-commerce. This explored the first step done by students when using e-commerce, the reasons behind using online transaction, and the features used to complete the transactions.
2. To obtain student preferences and provide website design that were suitable for students.

3. To investigate the expectations of students regarding features that should be provided in e-commerce.

The process of data analysis was defined as the following steps [5]:

1. Prepared the data for the analysis i.e. transcribing form audio file to text file;
2. Read through the data to obtain a general sense of material;
3. Coded the data i.e. located text segments and put a label;
4. Coded the texts for description to be used in the research report;
5. Coded the texts for themes to be used in the research report.

The output of the data was grouped into three components of web design namely information, navigation, visual design, and presented as a list of e-commerce website features.

26.4 Result and Discussion

Eleven participants, five female and six male from Universitas Teknologi Yogyakarta, Universitas Islam Indonesia, and Universitas Negeri Yogyakarta were recruited in the present study. Interview process was done in four days with an average length of interviews about one hour for each participant.

The first step of the data analysis was done in one week. The next process was reading the data and finding the keywords for the main themes. For examples keywords such as bank transfer, internet banking, and credit card were the themes of order and payment. Whereas, company profiles, contacts, and email or mailing address were the themes of contact. The main themes were used as guidelines to provide facilities in e-commerce. The facilities in e-commerce were grouped as shown in Table 26.1.

The grouped facilities (Table 26.1) were then used to provide a template of facilities that should be available in e-commerce for Indonesia students. The keywords in the grouped facilities were used as candidate names of menus in the e-commerce website. Table 26.2 shows the candidate names of menus.

Table 26.1 Grouping facilities

Transaction	Contact	Feedback	Category	Design
Credit card	Mailing address	Complaints facility	Fashion	Simple
Bank transfer	Pin BBM	Forum	Electronic	Contrast in color
Internet banking	Mobile phone	Order tracking	Gadget	Have a sitemap
ATM	Email address	Testimony	Computer	Easy to use
COD	Messenger	Email form		Help
Compare	Profile	Recommendation		Error handling
Shipping	Form order	Review form		Attractions

Table 26.2 The candidate names of menus in e-commerce

Contact	Product and order	Feedback	Support
Company profile	Product catalogues	Order tracking	Help and documents
Contact name	Product categorization	Complaints form	Error handling
Phone number	Product search	Review form	
Mailing address	Product order	Testimonial form	
Email address	Payment method		
Messenger	Delivery method		

Table 26.3 The mostly click menus and actions

Menu name	Description	Action
Profile	Users need to know history of vendor before transaction	Closed if no profile
Contact	Users need phone number to contact vendor before transaction	Closed when phone number not found
Payment method	Users find a payment method with sequence:(1) cash; (2) bank transfer; (3) ATM transfer; (4) internet banking	Closed when the payment method is only credit card
Help	Users find online documentation to support their transactions	Closed when the website does not have help
Product search	Users find and compare the product before transaction	Closed when search and compare the products are difficult
Testimonial and review	Users need to know vendor reputation in online transaction	Trust will drop when vendors do not have a review

Indonesian students always look at the vendor reputation before proceeding to the transaction. The first step was clicking each menu known as the mostly clicked menu to find detailed information particularly about vendor. Therefore, mostly click menus could identify Indonesian students when they were accessing a website. Table 26.3 represents the mostly click menus and action of Indonesia students.

Based on the results as shown in Table 26.3, Indonesian students preferred accurate information about vendor. In addition, help and documentation were necessary facilities to support users when they have difficulty in transaction process. Supporting facilities in video format or diagram block were the most preferable for users. Users trust also increased when vendors released testimony and reviewed their website. Trust was very useful to influence transaction and intentions of users to use the website. Moreover, users need to find information easily about product such as price and specification before purchasing. Attractive catalogues and easy product searching were the important facilities to support users in a transaction. As a consequent,

Table 26.4 The group of menu candidates in each of website design components

Information	Navigation	Visual
Contact detail	Product catalogue	Simple
Order tracking	Product categorization	Contrast in color
Testimonial and review	Product searching	Attractions
	Transaction order	
	Transaction payment	
	Transaction delivery	
	Help and documentation	
	Feedback	
	Error handling	

users will close the web page or cancel the transactions when these condition was not fulfilled.

According to the three components of website design, namely information, navigation, and visual design, the following process were grouping menu candidates (Table 26.2) into each component of the website design. The purpose of the process was to facilitate data analysis on the three components of website designs to provide a website which was suitable with the users. The menu candidates in each group of website design components were shown in Table 26.4.

Regarding the three component design of the website, contact detail, tracking order, and testimonial and review were grouped into information design because they were related to the accuracy of the information. Whereas, transaction process such as search, order, and payment until supporting and feedback were grouped into navigation design. This was associated with the transaction process, thus required a good navigation for the users. In order to make a comfort transaction, website design should consider the layout and color combination. A simple, attractive, and matching color combinations were preferable. These elements were grouped into visual design which related to aesthetic and emotional appeal to increase users comfortability.

Indonesian students characteristics in online transaction focused before the transaction process. Based on the mostly click menus, Indonesian students had a priority in information design which indicated visiting company profiles, contact detail, and review of vendor. They will cancel the transaction and left the website when the information about vendor was not found. Based on the discussion above, we suggest that Indonesian students characteristics were:

1. Carefully

Indonesian students were very careful in online transaction. Budget limitation may be a reason why Indonesian students were very careful in online transaction. As users, they did not want to be deceived, for example, because it was paid, but the products were not delivered.

2. Maximum tolerance to be fooled

Indonesian students were very sensitive in online transaction because they do

not have enough funds to trial transaction. Hence, they have a maximum tolerances to be fooled in online transaction is about 50.000 IDR. it is usually for the first transaction without a recommendation from someone who had done the transaction.

3. Budget orientation

In purchasing, Indonesia student preferred budget orientation rather than specification. For examples, specifications of the computer will be reduced to adjust the budget. Therefore, when the budget did not meet with the specifications, the purchasing will be delayed until the funds were sufficient.

4. Compare product and price

To get a product that suitable with budget and specifications, Indonesian students usually doing comparisons with the same as well as different brands in the same or another website. To facilitate these characteristic, comparison facilities should be provided to satisfy the users during the transaction process.

5. Cash payment was recommended

Cash payment was preferred by Indonesian students because on the average Indonesian students have no credit or debit cards. Therefore, they preferred transaction by cash on delivery (COD). COD method has two advantages for Indonesian student i.e. (1) Easiness of payment; (2) Transaction certainty.

6. Simple procedure in transaction

Indonesian Student preferred simple transactions such as order, payment, and delivery. Too many procedures made users frustrated and left the website meaning that the transactions were canceled. For examples, procedures of member registration which had too many fields must be filled. The challenge for e-commerce owners was how to make the transaction short, but represented everything.

7. Fashions and electronics were the mostly product purchased

Based on the product purchased, Indonesia student was mostly purchased fashions and electronics when using online transaction. Female and male students have different in product categories to purchased. Female students preferred fashions such as clothes, bags, and shoes, whereas, male students preferred electronic such as computer peripherals e.g. flash disk, CD or DVD writer, and active speaker.

The present study found some characteristics of Indonesian student in online transaction which described in the mostly click menus and the facilities to support the transaction. These findings were important to improve and increase online transactions based on website design. This was similar to the finding of the research [11] that a website design was identified as a key factor for the acceptance and success of a website. The analysis using three components of the website design made easier for website developers to develop their website based on users characteristics in using the website. Furthermore, based on users characteristic a website should be automatically adjusted the structure with users which was known as an adaptive website. To make an adaptive website, the website should have knowledge to identify the users to provide appropriate facilities as found in the present study.

26.5 Conclusions

This present study focused on finding the Indonesian student characteristics in online transaction. In the future, these characteristics as representative knowledge can be used as users model to identify the users when accessing a website. We concluded that:

1. Indonesian students were more concerned about information design and navigation design rather than visual design;
2. Regarding the visual design, Indonesian students more preferred a simple design and contrast in color combinations;
3. In the navigation design, Indonesian students need easy to use website especially to compare price and specification of product;
4. Trustworthy information about vendor is a factor that promotes consumer trust.

26.6 Future Research

Future study is recommended to provide an adaptive e-commerce for Indonesian student using characteristics as the results of the present study. Indonesian student characteristics and preferable website design should be modeled as user model and design model. Both models will be saved in a database such MySQL, MS SQL, or Oracle in order to easily modify the data and could be implemented into other website categories such as entertainment, news, or communication. Furthermore, website developers should provide an engine to communicate the models and website to identify the users. The engine could be provided with web programming language such as PHP Hypertext processor (PHP), Active Server Page (ASP), or Java Server Page (JSP).

Software testing is an important factor in the development of computer applications. Similarly, the communication engines between website and models that have been created also should be treated with internal and external tests. Internal test was done by developer teams, whereas external test was done in the participants of the present study. The purpose of tests is to find that the engines working correctly which indicates that the engines could identify users. The tests ignored inspections and only concerned with the output of computer applications. If the output of data processing is true, then the inspection of computer application is running well. The software for the tests is known as black box testing [17].

Acknowledgments We thanks to Universitas Teknologi Yogyakarta especially the students of faculty of information technology and business, Universitas Islam Negeri Sunan Kalijaga, Universitas Negeri Yogyakarta for their contributions to the present study. Their participation in the present study was appreciated.

References

1. Abran, A., Khelifi, A., Suryn, W.: Usability meanings and interpretations in ISO standards. *Softw. Qual. J.* **11**, 325–338 (2003)
2. Alpert, S., et al.: User attitudes regarding a user-adaptive eCommerce web site. *User Model. User-Adapt. Interact.* **13**(4), 373–396 (2003)
3. Al-Qaed, F., Sutcliffe, A.: Adaptive decision support system (ADSS) for B2C e-commerce in ICEC 2006, Fredericton, Canada (2006)
4. Butkiewicz, M., Madhyastha, H.V., Sekar, V.: Understanding website complexity: measurements, metrics, and implications. In: *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, pp. 313–328. ACM, Berlin, Germany (2011)
5. Creswell, J.W.: In: Smith, P.A. (ed.) *Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson, Boston (2012)
6. Cyr, D., Trevor-Smith, H.: Localization of web design: an empirical comparison of German, Japanese, and United States web site characteristics. *J. Am. Soc. Inf. Sci. Technol.* **55**(13), 1199–1208 (2004)
7. Dianne, C.Y.R.: Modeling web site design across cultures: relationships to trust, satisfaction, and e-loyalty. *J. Manag. Inf. Syst.* **24**, 47–72 (2008)
8. El Said, G.R., Galal-Edeen, G.H.: The role of culture in e-commerce use for the Egyptian consumers. *Bus. Process Manag. J.* **15**, 34–47 (2009)
9. Fang, X., Salvendy, G.: Customer-centered rules for design of e-commerce web sites. *Commun. ACM* **46**(12), 332–336 (2003)
10. Flavian, C., Guinalu, M., Gurrea, R.: The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Inf. Manag.* **43**(1), 1–14 (2006)
11. Flavian, C., Gurrea, R., Orus, C.: Web design: a key factor for the website success. *J. Syst. Inf. Technol.* **11**, 168–184 (2009)
12. Flesca, S., et al.: Mining user preferences, page content and usage to personalize website navigation. *World Wide Web* **8**(3), 317–345 (2005)
13. Garrigs, I., Gomez, J., Houben, G.-J.: Specification of personalization in web application design. *Inf. Softw. Technol.* **52**(9), 991–1010 (2010)
14. Helander, M.G., Khalid, H.M.: Modeling the customer in electronic commerce. *Appl. Ergon.* **31**(6), 609–619 (2000)
15. Ho, S.Y., Bodoff, D., Tam, K.Y.: Timing of adaptive web personalization and its effects on online consumer behavior. *Inf. Syst. Res.* **22**, 660–679 (2011)
16. Hsieh, H., Chen, C.-H., Hong, S.: Incorporating culture in website design: a comparison of Taiwanese and Australian website characteristics, In: Rau, P.L.P. (ed.) *Cross-Cultural Design. Cultural Differences in Everyday Life*, pp. 393–403. Springer, Berlin Heidelberg (2013)
17. Jangra, A., et al.: Exploring testing strategies. *Int. J. Inf. Technol. Knowl. Manag.* **4**, 297–299 (2011)
18. Kamoun, F., Halaweh, M.: User interface design and e-commerce security perception: an empirical study. *Int. J. E-Bus. Res.* **8**, 15–32 (2012)
19. Kim, J., Lee, S., Kim, S.: Understanding users in consumer electronics experience design. In: *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pp. 189–194. ACM: Montreal (2006)
20. Kuster, I., Vila, N.: Successful SME web design through consumer focus groups. *Int. J. Qual. Reliab. Manag.* **28**, 132–154 (2011)
21. Lee, S., Koubek, R.J.: The effects of usability and web design attributes on user preference for e-commerce web sites. *Comput. Ind.* **61**, 329–341 (2010)
22. Liang, T.-P., Chen, H.-Y., Turban, E.: Effect of personalization on the perceived usefulness of online customer services: a dual-core theory. In: *Proceedings of the 11th International Conference on Electronic Commerce*, pp. 279–288. ACM, Taipei (2009)
23. Liao, H., Proctor, R.W., Salvendy, G.: Chinese and US online consumers preferences for content of e-commerce websites: a survey. *Theor. Issues Ergon. Sci.* **10**, 1942 (2009)

24. Ling, C., Salvendy, G.: Importance of usability considerations to purchase intention on e-commerce website with different user groups. In: HUMAN FACTORS AND ERGONOMICS SOCIETY, SAGE (2006)
25. Mobasher, B., Cooley, R., Srivastava, J.: Creating adaptive web sites through usage-based clustering of URLs. In: Knowledge and Data Engineering Exchange, Chicago (1999)
26. Purwati, Y.: Standart features of e-commerce user interface for the web. *J. Arts Sci. Commer.* **2**, 77–87 (2011)
27. Rahimnia, F., Hassanzadeh, J.F.: The impact of website content dimension and e-trust on e-marketing effectiveness: the case of Iranian commercial saffron corporations. *Inf. Manag.* **50**(5), 240–247 (2013)
28. Rianto, S., Nugroho, L.E., Santosa, P.I.: Adaptive E-commerce using group personalization: a preliminary study on Indonesian students. In: 2nd Asia-Pacific Conference on Computer Aided System Engineering APCASE 2014. Bali Dynasty Resort, Bali, Indonesia (2014)
29. Singh, N., Pereira, A.: *The Culturally Customized Web Site*. Routledge, New York (2012)
30. Song, G., Salvendy, G.: A framework for reuse of user experience in web browsing. *Behav. Inf. Technol.* **22**, 79–90 (2003)
31. Tan, G.W., Wei, K.K.: An empirical study of web browsing behaviour: towards an effective website design. *Electron. Commer. Res. Appl.* **5**(4), 261–271 (2006)
32. Teevan, J., Dumais, S.T., Horvitz, E.: Personalizing search via automated analysis of interests and activities. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development In Information Retrieval, pp. 449–456. ACM, Salvador (2005)
33. Wang, Y.D., Emurian, H.H.: Trust in e-commerce: consideration of interface design factors1. *J. Electron. Commer. Organ.* **3**(4), 42–60 (2005)
34. Wen, H.J., Chen, H.G., Hwang, H.G.: E-commerce web site design: strategies and models. *Inf. Manag. Inf. Secur.* **9**, 5–12 (2001)

Chapter 27

Optimizing Financial Markets in C# .NET

Billy Leung, Zenon Chaczko and Jan Nikodem

Abstract This paper addresses the financial markets problem space with particular emphasis on trading systems seen in banking today. The paper starts off by describing the business problems in banking, demonstrating the undeniably significant contribution technology has provided to financial markets globally. With the advent of modern computing, cross-regional trading transactions can now be executed within milliseconds, a reality that is impossible without the advancements of software systems technology. We then move on and discuss the anatomy of a trading system and how it fits in with the banks ecosystem of vital inter-working components. The paper explores trading system concepts by walking through the development of a C# Foreign Exchange trading application using current popular technologies seen in major institutions today Leung, B.: Software architecture diagrams document. Front-Office Forex project at BJSS Ltd. (2011) [1]. The paper concludes with one of the most important discussions in engineering; designing quality financial software with accuracy, high availability, performance, high throughput and explores how they can be addressed by middleware, compute grid technologies, and simply by sound engineering design.

27.1 Background

Trading, one of the fundamental operations of banking has been a source of interests by financial institutions, physicist, mathematicians, engineers and numerous different disciples since the birth of trading. Trading progressed from the old age marketplace

B. Leung (✉) · Z. Chaczko
University of Technology, Sydney, Australia
e-mail: bleung.uts@gmail.com

Z. Chaczko
e-mail: zenon.chaczko@uts.edu.au

J. Nikodem
Wroclaw University of Technology, Wroclaw, Poland
e-mail: jan.nikodem@pwr.wroc.pl

meetings where traders come to trading pits shout out what they have on offer, and prospective buyers will cease the opportunity and make a purchase by signaling a corresponding buy hand signal [2]. In today's marketplace there are more than millions of transactions per second, prices of products change in the order of sub-milliseconds, and the amount of products traded from countries in disparate regions around the world well surpass the ability of a human being to consume. The reasons why computing and software systems are undeniably one of the main contributors to growth in banks and subsequently global economies is the fact that not only can we add more computing power to scale operations but on the other end of the scale, getting human traders together from different parts of the world to make cross regional trades and having them execute the trades, provide reports all in a matter of seconds is simply impossible.

This paper examines the thought process required to build trading systems to support such high volumes, high throughput, solve cross regional communications and takes a look at the business drivers which make these issues mission critical.

27.1.1 Objectives

The aim of this paper is to examine the business problem of front office trading systems within the banking ecosystem and by taking an engineering approach disseminate the complexities and issues seen in this extremely fast paced, highly reliant on accuracy and high availability environment. The approach utilized is purely pragmatic with the focus being to engineer a functional Foreign Exchange application in C# .NET using latest popular technologies seen in banking today [3]. By using FXTrader C# as a basis, the paper will uncover crunch points within the architecture of the system in which considerations of scalability—vertical and horizontal, security, low latency and high throughput becomes a serious non-functional requirement with potential direct business impact. The demo will be written using the latest Microsoft's .NET technologies WCF, WPF, C# .NET and Microsoft's Reactive Extensions (RX) technology to address real-time event streams (especially useful when dealing with streams of prices or trades) which has gained much attention in recent years in the banking industry. The outcome of this paper is a working Foreign Exchange C# blueprint (open source and available on sourceforge) with architectural critiques on considerations to potentially evolve this application to bring to market.

27.1.2 Evaluation of Effectiveness

It is important to recap the purpose of the paper and the method proposed in order to evaluate the effectiveness of the proposed method. The purpose of this paper is to explore aspects that are essential in trading software such as precision and accuracy to honor the financial impact of errors, scalability to honor organic growth and cross

regional distribution and high throughput to support the demands of financial markets. Built in to FXTrader is a scenario generation ability that allow the user to examine certain complex scenarios such as increasing/decreasing the demand for requests to simulate scalability, switching off threading optimizations to identify its consequent design time issues and a range of other parameters that the user can utilize creatively in order to reproduce real life scenarios in high pressured environments. Having these parameters help demonstrate the potential issues that can occur, and the latter of the paper will provide input on its impact to the business and suggest various solutions and technologies that can be utilized to address these problems. In essence the paper concludes by addressing outstanding issues with insight, research and discussion of the problem space.

27.2 Business Problems in Banking

Banking operations are diverse, this paper focuses on the fundamental business functions of front office banking i.e. trading systems, and how banks/corporations/wealthy individuals can trade with each other cross regionally, efficiently, and accurately. This translates to cross regional availability servers, high performance/low latency/high throughput, and quality of code in engineering. In addition, trading systems usually have service level agreements (SLAs) in place which translates to a set of non-functional software requirements in the software engineering world. A typical example of this is that the trading system must be able to execute a trade within one second from when the price of the product has been published. There are common misconceptions that engineers see non-functional requirements as a nice to have, in banking it is critical to the company and a lot of the time can incur a direct financial impact. These can be serious issues in the banking world and at times can impact the economy. The final section of this paper will discuss these issues in detail.

27.3 Workshop—C# FX Trading System

27.3.1 Overview

This workshop illustrates the complexities of engineering a Foreign Exchange (FX) trading application. The FX app will be written in C# .NET Visual Studio 2010 and aids as a tool for discussing financial applications in banking. This walk-through will follow the major SDLC process, requirements > architecture > design and follows with a discussion of the major issues involved in the trading application. The benefit of using FX trading applications as an example is that they require less savvy finance related knowledge which allows the focus to retain on the engineering aspects of trading systems. FX trading applications are commonly known for being a sound basis for furthering into more complex financial fields.

27.3.2 Scope

User has knowledge of:

- C# and Visual Studio 2010—Familiar with server side development
- Familiar with Service Oriented Architecture, Message Queues, High performance computing
- Basic knowledge of finance
- Basic WCF, WPF knowledge
- Basic knowledge of Microsoft Reactive Extensions (RX).

27.3.3 Business Problem

The FXTrader C# application is designed for the fictional FXWorld Bank organization. Focus on the FXWorld bank boundary and pay attention to firstly how clients interact with FXWorld Bank (in this case the client is Bilbo's bank) and how FXWorld sources currencies from the Wholesale FX Market to build an inventory (Fig. 27.1).

27.3.4 Business Requirements

Follow the sequence in the major use case from zero to two to understand firstly how commissions are preset by FXWorld, and how Bilbo's bank is presented with rates that it then executes against. Also, take note of what is happening behind the scenes with the Trader actor in FXWorld. Traders source currencies from the wholesale market and apply trader and sales commissions which determine the final client rate (Table 27.1).

Consequently, the business requirements we can draw for designing the FXTrader system are as follows.

27.3.5 Architecture

The main focus of the demo application will be with the FXTrader server component as this is where the majority of the crunch points in the system lie (as per the notes on the diagram above). This will be explained in the final part of the document in detail.

The FXTrader GUI and other bank services will be discussed briefly but will not be included in the design of this demo. With these points in mind, the consequent architectural logical view is depicted below (Fig. 27.2).

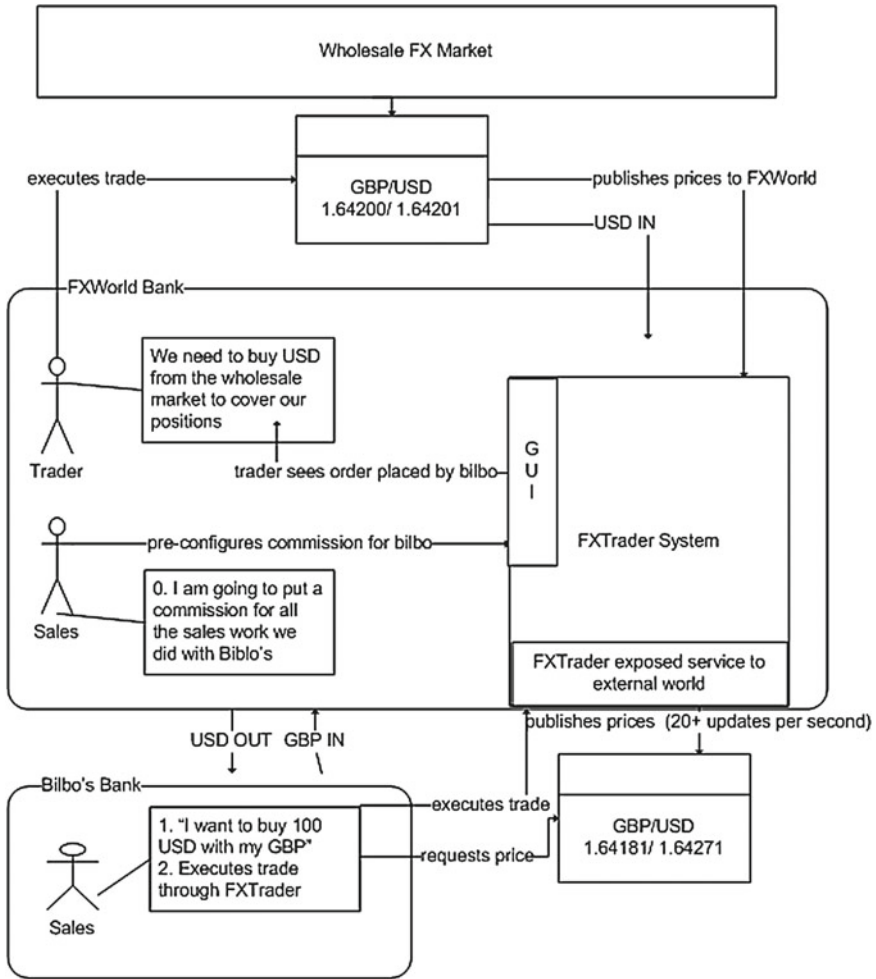


Fig. 27.1 The FX business problem

The wholesale market price pump will play the form of ExternalSimulator. Price-Generator and Bilbo's eTrader GUI as is are important aspects in the design of the demo. These components are crucial input/exit points in and out of FXTrader. The demo shall provide configurable parameters to help demonstrate the effects of various scenarios e.g. faster price pumps and increased price and trade volumes for external clients.

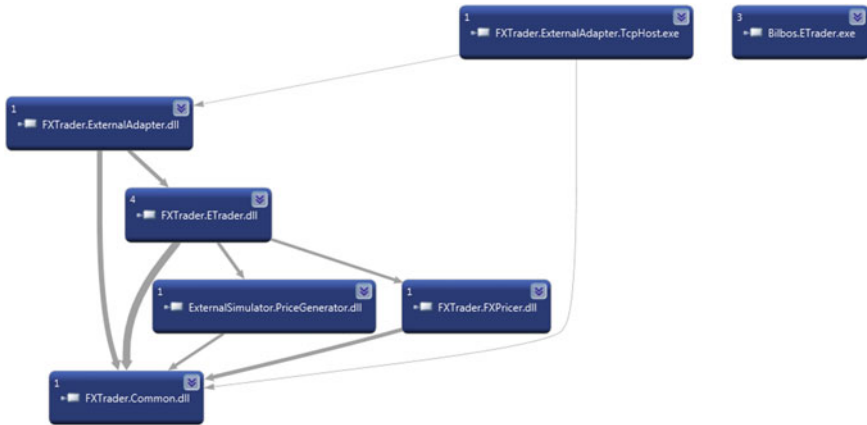


Fig. 27.3 FXTrader physical view

27.3.6 Physical View

The below diagram depicts the physical deployment of the FXTrader application. This is supplied to give an idea of the binaries involved to run the application, the main being FXTrader.ExternalAdapter.TcpHost; the WCF server host executable and Bilbos.ETrader the WPF/WCF fat client GUI to see rates and execute trades (Fig. 27.3).

27.3.7 Modeling FXTrader

The below diagram depicts a high level design of FXTrader and the responsibilities of each of these components.

This section will make technology decisions on the design of FXTrader to meet the business requirements stated in the previous section and the unwritten non-functional requirements in trading systems which will be discussed in future sections. It is important to take note of the rapid price feeds, rapid responses sent to many client connections and the ability to derive statistics, FX pricing, the non-functional requirements of scalability and high throughput.

27.3.8 Modeling Real-Time Events Using RX.NET

Microsoft’s Reactive Extensions (RX) is .NET’s event model for asynchronous programming and is a perfect choice for modeling streams of price events, trade events, in a high performing manner. RX’s basic concept of Observable<T> stream is a powerful way to represent a series of events.

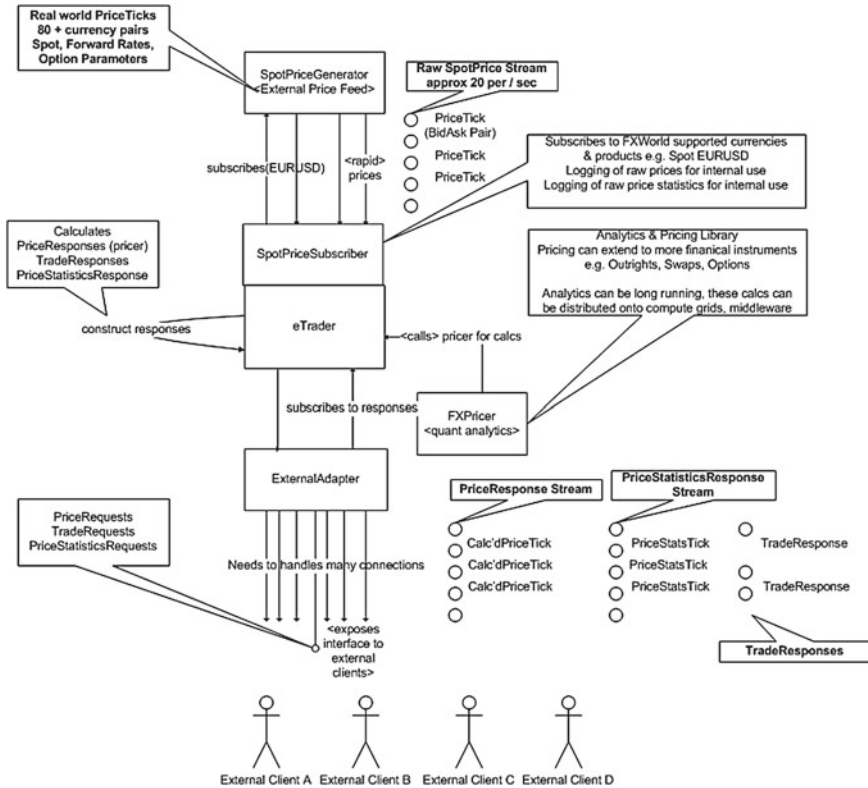


Fig. 27.4 High level design

In FXTrader’s case, we have an Observable<PriceTick> which can be subscribed to, aggregated, projected and handed out throughout the system. Incidentally, this allows us to address the core functional requirements of this project:

- Logging the raw spot price stream
- Aggregating the raw spot prices to derive statistics and logging the raw price stats
- Calculating the client price using FXPricer on each price in the raw spot price stream i.e. generating a stream of SpreadedPriceResponse to send to the client
- Aggregating the SpreadedPriceResponse stream to derive statistics for the client i.e. generating a stream of PriceStatisticsResponse to send to the client.

Referring to the high level design diagram on Fig. 27.4, we can see the crunch points of the system. The areas where there requires ultra fast streams of data, high throughput, lie in the boundaries of the system where the stream of updates come through. Namely like raw spot price stream, and the PriceResponse/TradeResponse stream sent back to the clients (Fig. 27.5).

The figure above depicts a RX marble diagram and demonstrates that we can our issues solve and model the hardest problems entirely using Reactive Extensions.

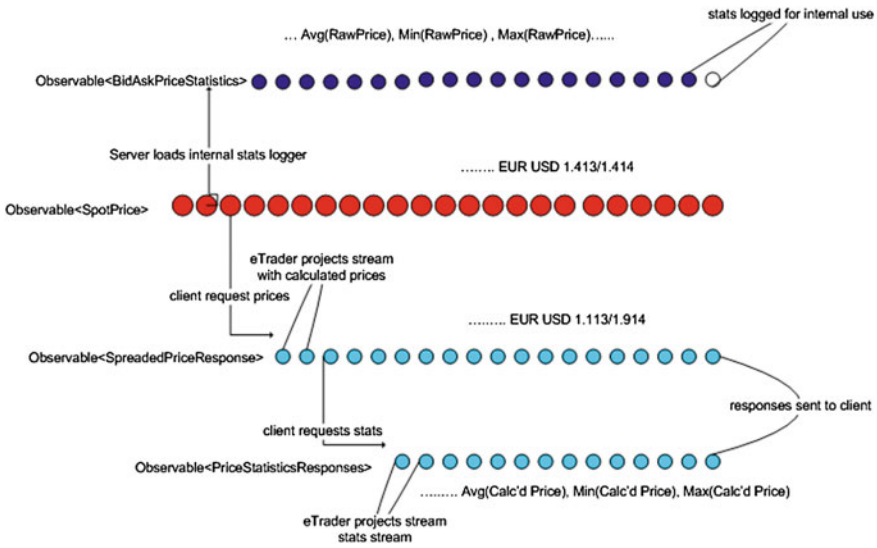


Fig. 27.5 Marble diagram of the streams within FXTrader

As described in the diagram above the main source of prices is the raw price spot stream Observable<SpotPrice> in which originates from the price pump SpotPrice-Generator. Observable<SpotPrice> projects itself into three separate real-time event streams;

- Observable<SpreadedSpotPrice> the calculated client price computed by the quantitative library. This is the end client price that clients can subscribe to.
- Observable<PriceStatisticResponse> the real-time statistics based on the client price average, maximum, minimum. Clients can subscribe to this feed to receive real-time statistics on the offer prices.
- Observable<BidAskPriceStatistics> the real-time statistics based on the raw price coming from the wholesale market. Financial analysts within Bilbo’s bank can analyse these prices to manage risk and assess whether the prices of FX products are appropriate for resale.

For example when a client requests a price, the source SpotPrice stream projects into a SpreadedPriceResponse stream using the FXPricer’s calculate method to transform the stream; the responses are consequently streamed to the client (Fig. 27.6).

27.3.9 The contract between FXWorld and External clients

WCF is the chosen distributed technology for this demo, using TCP hosting end-points provide fastest client server interactions. In reality, FX systems commonly use a TCP variant over the FIX (Financial Interchange Protocol). However, this design can act as a good starting point and also for smaller FX systems it may suffice.

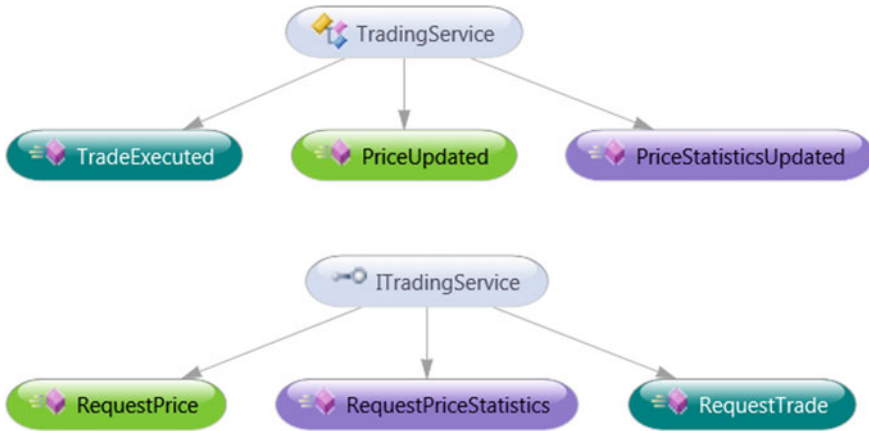


Fig. 27.6 WCF contract

The WCF contract between FXWorld’s FXtrader ExternalAdapter component and Bilbo’s Banks’ eTrader component is shown in the above diagram. It exposes simple request/response methods for prices, statistics, and trades.

27.3.10 External Client’s GUI—Bilbo’s eTrader

The external Client GUI not only acts as a trading interface for Bilbo’s but also acts as a simulation tool to help visualize the real-time pricing of pricing event streams (Fig. 27.7).

27.3.11 Bank Ecosystem in Reality

In reality the technology systems in a bank would most likely look more like the figure below, paraphrased from [1]. Most banks will be using a service oriented architecture, and have separate systems that handle configuration/static data, a separate service that handles market prices for the bank as a whole that manages all business areas, risk management systems which helps traders assess risk and devise strategies, a trade management system that holds trade execution information that connect to middle office systems for audit and settlement reasons (Fig. 27.8).

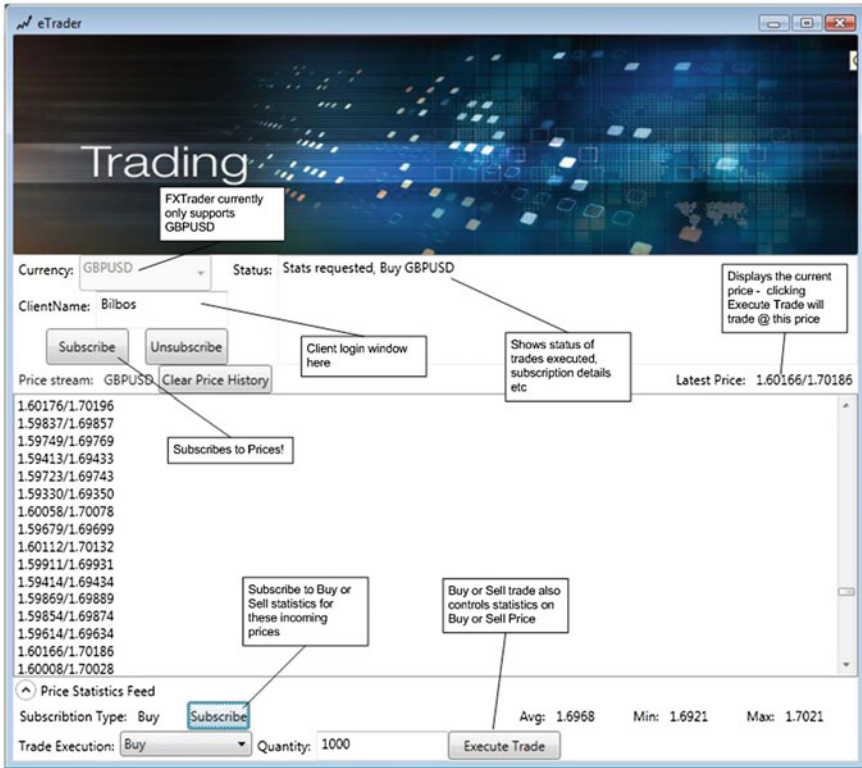


Fig. 27.7 Bilbo's eTrader GUI

27.4 Discussion of Crucial Issues in Financial Systems

27.4.1 SLAs, NFRs, QoS

All trading systems will have some form of SLA agreements to meet quality expectations of trading in the market.

For example, FX pricing must meet strict regulatory requirements to comply with industry regulations to eliminate the risk that market players over/under price which is a legal violation in the form of market manipulation. This translates technically into correctness.

There are many legal regulations policed by government bodies such as the Financial Services Authority (FSA) in the UK and the Securities and Exchange Commission (SEC) in the U.S. Top tier investment banks normally have a representative group that polices the abidance of these laws. Non compliance is a serious legal issue and implies direct financial and reputation impact.

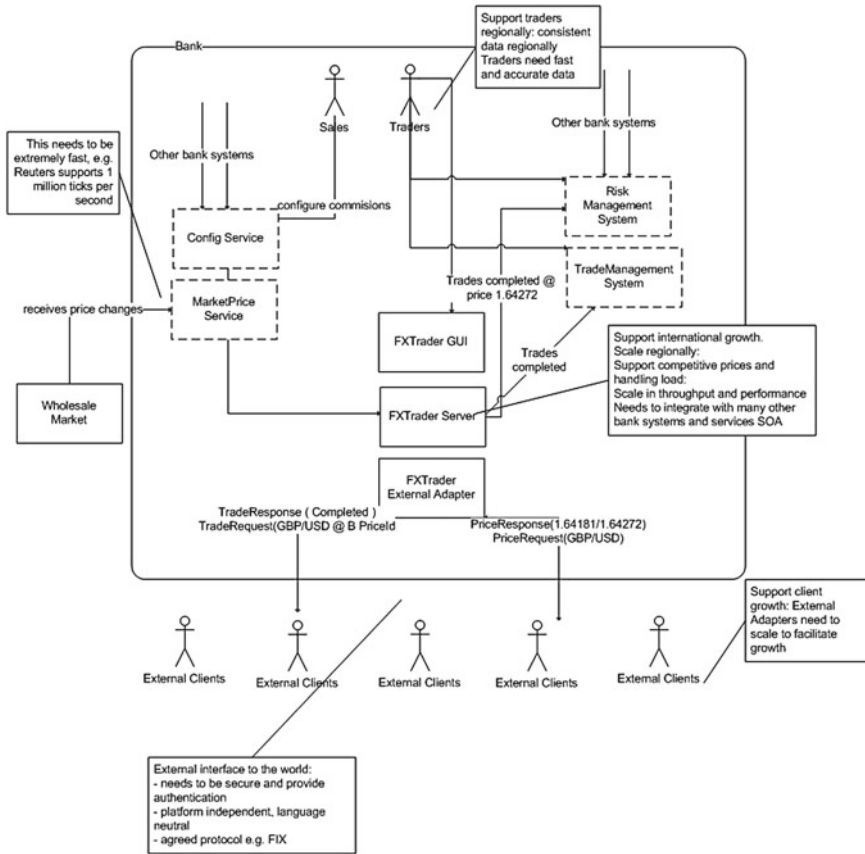


Fig. 27.8 Bank ecosystem

To the inexperienced eye, engineers maybe set to believe that this is purely about the calculation correctness of Quantitative Libraries. Correctness of Quantitative libraries play a significant role in the correctness of pricing, however the next section surprisingly uncovers how performance, scalability and throughput in real-time trading systems can have a direct impact on the correctness of prices as well.

In an engineering point of view, the business enforced SLAs are viewed as QoS requirements on the distributed networking nature of the system and consequently translated to NFRs in software systems.

27.4.2 Running Scenarios Using FXTrader

To fully appreciate the future sections, please download the FXTrader source code available online <http://sourceforge.net/projects/fxtrader/>.

After you have downloaded the package, extract the latest binaries FXTrader_x.x.zip and navigate to the 'bin' sub folder. Follow the steps below to start up the FXTrader application

To simulate a scenario, you must perform step 1 to modify the server configuration.

1. [Key Step] Modify server\FXTrader.ExternalAdapter.TcpHost.exe.config to the values specified in each scenario, configuration changes are enclosed under the <appSettings> xml element.
2. Launch the server—execute server\FXTrader.ExternalAdapter.TcpHost.exe; this will start a WCF TCP host for the FXTrader application, review the Physical View and Architecture section of this document for more details.
3. Launch the client by executing Bilbos.ETrader.exe; this will connect via WCF to the Server.
4. Click Subscribe on the client GUI and you should see rates coming through the console; whilst rates are being received you can utilize the other functionality of the GUI such as Execute Trade, Price Statistics Feed → Subscribe.
5. Analysis of server interactions is driven from tailing the server log file bin ServerTradingService.txt. Load the BareTail tool and tail this log to observe server interactions.
6. You also have access to tail the client logs eTrader.txt and have complete access to the source code in the root directory.

27.4.3 Scenarios

In light of all the theory discussed about enforcing non-functional qualities, this section demonstrates these concepts using the FXTrader application as a vehicle to help visualize these problems in reality. These scenarios are extracted from real life experiences in FX Trading systems [1].

27.4.3.1 Scenario 1: Consider the Threading Model in RX

This scenario demonstrates how it is crucial to understand the multi-threading constructs used in real-time trading applications. As threading analysis techniques in .NET is widely available, this scenario looks at the threading constructs in RX when prices are requested and responses sent back to the client (Fig. 27.9).

Modify server configuration and restart the server;
DisableThreadingConstructs = true.

```

eTrader.txt  TradingService.txt
[INFO ] [14-02-16 04:45:15,608] [1] [Program] -- begin config --
=====
Simulation:
=====
PriceTickIntervalInMilliseconds=1
PriceLongOperationTimeInMilliseconds=-1
TradeLongOperationTimeInMilliseconds=-1
DisableThreadingConstructs=True

```

Fig. 27.9 Disable threading constructs configuration

```

public void RequestPrice(PriceRequest request)
{
    Logger.InfoFormat("PriceRequest {0} - {1}, client = {2}.", request.ProductType, request.CurrencyPair, request.ClientName)
    _eTrader.AuthorizeClient(request.ClientName);

    SetupCallbackChannel(request.ClientName);

    Logger.InfoFormat("PriceStream channel setup for {0}, streaming {1} - {2} to client.", request.ClientName, request.ProductType, request.CurrencyPair);

    IDisposable disposable;
    var primaryStream = eTrader.GetPriceResponseStream(request);
    if (Constants.DisableThreadingConstructs)
    {
        disposable = primaryStream
            .Subscribe(priceResponse => PriceUpdated(priceResponse),
                e => Logger.WarnFormat("PriceResponse stream error, Reason: '{0}'", e.Message));
    }
    #region Threading Considerations
    else
    {
        disposable = primaryStream
            .ObserveOn(NewThreadScheduler.Default)
            .Subscribe(
                class System.Reactive.Concurrency.NewThreadScheduler
                Represents an object that schedules each unit of work on a separate thread. '{0}''", e.Message));
    }
    #endregion
    _clientSessions.AddStream(request.ClientName, disposable, primaryStream);
}

```

Fig. 27.10 Disable threading construct coding issues

With threading constructs turned off, prices arrive incorrectly to the client affecting *price correctness*; the impact can be detrimental and has a direct financial impact to the company.

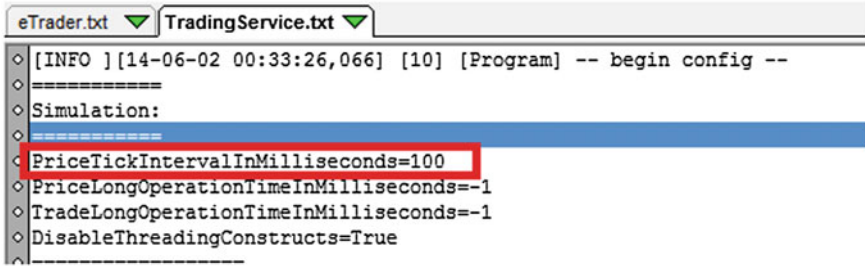
The below screenshot identifies the offending code; the erroneous code branch is highlighted and the correct code is in the else branch (Fig. 27.10).

With disabling constructs off, `ObserveOn` will default to sending `PriceUpdated` notifications to observers (i.e. the client) on the current thread. The result is the server and client will seem to hang as the notification to observers is being executed so quickly every 100ms and the system will seem to block [4].

The net effect of this is the client will not be receiving prices on time, as a non-functional requirement is that end users should be able to see prices every 100ms.

The fix for this is to spawn a separate thread when the server sends notifications to the client by using the `ObserveOn(NewThreadScheduler)` construct as above. This will stop the server from blocking and will correctly send prices back to the client.

This scenario demonstrates correctness and performance.



```
eTrader.txt TradingService.txt
[INFO ] [14-06-02 00:33:26,066] [10] [Program] -- begin config --
=====
Simulation:
=====
PriceTickIntervalInMilliseconds=100
PriceLongOperationTimeInMilliseconds=-1
TradeLongOperationTimeInMilliseconds=-1
DisableThreadingConstructs=True
```

Fig. 27.11 Simulate faster price ticks

27.4.3.2 Scenario 2: Simulate Faster Price Ticks

This simulation allows the engineer to simulate when the external party in this case Wholesale Market pumps prices quicker than the system can handle (Fig. 27.11).

Modify server configuration and restart the server;
PriceTickIntervalInMilliseconds = 100

What usually happens in this scenario is engineers will work with the external party and agree on an SLA on how fast prices will be received from the wholesale market business. This in turn translates to a non functional requirement in which FXTrader shall support price ticks at 100 ms.

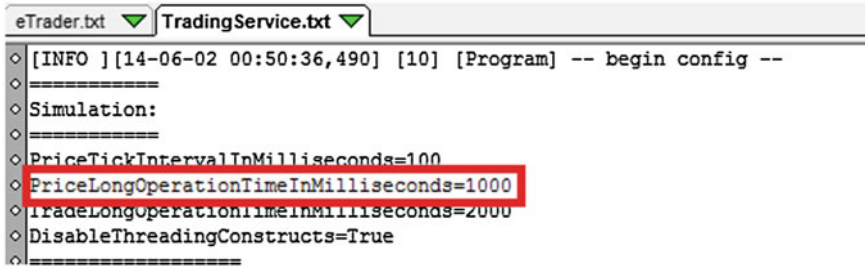
This configuration value allows the engineers to conduct performance testing of the system to identify what it can cope with or areas in the design it can optimize to allow faster throughput.

27.4.3.3 Scenario 3: Simulate Slower Pricing Execution—Dependency on Quant Libraries and Distributed Computing Systems

The complexity of quant calculations and its throughput normally merits a separate software division. In the scope of FXTrader, a method call from a referenced API library is called to perform the calculations.

Complex pricing may take a long time to compute and many banks will utilize grid computing or cloud computing to leverage computation power in a distributed fashion. In order to ensure FXTrader operates correctly when pricing calculations are slow, this simulation provides engineers with the ability to modify the calculation times and observe the affects (Fig. 27.12).

Modify server configuration and restart the server;
PriceLongOperationTimeInMilliseconds = 1,000.

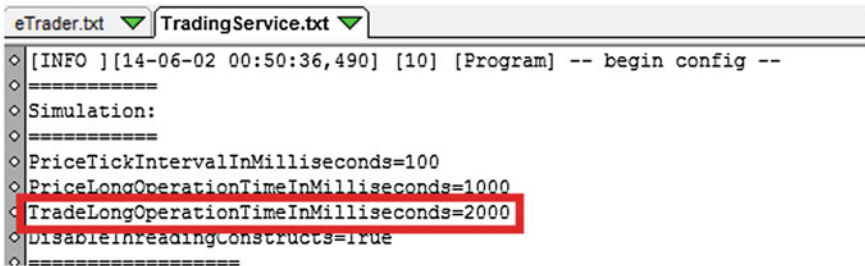


```

eTrader.txt ▼ TradingService.txt ▼
[INFO ] [14-06-02 00:50:36,490] [10] [Program] -- begin config --
=====
Simulation:
=====
PriceTickIntervalInMilliseconds=100
PriceLongOperationTimeInMilliseconds=1000
TradeLongOperationTimeInMilliseconds=2000
DisableThreadingConstructs=True
=====

```

Fig. 27.12 Slower pricing execution configuration



```

eTrader.txt ▼ TradingService.txt ▼
[INFO ] [14-06-02 00:50:36,490] [10] [Program] -- begin config --
=====
Simulation:
=====
PriceTickIntervalInMilliseconds=100
PriceLongOperationTimeInMilliseconds=1000
TradeLongOperationTimeInMilliseconds=2000
DisableThreadingConstructs=True
=====

```

Fig. 27.13 Slower trade execution configuration

27.4.3.4 Scenario 4: Simulate Slower Trade Execution—Dependency on Trade Capture Systems

Due to the organizational structure of a bank, trade capture is considered a middle office function so this is normally handled by a separate division. Trade capture systems capture execution of trades so that it can be followed up by settlement and legal departments (Fig. 27.13).

Modify server configuration and restart the server;
 TradeLongOperationTimeInMilliseconds = 2,000

27.4.3.5 Scenario 5: Logging Affects Performance and Price Correctness

System logging can affect performance and consequently not meet NFRs of sending prices to clients on time.

Logging can be enabled or disabled and logging intervals can be modified to see the effects of logging. When releasing FXTrader to production, the benefit versus cost of logging should be considered carefully.

Modify server configuration and restart the server;
 LogRawPriceStatistics
 RawPriceStatisticsReadIntervalInMilliseconds

```

LogRawSpotPriceLogging
RawSpotPriceLoggingThrottleInMilliseconds
LogPriceResponse
LogPriceStatisticsResponse.

```

27.4.3.6 Scenario 6: Real Life Case Study—Off Market Pricing

This is an interesting engineering problem that has direct financial consequence on banks with real-time trading systems as it incurs a financial loss. With the advent of real-time trading systems, algorithmic trading (i.e. computer traders as opposed to human traders) companies tend to leverage the fact certain trading systems are not on par with their throughput and ability to handle the scale of operations.

The scenario can be explained as follows:

1. FXTrader is publishing prices to Bilbo's bank at 1 price tick/100 ms.
2. A trader in Bilbo's bank executes a trade at a certain price.
3. FXTrader processes the trade and acknowledges Bilbo's banks trade execution.

There is non-functional requirement embedded at point c. The processing of the trade must complete quickly so that the market has not moved outside the favor of the bank when transaction has been executed.

In this scenario, monitoring what Bilbo's sees is important, hence we can examine the client logs (Fig. 27.14).

In this example, a trade was executed with a notional amount of 1 million at 5.69227, the real market price dropped to 1.6296 between the time of booking the trade and replying with a confirmation. Thus, the slowness of the response to client **has caused the company to lose 4 million USD** in this hypothetical example. This example is purely hypothetical to demonstrate the direct financial impact of this subtle non-functional requirement, in reality these major currencies will unlikely fluctuate so much.

The importance of this example is for engineers to focus on the performance of FXTrader between TradeExecuteRequest and TradeExecuteResponse. Remember between these two states FXTrader needs to accept the request, process it which includes communicating with other banking services and send a confirmation response promptly to the client.

N.B. Modifying parameters in Scenario 4: Simulate slower trade execution can slow down processing and help illustrate this issue.

```

[14-04-02 01:26:51,686] PriceReceived: ccyPair=GBPUSD, Price=1.59207/5.69227
[14-04-02 01:26:51,723] TradeExecuteRequest Buy 1,000,000 @ 5.69227
-- Market prices move between TradeExecution and TradeResponse - if price moves under 5.69227
[14-04-02 01:26:51,810] PriceReceived: ccyPair=GBPUSD, Price=1.59276/1.6296
-- Price has dropped to 1.6296 --
[14-04-02 01:26:51,928] TradeExecuteResponse Latency=0.241 ms
-- Trade accepted at 5.69227 --
|

```

Fig. 27.14 Off market pricing scenario

FXTrader demonstrates a confined example and in this case had a latency of 200ms. In reality, banks have many inter-networked services; suffer from network latency and the many complex computing issues results in response times of three to five seconds [1]. With slow response times as such, the market is likely to have moved in the unfortunate case against the banks favour.

The next section discusses important trading system considerations in attempt to describe the complexities of trading systems.

27.4.4 Important Considerations

This section outlines the complexities of trading systems by listing important trading qualities considerations.

27.4.4.1 Correctness

Ensuring prices are published correctly is governed by financial authorities. Financial Services Authority (FSA) in the UK and the Securities Commission (SEC) in the US would normally have a representative group within banks to police the correctness of prices. Compliance can intervene in cases of mispricing and in the worst case can trigger legal implications.

- Pricing and analytics
- Timing correctness affects correctness
- Ensure price sequencing is correct as prices are time dependent.

27.4.4.2 Throughput

Throughput of a trading system as we have seen in 4.3 Scenario has adverse implications sometimes leading to financial damage. The below ideas can help engineers validate the effects of throughput.

- Up the price tick rate coming in from wholesale market
- Up the statistics collection interval
- Disable logging
- Threading analysis can cause prices to come out of sequence or system to hang.

27.4.4.3 Performance

At a more technical level, the core performance of the software has a direct impact on the throughput of the system. The following list ideas on how to tackle this complex problem.

- Understand hardware and CPU cores the software runs on
- Add more eTrader, External adapters so they can be deployed on separate machines to get more memory/performance resources
- Code optimization by using software profiles such as Redgate's products
- Memory optimization have a significant impact to performance considering using ValueType versus Object types ensure efficient memory utilization
- Pricing Performance can be optimized in .NET by inter-operating with C++ quant libraries.

27.4.4.4 Support Growth

- Client on boarding—consider expanding the system with more client connections it is not uncommon to support 10–20 clients at a time. Running more Bilbo's eTrader instances can help the engineer simulate client growth.
- Higher volume trades—clients may increase their volume of trading according to business growth. This results in more requests issued on the server side.
- More product supports—Supporting more FX products such as derivatives may increase the event streams supplied to FXTrader with derivative information.
- More currency pairs—business growth into different currencies require support for more currency pairs, an amount of 80+ currencies is not uncommon.
- Pricing performance—Leveraging multithreading, grid computing etc. to support the complexities of pricing and growth of product types and currencies.

27.4.4.5 Scaling

- Regional Scaling—Growth can span cross-regionally, generally a bank will cater for the Asian market, European Market and US Market. Analyzing network latency between components becomes crucial here.
- Scaling for throughput—Deciding on either a horizontal or vertical scaling strategy will assist in throughput.

27.4.4.6 Other Unwritten Constraints

- Harder to deploy—the more the system grows the harder the system becomes to deploy, consider modern deployment strategies
- Outages are a catastrophe in a real-time trading environment—releases are commonly constrained to be executed outside trading hours
- Pass audit—the strictness of financial regulations may involve software releases to pass audits especially pricing correctness with securities committees.

27.4.5 Technologies

This section lists modern technologies used in banks today which provide features that address the issues in trading systems.

- Technologies for price ticks—middleware, Message Queues, Reuters RMDS
- Technologies for scaling pricing—Grid computing, Quant Libraries/Analytics
- Technologies for intrabank services—middleware, Message Queues, SoA REST-FUL Webservice, JMS JBOSS
- Message Queues—IBM Webshpere, RabbitMQ, 29West, Microsoft MQ.

27.5 Conclusion

This paper has successfully explored the nature of real-time front office trading systems. The complex problem of real-time ticking events are addressed by the usage of the modern Reactive Extensions .NET. The implementation of FXTrader has provided a vehicle to visualize these non-functional requirements helps simulate certain real life complex scenarios and provides engineers with basic tools and approach to examining these problems in a pragmatic manner. Lastly a thorough discussion of important considerations and suggested technologies completes the paper by providing descriptions on how to resolve these real life banking issues.

27.6 Acknowledgment

The author would like to acknowledge the Front Office FOREX team at BJSS for their guidance and support and special thanks to Ronnie for introducing the author to the power of RX and it's applications in real-time trading.

27.7 Source Code

FXTrader is an open source project written wholly by the author. The latest binaries and source code is available for download on sourceforge <http://sourceforge.net/projects/fxtrader/>. Feel free to leave comments and feedback via the sourceforge feedback page.

References

1. Leung, B.: Software Architecture Diagrams Document. Front-Office Forex project at BJSS Ltd. (2011)
2. Daniel's Trading Editorial Team: Hand Signals. <http://www.danielstrading.com/education/futures-options-101/hand-signals/> (2014). Viewed 15 Jan 2014
3. Weinstein, S.: Usage of Reactive Extensions in the "Real World". <http://stackoverflow.com/questions/5607820/usage-of-reactive-extensions-in-the-real-world> (2012). Viewed 1 Jan 2014
4. Microsoft Developer Network. [http://msdn.microsoft.com/en-us/library/hh242963\(v=vs.103\).aspx](http://msdn.microsoft.com/en-us/library/hh242963(v=vs.103).aspx). Viewed 10 Apr 2014

Chapter 28

The Model of Customer Trust for Internet Banking Adoption

Shidrokh Goudarzi, Wan H. Hassan, Mir Ali Rezazadeh Bae and S.A. Soleymani

Abstract The use of the Internet has increased dramatically over recent years and is now regarded as the best channel for distribution of products and services of various types of businesses, such as internet banking services. This paper extends an area of information systems research into a financial services context by looking into the element of trust in Internet banking. As more financial institutions are currently seeking ways to boost Internet banking adoption rates, trust is also being examined as a significant issue in the relationship. This can be attributed to the fact that bank customers are concerned about the security involved in processing such sensitive material as financial information. Moreover, significant factors of trust are involved and these include: accessibility, privacy, security, quality, usability, users knowledge and disposition to trust. These can all have an impact on customer trust in adopting internet banking. Based on previous models with aforementioned variables that are theoretically justified as having an influence on trust, a relevant research model was developed to test eight (8) hypothesized paths among the study variables. These include, namely: accessibility, privacy, security, quality, usability, users knowledge, disposition to trust, trust, as well as the rate of internet banking adoption. Data was collected by survey questionnaires from a sample of 150 internet banking users. The Smart PLS tool was used for data analysis. The results of the data analysis generally support the model, as well as all of the proposed hypotheses. In summary, the results of this research have shown that accessibility, privacy, security, quality, usability, users knowledge and disposition to trust were found to have significant influence

S. Goudarzi (✉) · W.H. Hassan
Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia,
54100 Kuala Lumpur, Malaysia
e-mail: shidrokhgoudarzi@gmail.com

W.H. Hassan
e-mail: wanhaslina@ic.utm.my

M.A.R. Bae · S.A. Soleymani
Faculty of Computing, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
e-mail: don.of.delta@gmail.com

S.A. Soleymani
e-mail: ahmad.soleymani@gmail.com

on customer trust. Trust, in turn, was found to be an important factor in fostering a positive attitude toward adopting the services. Several implications for both research and practice have emerged and are discussed later in this study.

Keywords Trust · Internet banking · Adoption · Smart PLS · Customer

28.1 Introduction

Internet banking has become the most significant information technology used by most of the customers worldwide with the promise of greatly improving operational efficiency and enhancing financial performance. The usage of the internet for conducting all financial transactions can have positive results for the banking system. Unfortunately, the low range of adoption of technology in some developing countries can be the consequences of poor economic circumstances, inadequate education and lack of convenient infrastructure. There are probably also other reasons such as trust that play an important role in this regard. It is a tangible fact that the bank has expended a great amount of money in carrying out internet banking reforms. The main goal of the bank is to increase the range of use of internet banking to the highest level. Hence, the usage of internet banking has become a significant issue for the banking industry. In other words, lack of adoption in internet banking can result in negative consequences for customers and banks, since customers utilize internet banking at a minimum rate whilst still requiring access to information for their activities. Some researchers have shown that lack of trust can be the most significant long-term barrier to recognizing all of the potential benefits of internet banking. It indicates that one of the problems is the lack of an effective relationship between technology trust and e-banking adoption. Furthermore, data and network security, as well as privacy issues are other problems. Haque et al. [21] stated that customers are wary of using internet banking and purchasing through the internet as they think that any mistake or error could lead to losing money. Banks should make their customers more aware of their service quality attributes, the firm guarantee about e-banking security and the regulations governing e-banking. Lack of internet awareness is one of the most important problems associated with customers of banks, due to the fact that this service is still not widely accepted. It is clear that customers are still not completely confident with using ATM cards or telephone banking. The high level of awareness, as well as public education and public training could demonstrate the benefits of using new systems and could thereby encourage them to adopt internet banking transactions, since customer knowledge plays an important role in increasing the level of customers trust. The failure in the adoption of internet banking can be the result of different problems. Moreover, these problems can sometimes lead to rejection of the system. These problems relate to technical, social and organizational factors. Hence, in order to prevent the failure of internet banking adoption, researchers are studying important issues such as trust [23].

Sanayei and Noroozi [36] generated Internet Banking Adoption Model. In this study Perceived usefulness and perceived ease of use have an influence on intentions to transact positively, but internet banking security needs customers trust before individuals start identifying sensitive information and transferring that information to their bank through the Internet. Thus trust act as a rather universal capable for internet banking. Suh and Han [40] said that: Trust is the groundwork of commerce. It performs an important role in increasing customers trust in most sectors [20] such as the financial system [46]. According to [20], providing and maintaining trust-based connection between banks and customers is being identified as an extremely important issue in the financial industry. According to General Trust Perception Model (GTPM), the level of perception of trust among users is recognized. Base of the models for the perception of trust from past research, they could offer a model for the perception of trust generally that includes a large set of factors namely (GTPM). This model is very useful in various types in different situations. In this model all of the factors have their weights also weight of each factor is related to its application. Furthermore, weight of factors is not fixed because the level of awareness of Information Technology is different between customers [9].

This research aims at responding to a question on developing the trust model for using internet banking, namely: How can a trust model for Internet Banking service be developed? For this purpose, we aim to identify and describe the efficient factors relating to customer trust in internet banking. Also, we try to clarify the effect of trust on the adoption of internet banking. Moreover, we develop a model of trust towards increasing adoption of internet banking. To continue, we evaluate and validate the trust model in internet banking.

This study utilizes the new model on trust and internet banking adoption. Most effective factors that have highest effects on trust are selected. Also, trust as an important factor can determine the level of adoption to Internet banking among customers. A review of the literature on different aspects of the generating of adopting models for Internet banking shows a number of gaps in our understanding. These gaps fall into the following categories.

The users of a website can act electronic activates, because features such as quality of the website, security, privacy, reliability, usability are the main factors in internet banking that can affect on trust, because one aspect in increasing the level of adoption in internet banking is related to properties of the website.

On the other hand, users play an important role in increasing the level of adoption in internet banking. Because the properties of users such as the disposition to trust and users knowledge are the main factors that can effect on trust to internet banking activities. In the other words, these factors relate to the customers for doing internet banking activities.

The rest of the paper is organized as follows. Section 28.2 presents the background of different studies that have been presented and published and parameters used to research model for Internet banking adoption. Section 28.3 deals with current research methodology. The anticipated research results are illustrated in Sect. 28.4. Finally, Sect. 28.5 concludes our study.

28.2 Theoretical Background and Research Model

Figure 28.1 represented our research model. Each element of trust was assumed to affect the Internet banking adoption positively. Additionally, trust was assumed to moderate the relationships among each component of trust and the Internet banking adoption.

28.2.1 Trust and Internet Banking Adoption

Many studies have demonstrated that the significant relationship exists between trust and online banking or any e-commerce adoption. We can claim that trust happens when one party has confidence in an exchange integrity and reliability with her/his partner. For instant, Chen and Barnes [3] realized that trust substantially plays an important role on online purchasing intent, web site loyalty [12], internet banking commitment [29], electronic banking adoption [33] and behaviour intention in adopting online information service. Yousafzai et al. [46] concluded that trust in online banking and infrastructure of internet banking can cause reduction in users transaction-specific uncertainty. When individuals trust others, they accept that others will behave as they are expected and it can affect in reducing the complexity of the interactions. Studies of online banking [24, 25, 29] have shown that trust is a determining factor in making online banking processes. The uncertainty that an individual often can accept trust as a necessary component [19, 32]. Otherwise the consumer is antipathetic to usage internet banking services [24, 29].

Since of the importance of trust in Internet banking, customer trust is a significant component impacting the growth of Online banking [40]. Trust on the internet can gain through long-term internet usage. Moreover trust has been founded as an important factor in the adoption decision [14]. According to the [24], they could develop

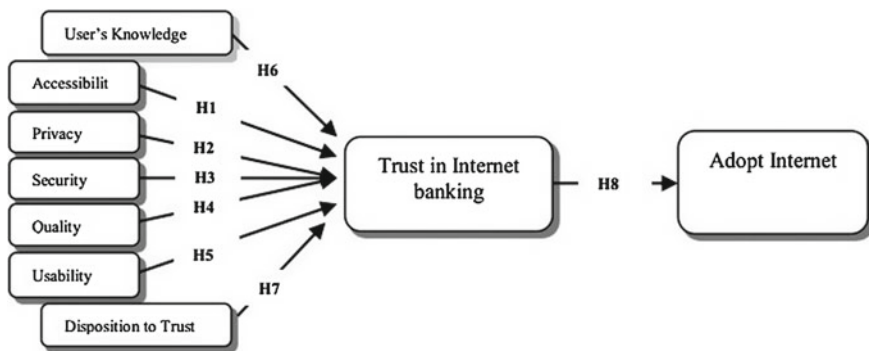


Fig. 28.1 Research model

a framework linking trust and attraction to customer commitment. Their framework has two main parts. First, it analyses the important effects of trust and attraction on commitment. Second, it analyses the effect of shared values, communication and opportunistic behavior of trust and communication on attraction. Mukherjee and Nath [29] view the customers tendency towards e-commerce and the level of usage of e-commerce technology is the consequences that they trust the electronic system as an empower for their trust in online banking. Also, they indicated that lack of experience may lead to concern about using the internet or may lead to avoidance of usage the internet and hence to a lack of trust. Mwesigwa and Nkundabanyanga [30] state that organizations acting business online must boost trust rapidly in order to succeed.

Papers [24, 29] realized that trust plays an important role in expanding and keeping successful relationships in the financial tasks sector because many of the activities are complicated and there is an online relationship between the bank advisor and the consumer. Transactions can be usually completed by these technologies and it is not necessary that the parties meet each other and directly facing each other. The parties will be concerned that their personal information and details and money will be moved to a third party without their awareness [44]. Clients behavior towards online banking is motivated through trust, which performs an essential part in boosting availability within the online banking environment. The concept of trust is more essential in online banking in contrast with the offline banking because transactions and the parties involved in the financial activities are worried about access to essential files and information transferred through the Internet [1, 40]. The role of trust in the expansion and keeping of effective connections is likely to be of particular importance in the financial activities sector because of the complication of many of the items. The level of customers trust in the internet banking will be negatively influenced by the belief that he/she is controlling in a high degree of risk even though the risk level may be basically low (perceived risk). The creation of trust in a relationship is a type of assurance versus risks and unexpected behavior.

28.2.2 Factors of Trust in Internet Banking Adoption

There are many factors identified by the past studies which will affect customers trust in online banking vendors. The all factors that can affect on trust are divided in two groups. One group is related to the characteristics of the website and another one is related to the characteristics of the customers [9]. We can realize that some of the listed factors refer to properties of the website (e.g. Privacy, security), while others refer to properties of the user (e.g. Users knowledge and disposition to trust). In this study one of the objectives is evaluating the level of trust among users. According to the models for the perception of trust from past research, they could offer a model for the perception of trust generally that includes a large set of factors namely (GTPM) by [9].

This model is very useful in various types in different situations. This study aims to collect two sets of important factors that can affect on trust in internet banking. One set relates to the most effective factors of a website and another set is related to properties of customers.

Based on the results, the following factors were observed, namely: accessibility, privacy, security, quality, usability, users knowledge and disposition to trust have significant effects upon trust. Accordingly, trust can be considered as a critical importance factor relating to the adoption of Internet banking. The concept of this study is relatively new because the selected factors of trust are related to the characteristics of customer and website.

Accessibility of computer, fast speed relationship via Internet and their basic knowledge are requirements to access the internet banking services, it is one of the important factors regarding the internet banking adoption [35]. The adoption of internet banking is not achievable without appropriate and direct accessibility of the internet. Lichtenstein and Williamson [27] conducted a research on adoption of internet banking and they could find that so many users in their research oriented to the high level of internet's accessibility in doing all their banking activities by connecting to the internet all day readily available. They mentioned that the level of accessibility can effect in choosing their banking channel. The perceptions of adequate internet accessibility can be highly important and it is necessary for customers that can use of internet banking with dedicated and unchallenged access [27].

Hypothesis 1: Accessibility is positively influences trust towards the use of Internet banking.

Yousafzai et al. described privacy as users perception relating to their capability to monitor, observe and control the collection, use, disclosure, and following accessibility of their information provided to the bank during an online financial transaction [46]. The consumer has concerns related to the significant matters in online activities such as internet security, privacy and trust. Also in many researches privacy have been indicated by (e.g. [15, 31]). The privacy as a significant factor can have an effect on consumer internet banking choices. It was found that 80% of Phishing attacks, luring of an internet user to reveal personal details (like passwords and credit card information), disturb the financial services sector globally in 2005 (IDC 2005).

Hypothesis 2: Privacy is positively influences trust towards the use of Internet banking.

Security refers to the safety of information or systems from unexpected inflows or outflows. The concern of the lack of security is one of the factors that has been found in most studies as impacting the growth and improvement of e-commerce. Users are worried about the level of security present when offering important information online [37], and will execute transactions only when they generate a specific level of trust. Of course, banks should try to boost their security and privacy standards [34] and perform so carefully [39] by examining various techniques for improving their security measures. One study by Chung and Paynter [6] realized that one of the barriers to the adoption of internet banking is that consumer fears of transaction security. The security plays an important role in internet banking transactions and it has been recognized as a main consumer concern in many internet banking adoption

researches [38]. In the significance of customer adoption of internet banking, trust can be related to consumer perception of security and privacy issues [45].

Hypothesis 3: Security is positively influences trust towards the use of Internet banking.

In the Internet environment, a very well-designed website of an online seller should offer consumers with adequate information for decisions making in purchasing, navigating in the user-friendly environment, interacting with the modern vendor, and buying procedure easy. It should impact users to trust an online salesperson as does a vendor in the traditional marketing environment. In overall, a good quality of the website should improve the growth of consumer trust in online environments.

Hypothesis 4: Quality of the website is positively influences trust towards the use of Internet banking.

Customers like a website that is easy and simple for navigating to achieve their research, to acquire complete information exactly, and to enhance their decision making process. Gefen and Straub have recommended usability and navigability will help customers discover a website as easy to use [16]. Usability describes the activities of the user that has interaction with the web site, for navigating, browsing and performance of tasks easily until completion of the transaction [10]. Diniz et al. describes usability into the three categories namely layout, data entry and user on command [10]. The layout means the display of information on the web page; data entry means the fields for user information collection and user on command means the level of control users over the web page.

Hypothesis 5: Usability positively influences trust towards the use of Internet banking.

The users knowledge related to the experience the user has about the web, risk and security matters (e.g. Level of knowledge about internet security, https, digital certificates, reputation system. Costante et al. [9] argued that the users knowledge has an effect on trust perception in the IT Security field. In [20] indicated that the bank managers might focus on training and promotion programs for improving their customers knowledge regarding security mechanisms and other concepts and these programs can effect on customers trust.

Hypothesis 6: Users knowledge is positively influences trust towards the use of Internet banking.

A consumers disposition to trust is a usual trend to display trust between humans and to accept a trusting position toward others [15]. Due to the fact that consumers have different developing experiments, character types, and social qualification, they differ in their natural disposition to trust. If a consumer has a high disposition to trust others in general, this tendency is very likely to absolutely impact his or her trust in a particular marketing party, whereas a user with a low disposition to trust others in general is likely to grow a comparatively lower trust in a specific selling party. Disposition to trust can be defined in the two different groups namely fast and slow. Fast trust can happen when customer visit a website as a first time then trust to the website with the lack of the last experience. Slow trust beliefs the site gradually with a high level of good experiences [8]. In several studies the disposition to trust is realized as a significant factor of consumer that can impact on the initial trust of customer in an online environment [15, 28, 42]. Tendency to trust or disposition to

trust is the context to which a customer shows an orientation to be depended on others throughout a various situations and persons [28]. Grabner-Kräuter and Faullant found that not only customers may be the target of disposition to trust, but tendency to trust can be a genetic tendency both across different circumstances and different targets of trust [20].

Hypothesis 7: Disposition of trust is positively influences trust towards the use of Internet banking.

The value of trust is based on the potential use of the technology to increase information sharing. Trust boosts the possibility of a dealing partners motivation to improve the amount of information sharing and to discover new, mutually beneficial agreements [8]. Trust on the internet can gain through long-term internet usage. Moreover trust has been founded as an important factor in the adoption decision [14]. Suh and Han [40] by using a Web-based survey found trust an important factor in consumer adoption of internet banking, while Rexha et al. [33] found similar results in Singapore. Many researchers have indicated that trust is a significant factor for understanding customer behavior and relates to internet banking. Trust is not only a simple issue, but also the most significant barrier to realizing the strength points of Business to Customer e-commerce [17]. The role of trust could be even more important in an e-commerce setting, since customers in an online environment cannot access to business deal directly with all of the their interactions with the firm, or in interactions with the staffs of the firm [43].

Hypothesis 8: Trust is positively influences customers towards the use of Internet banking.

28.3 Research Methodology

This research focuses on students and staffs in the Faculty of Computing in the university who have Internet banking accounts. We used some criteria to choose suitable respondents from among the target population. The first criterion is the fact that postgraduate students generally have sufficient knowledge of the concept of trust in the online environment. The second is the requirement that students and staff should have an online account, as well as having adequate experience and knowledge about the internet banking process for the purposes of this case study (Faculty of Computing). These criteria ensure that the respondents have sufficient knowledge and suitability to be able to answer the questionnaire. First, the faculty is selected; second, we asked staffs of the faculty to participate in this study; third, questionnaires in hard-copy format were distributed to staff and students.

A survey questionnaire was designed to collect data from samples. A total of 210 questionnaires was distributed, of which 165 were collected and 150 were valid for analysis. This sample size is enough for structural equation modeling [4]. The items in the questionnaire were adapted from previous studies. All items were measured using a 5 point scale from a range of “strongly disagree” to “strongly agree”. Table 28.1 shows the demographic variables of the respondents.

Table 28.1 Characteristics of the respondents

Measure	Categories	Frequency	Percent
Group	Staff	52	34.6
	Student	98	65.4
Age	Greater than 46	21	14
	Between 36 and 45	39	26
	Between 25 and 35	49	32.67
	Less than 25	41	27.33
Education	Diploma	15	10
	Bachelor/UP	45	30
	Master/PG	53	35
	Doctorate/PhD	38	25
Length of experience	Years > 5	61	40.67
	3 < Years < 5	66	44
	Years <3	23	15.33

28.4 Results of Reliability and Validity

The structural equation modeling approach was used to validate our research model. To perform the analysis, partial least squares (PLS) were employed. PLS offers several strengths according to the [5] it places the least demands on measurement scales; it is appropriate for conditions with little theoretical development; it prevents identification problems of recursive models; it prevents factor indeterminacy problems; it makes no assumptions about the data; it assumes the errors are uncorrelated; it works well with small samples. Data analysis was carried out in two steps. Second, the path coefficient in the research model was assessed through the structural equation model. The significance of each path coefficient using t-tests; to do this, we employed bootstrapping (with 150 sub-samples).

Table 28.2 shows the Inserting the values of AVE in table Latent Variable Correlations. As we can see the root of all values in the correlation values are larger than they are under the related value. After carrying out the necessary investigations relating to the validity, Reliability research tool it comes to measurement. Also, Table 28.3 shows the factor loadings of the measurement items, composite reliability, AVE or Communality and Cronbachs Alpha. The factor loadings of all items exceed there commended level, with 0.60 representing convergent validity and all t-values are also greater than 1.96 [7]. We checked two measures of reliability: composite reliability and Cronbachs Alpha.

PLS can test the convergent and discriminate validity of the scales. Table 28.3 presents the commonality of the measurement items. If this amount is more than 0.5 for all variables can be argued that our questionnaire has good concurrent validity [18]. But there is another view in this regard; some scholars have argued that the root of each of the values for each variable in this table must be greater than its correlations with other variables.

Table 28.2 Values of AVE in table latent variable correlations

Dimension	AVE	1	2	3	4	5	6	7	8	9
Adoption	0.78595	0.88653								
Accessibility	0.7981	0.45454	0.89337							
Disposition of trust	0.82176	0.29528	0.14077	0.90651						
User's knowledge	0.74837	0.56743	0.25906	0.22937	0.86508					
Privacy	0.88270	0.45973	0.42326	0.27994	0.39145	0.93952				
Quality	0.92711	0.25831	0.18418	0.02310	0.13757	0.22521	0.96286			
Security	0.89454	0.57479	0.28506	0.22850	0.33087	0.3938	0.22579	0.94580		
Trust	0.84709	0.73018	0.52265	0.39198	0.52183	0.6052	0.34552	0.62637	0.92037	
Usability	0.81389	0.40897	0.21953	0.0877	0.27986	0.23759	0.00033	0.22400	0.36138	0.90216

Table 28.3 Overview

Construct	Items	Item loading	Composite reliability	AVE	Cronbach alpha	R Square	Commonality
Accessibility	Acc1.1	0.837	0.959	0.798	0.949		0.79812
	Acc1.2	0.903					
	Acc1.3	0.899					
	Acc1.4	0.889					
	Acc1.5	0.918					
	Acc1.6	0.91					
Privacy	Priv2.1	0.929	0.967	0.882	0.955		0.882702
	Priv2.2	0.942					
	Priv2.3	0.942					
	Priv2.4	0.943					
Security	Sec3.1	0.945	0.971	0.894	0.96		0.894549
	Sec3.2	0.946					
	Sec3.3	0.945					
	Sec3.4	0.945					
Quality	Qual4.1	0.959	0.984	0.927	0.98		0.927117
	Qual4.2	0.964					
	Qual4.3	0.961					
	Qual4.4	0.965					
	Qual4.5	0.962					
Usability	Usab5.1	0.909	0.956	0.813	0.943		0.813896
	Usab5.2	0.935					
	Usab5.3	0.923					
	Usab5.4	0.877					
	Usab5.5	0.862					
knowledge	Know6.1	0.825	0.899	0.748	0.831		0.748378
	Know6.2	0.877					
	Know6.3	0.89					
Disposition to trust	DisTr7.1	0.922	0.948	0.821	0.927		0.821767
	DisTr7.2	0.881					
	DisTr7.3	0.876					
	DisTr7.4	0.943					
Trust	TR8.1	0.923	0.943	0.847	0.909	0.703	0.847094
	TR8.2	0.919					
	TR8.3	0.918					
Adoption	AD9.1	0.923	0.943	0.785	0.91	0.533	0.785951
	AD9.2	0.919					
	AD9.3	0.918					

In this case, we can claim that there is divergent validity in that variable. After calculating the square root of AVE values, for comparison the results with correlations, we can check the Latent variable correlation table. In this purpose, Instead of being the ones above in diameter matrix, we insert the square root of AVE of Corresponding to the variables related. The square root of the AVE of each construct should be much larger than the correlation of the specific construct with any of the other constructs in the model [5].

As shown in Table 28.3, our scales had Cronbachs Alpha and composite reliability exceeding 0.80 demonstrating adequate reliability. To test for convergent validity we used Fornell and Larckers assessment criteria [13]: factor loadings of all items should be significant and surpass 0.7 and the average variance extracted for each construct should surpass 0.5. Overview table shows a summary of the reliability and validity of the information.

28.5 Structural Research Model Test

This section includes the analyses of the structural (interior) model done through examining the paths between the constructs in the model for determining the important elements of the variables. The outcomes of bootstrapping technique in PLS can aid to determine the important paths [41]. In the current study, the PLS technique was employed to evaluate the model and examine the assumed associations for the reason that this software can function for little or average sizes of samples [4].

PLS is capable to examine the paths in the model for every bootstrap sample automatically provided through the processes used in bootstrapping process. Such an approach has been utilized for analysing the significant regressed [2].

Bootstrapping was used in the present research to form 150 sub-samples. T-values are obtained based on the bootstrapping method which coordinates with inner and outer model paths. In addition, in order to examine the hypothesis, the probability value (P-value) is used. We used T-distribution calculations in Excel (probabilities) for examining the (P-value), according to the TDIST equation:

$$P(\text{value}) = \text{TDIST}(\text{t value}; \text{sample}; 2)$$

We presented the results of the hypothesis assessment in the Table 28.4. The t-distribution is used in the hypothesis testing of small sample data sets. P-value < 0.05 implies the significance of the related hypothesis (e.g. [22]). As we can see the P-value for each hypothesis is calculated according to its (T-value).

For preceding the study, according to the evaluation and prediction of the structural model, some data about the t-values, path coefficients (β), p-values (p), squared R (R^2) are identified in the details.

Squared R (R^2): The R^2 shows the expected effect of the model R^2 dependent variables through estimating the percentage of a constructs variance in the model [22].

Path coefficients (β): Path coefficients (β) shows how strong and significant the associations between dependent and independent variables are [26]. It means that,

Table 28.4 Hypothesis assessment

Hypothesis	T-value	P-value	Original sample (O)	Result
H1: Accessibility → Trust	3.237632	0.00148	0.216161	Support
H2: Privacy → Trust	2.654197	0.00881	0.197087	Support
H3: Security → Trust	4.263915	0.0000354	0.31781	Support
H4: Quality → Trust	2.812208	0.00558	0.168756	Support
H5: Usability → Trust	2.0528	0.0418	0.128705	Support
H6: Users knowledge → Trust	2.607093	0.0101	0.181943	Support
H7: Disposition to trust → Trust	2.691506	0.00792	0.184629	Support
H8: Trust → Adopt Internet Banking	15.63408	2.66E-33	0.730148	Support

a path coefficient reveals the immediate influence of a variable (considered as cause) that is supposed to result in a different variable (considered as effect). Since a Path coefficient can be identified based on the correlation, it is standardized while a path regression coefficient cannot be considered standardized.

Hypothesis testing: According to the [5], for conducting the hypothesis testing the path significance can be determined via t-tests values by using the bootstrapping procedure. Commonly, the acceptable value for t-value is larger than two. (T-value > 2) means significant level [11].

P-value: The P-value can be considered as a quantitative measure of the numerical importance of testing a hypothesis. Furthermore, regarding the studies conducted formerly, P-value < 0.05 implies the significance of the related hypothesis (e.g. [22]).

In summary, the formulated hypotheses (H1, H2, H3, H4, H5, H6, H7 and H8) were supported by the data. The preceding constructs together explained 0.703 of the variance in the dependent construct: Trust and Trust showed 0.533 of the variance in the dependent construct: Adopt Internet Banking.

28.6 Conclusion

Based on the results, the following factors were observed, namely: accessibility, privacy, security, quality, usability, users knowledge and disposition to trust have significant effects upon trust. Accordingly, the trust can be considered as a critical

importance factor relating to the adoption of Internet banking. The concept of this study is relatively new because the selected factors of trust are related to the characteristics of customer and website. Importantly, this study proves that trust is a crucial factor that should be factored into the adoption of internet banking services. In summary, it was shown that where trust exists, adoption of internet banking would be assured.

This research concerns the effort to investigate information about the relationships among the selected factors of trust. Empirical information about the relationships among the selected factors of trust is then provided. These research findings support the fact that selected factors of trust have a positive influence on trust for adopting customers. It was shown that where trust exists, appreciable levels of adoption will be acquired.

Bank managers should ensure that customers are able to conduct financial activities when an adequate level of trust has been acquired. It is safe to suggest that with such long-terms of its faction, long-term adoption of the internet banking services can be enhanced.

The implications of these results for project personnel selection are significant. On the customer side, a sound knowledge of internet banking, as well as general interpersonal skills are the major requirements for those to be assigned. On the bank side, these results can be useful for improving the level of adoption among their customers. This matter can lead to an increased level of profit for banks.

References

1. Alsajjan, B., Dennis, C.: Internet banking acceptance model: cross-market examination. *Elsevier, Bus. Res.* **63**, 957–963 (2010)
2. Austin, P.C., Jack, V.Tu.: Bootstrap methods for developing predictive models. Taylor & Francis, *Am. Stat.* **58**, 131–137 (2004)
3. Chen, Y.-H., Barnes, S.: Initial trust and online buyer behaviour. Emerald Group Publishing Limited, *Ind. Manage. Data Syst.* **107**, 21–36 (2007)
4. Cheung, C.M., Chan, G.W., Limayem, M.: A critical review of online consumer behavior: empirical research. *IGI Global, J. Electron. Commer. Organ. (JECO)* **3**, 1–19 (2005)
5. Chin, W.W.: The partial least squares approach to structural equation modeling. *Erlbaum, Mod. Methods Bus. Res.* **295**, 295–336 (1998)
6. Chung, W., Paynter, J.: An Evaluation of Internet Banking in New Zealand. *IEEE, System Sciences* (2002)
7. Cronbach, L.J.: Coefficient alpha and the internal structure of tests. *Springer, Psychometrika* **16**, 297–334 (1951)
8. Corritore, C.L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. *Elsevier, Int. J. Hum.-Comput. Stud.* **58**, 737–758 (2003)
9. Costante, E., den Hartog, J., Petkovic, M.: On-line Trust Perception: What Really Matters. *IEEE, Socio-Technical Aspects in Security and Trust (STAST)* (2011)
10. Diniz, E., Porto, R.M., Adachi, T.: Internet banking in Brazil: evaluation of functionality, reliability and usability. *Electron. J. Inf. Syst. Eval.* **8**, 41–50 (2005)
11. Eckblad, J.W.: How Many Samples Should be Taken? *JSTOR, BioScience* (1991)
12. Flavián, C., Guinalíu, M., Gurrea, R.: The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Elsevier, Inf. Manag.* **43**, 1–14 (2006)

13. Fornell, C., Larcker, D.F.: Structural equation models with unobservable variables and measurement error: algebra and statistics. *JSTOR, J. Mark. Res.* 382–388 (1981)
14. Gartner, L.M., Greer, F.R.: Prevention of rickets and vitamin D deficiency: new guidelines for vitamin D intake. *Am. Acad. Pediatr.* **111**, 908–910 (2003)
15. Gefen, D.: E-commerce: the role of familiarity and trust. Elsevier, *Omega* **28**, 725–737 (2000)
16. Gefen, D., Srinivasan Rao, V., Tractinsky, N.: The conceptualization of trust, risk and their electronic commerce: the need for clarifications. *IEEE, System Sciences* (2003)
17. Gefen, D., Straub, D.W.: Managing user trust in B2C e-services. Indian University Press, *E-service J.* **2**, 7–24 (2003)
18. Gefen, D., Straub, D.: A practical guide to factorial validity using PLS-graph: tutorial and annotated example. *Communications of the Association for Information Systems*, vol. 16, (2005)
19. Gerrard, P., Cunningham, J.B.: The diffusion of Internet banking among Singapore consumers. *MCB UP Ltd, Int. J. Bank Mark.* **21**, 16–28 (2003)
20. Grabner-Kräuter, S., Faullant, R.: Consumer acceptance of internet banking: the influence of internet trust. Emerald Group Publishing Limited, *Int. J. Bank Mark.* **26**, 483–504 (2008)
21. Haque, A., Khatibi, A., Al Mahmud, S.: Factors determinate customer shopping behaviour through internet: the Malaysian case. *Aust. J. Basic Appl. Sci.* **3** (2009)
22. Ifinedo, P.: Examining the influences of external expertise and in-house computer/IT knowledge on ERP system success. Elsevier, *J. Syst. Softw.* **84**, 2065–2078 (2011)
23. Jayawardhena, C., Foley, P.: Changes in the banking sector-the case of internet banking in the UK. *MCB UP Ltd, Int. Res.* **10**, 19–31 (2000)
24. Kassim, N.M., Abdulla, A.K.M.A.: The influence of attraction on internet banking: an extension to the trust-relationship commitment model. *Int. J. Bank Mark.* **24**, 424–442 (2006)
25. Kim, K., Prabhakar, B.: Initial trust, perceived risk, and the adoption of internet banking. In: *Proceedings of the Twenty First International Conference on Information Systems* (2000)
26. Ko, D.G., Kirsch, L.J., King, W.R.: Antecedents of knowledge transfer from consultants to clients in enterprise system implementations. *JSTOR, MIS.* (2005)
27. Ko, D.G., Kirsch, L.J., King, W.R.: Understanding consumer adoption of internet banking: an interpretive study in the Australian banking context. *J. Electron. Commer. Res.* **7** (2006)
28. McKnight, D.H., Choudhury, V., Kacmar, C.: Developing and validating trust measures for e-commerce: an integrative typology. *INFORMS, Inf. Syst. Res.* **13**, 334–359 (2002)
29. Mukherjee, A., Nath, P.: A model of trust in online relationship banking. *MCB UP Ltd, Int. J. Bank Mark.* **21**, 5–15 (2003)
30. Mwesigwa, R., Nkundabanyanga, S.K.: Consumer attitude, trust, perceived risk and internet banking adoption in Uganda. *J. Bus. Econ.* (2011)
31. Nissenbaum, H.: Privacy as contextual integrity. *HeinOnline*, **79** (2004)
32. Pikkarainen, T., Pikkarainen, K., Karjaluoto, H., Pahnla, S.: Consumer acceptance of online banking: an extension of the technology acceptance model. Emerald. Group. Publ. Limited. **14**, 224–235 (2004)
33. Rexha, N., Kingshott, R.P.J., Aw, A.S.S.: The impact of the relational plan on adoption of electronic banking. *MCB UP Ltd, J. Serv. Mark.* **17**, 53–67 (2003)
34. Rotchanakitumnuai, S., Speece, M.: Barriers to internet banking adoption: a qualitative study among corporate customers in Thailand. *MCB UP Ltd, Int. J. Bank Mark.* **21**, 312–323 (2003)
35. Rupp, W.T., Smith, A.D.: Strategic e-commerce aspects of the e-banking/e-lending industry. IGI Global, *Strategies for Generating E-business Returns on Investment* (2005)
36. Sanayei, A., Noroozi, A.: Security of internet banking services and its linkage with users' trust: a case study of Parsian bank of Iran and CIMB bank of Malaysia. *IEEE, Info. Manag. Eng.* (2009)
37. Shim, S., Eastlick, M.A., Lotz, S.L., Warrington, P.: An online prepurchase intentions model: the role of intention to search: best overall paper award, the sixth triennial AMS/ACRA Retailing Conference, 2000. Elsevier, *J. Retail.* **77**, 397–416 (2001)
38. Siu, N.Y.M., Mou, J.C.W.: Measuring service quality in internet banking: the case of Hong Kong. *Taylor & Francis* **17**, 99–116 (2005)

39. Spence, L.J., Coles, A.M., Harris, L.: The forgotten stakeholder? Ethics and social responsibility in relation to competitors. *Wiley Online Library, Bus. Soc. Rev.* **106**, 331–352 (2001)
40. Suh, B., Han, I.: Effect of trust on customer acceptance of Internet banking. *Elsevier, Electron. Commer. Res. Appl.* **1**, 247–263 (2003)
41. Tenenhaus, M., Vinzi, V.E., Chatelin, Y.M., Lauro, C.: PLS path modeling. *Elsevier, Comput. Stat. Data Anal.* **48**, 159–205 (2005)
42. Teo, T.S., Liu, J.: Consumer trust in e-commerce in the United States, Singapore and China. *Elsevier, Omega* **35**, 22–38 (2007)
43. Urban, G.L., Sultan, F., Qualls, W.J.: Placing trust at the center of your internet strategy. *Sloan Manag. Rev.* **42** (2000)
44. Wang, Y.S., Lin, H.H., Luarn, P.: Predicting consumer intention to use mobile service. *Wiley Online Library, Inf. Syst. J.* **16**, 157–179 (2006)
45. Wang, Y.S., Wang, Y.M., Lin, H.H., Tang, T.I.: Determinants of user acceptance of internet banking: an empirical study. *MCB UP Ltd, Int. J. Serv. Ind. Manag.* **14**, 501–519 (2003)
46. Yousafzai, S., Pallister, J., Foxall, G.: Multi-dimensional role of trust in Internet banking adoption. *Taylor & Francis, Serv. Ind. J.* **29**, 591–605 (2009)

Chapter 29

Robust Storage Assignment in Warehouses with Correlated Demand

Monika Kofler, Andreas Beham, Stefan Wagner and Michael Affenzeller

Abstract In many warehouses manual order picking is one of the most time and labour intensive processes. Products that are often ordered together are said to be *correlated* or *affine* and order picking performance may be improved by placing correlated products close to each other. In industries with strong seasonality patterns and fluctuating demand regular re-locations of products might be necessary to ensure that the quality of the storage assignment does not deteriorate over time. In this chapter we study how to generate more robust assignments that are suitable for volatile warehouse scenarios with correlated demand. In a case study based on 13 monthly snapshots from a real-world warehouse robust slotting outperformed greedy re-locations by up to 9.6 %.

29.1 Storage Assignment in Dynamic Warehouses

For more than fifty years researchers have studied how the arrangement of parts within a warehouse—also called storage assignment or *slotting*—affects logistic performance. Out of all the warehouse processes, order picking has attracted the most attention for performance optimisation. Order picking is the process of retrieving customer orders from the warehouse; it is very labour-intensive and typically accounts for a large chunk of the warehouse operating expenses (roughly 55 % [19, p. 434]).

M. Kofler (✉) · A. Beham · S. Wagner · M. Affenzeller
Heuristic and Evolutionary Algorithms Laboratory (HEAL), University of Applied Sciences
Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria
e-mail: monika.kofler@gmail.com; mkofler@heuristiclab.com
<http://heal.heuristiclab.com>

A. Beham
e-mail: abeham@heuristiclab.com

S. Wagner
e-mail: swagner@heuristiclab.com

M. Affenzeller
e-mail: maffenze@heuristiclab.com

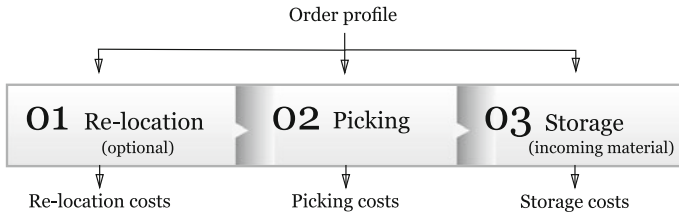


Fig. 29.1 Each period of the multi-period SLAP process consists of an (optional) re-location phase, followed by order picking and storage of incoming material for the subsequent period

Diverse warehouses such as container yards, automated storage and retrieval systems or high-racks have been modelled as *Storage Location Assignment Problems* (SLAP) and evaluated with respect to order picking performance. Most publications focus on *static* SLAP variants [9] and do not consider the costs for implementing a new assignment during operation, the impact of fluctuating order patterns, incoming and outgoing materials or re-locations.

In a dynamic environment, the expected performance gains in picking must be offset against the costs for re-locating materials. Figure 29.1 depicts the components of a multi-period SLAP model and the associated re-location, picking and storage costs.¹ One core issue is how to determine a balance between the three cost factors and achieve a good overall warehouse performance over time. Extensive re-locations are usually neither necessary nor economically feasible.² The question is therefore how to *identify* [11] materials that should be re-located and *how often* [16] and *how many* re-locations should be performed.

In [11] we proposed a greedy local search strategy that ranks re-location moves by their estimated improvement on picker performance and performs a fixed number of re-locations per period. The number of moves was determined empirically. By contrast, [16] suggested to monitor how many materials are stored in the “wrong” warehouse zone in each period and to assess the costs of re-locating them all versus projected picker performance gains. If the costs are smaller than the anticipated gains the moves should be implemented. [5, 15] proposed heuristics that re-locate materials to more favourable locations during idle times or interleave picking, storage and re-location tasks. In [5, 15] the capacity of the crane, which has to perform picking, storage and re-locations, is the main limiting factor.

One issue with all the above approaches is that they do not consider the *robustness* of an assignment. Some materials with strong demand fluctuations may be scheduled for re-location in every period, thus adding to the work load of the warehouse personnel. By contrast, robust assignments should require less maintenance to retain a certain picking performance. In this chapter, we will extend a robust re-location

¹ A detailed description of the multi-period SLAP and experimental results can be found in [11].

² Experiments in [13] showed that moving a limited amount of items can significantly improve picker performance. By re-locating only 60 pallets in a warehouse with more than 1,400 pallets a reduction of picker travel distances of 23% could be achieved. Conversely, almost the entire stock would have to be moved to realise a 60% reduction.

procedure initially proposed in [17] and compare it with conventional re-location strategies.

29.1.1 Overview of Storage Assignment Strategies

The three main families of assignment strategies are *random*, *turnover based* and *affinity based*. *Random* assignments are frequently used as a baseline in scientific research but are also adopted in real-world warehouses. For example, Amazon.com, the world's largest online retailer, shelves incoming orders via random storage. Random assignments tend to result in longer average picker travel times but also lead to better load balancing between pickers and scale well. By contrast, *turnover based* approaches rank materials by pick frequency or related key figures and assign in-demand materials to more favourable storage locations. The cube-per-order index (COI) [10] is a well-know representative of the family of turnover based assignment strategies. Finally, *affinity based* strategies place materials that are frequently ordered together (= *correlated* demand) close to each other to reduce the average order picker tour length. Various clustering [1, 4, 7, 18] and (meta-)heuristic approaches [6, 8, 11, 14] have been employed to generate affinity based assignments.

Both turnover and affinity based slotting approaches can lead to congestion issues when multiple pickers compete for concurrent access to a warehouse aisle or area. The introduction of warehouse *zones* can alleviate this problem. The allocation of materials to zones is performed via a turnover or affinity based strategy but within a zone the assignment to concrete storage locations is randomised. A-B-C analysis is a well-known variant of zone based storage, which allocates products to three classes according to their turnover frequency. Class A contains the 10% that are requested most frequently, i.e. the fast-moving parts, and class C the 70% least picked parts. The remaining 20% are class B parts. Different class splits such as 20%-30%-50% are also possible.

29.1.2 How Demand Volatility Affects Slotting

To illustrate the issue of robustness in warehouse assignments, we analysed 13 data snapshots from a real-world production warehouse from the automotive industry. Each dataset consists of roughly one month of planned orders. For more details on the data set refer to Sect. 29.2.

In each period we employed A-B-C analysis with a 10%-20%-70% split and recorded the classification history per material. As depicted in Fig. 29.2 we found that 38% of the materials exhibited a stable class assignment; most of them are slow-moving parts in class C but just over 3% of all parts exhibited constant high pick volumes across all periods and were therefore always assigned to class A. On the other hand, a classification pattern such as CBBCBABAACACA (as displayed in

Fig. 29.2 Volatility of A-B-C classification for materials in the benchmark dataset. 11 % of the parts are highly volatile and switch class assignment between five and ten times within 13 periods. By contrast, 38 % of parts are stable and placed in the same zone in each period

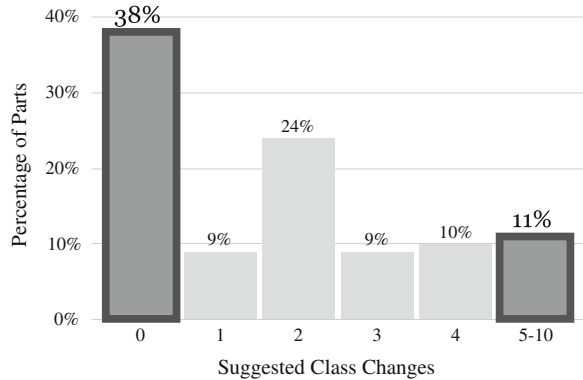


Fig. 29.3 A-B-C slotting suggested a placement sequence in zones CBBCBABAACACA for one of the most volatile parts in data set. This could trigger up to ten re-locations, depending on how the algorithm selects parts for movement

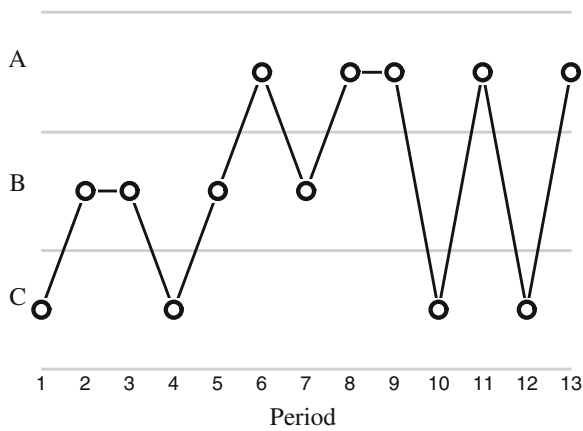


Fig. 29.3) is problematic, because a greedy slotting algorithm might suggest frequent re-locations. 11 % of the materials in the dataset required a move between classes in more than one third of the periods and were thus classified as very volatile.

29.1.3 Generating Robust Warehouse Assignments

Pierre et al. proposed a dynamic variant of the A-B-C storage policy [17] in which the classification is re-calculated regularly, for example every day. The total number of picks per period may vary but the relative size of the classes and associated storage areas is constant; thus the class limits l_{AB} and l_{BC} must be re-calculated in every period. Over time, this generates a suggested placement sequence such as the one displayed in Fig. 29.3. In each period a small number of moves is selected based on their suggested placement sequence. First, moves are generated for all products that are currently located in the “wrong” class and ranked by a priority scheme based on

three criteria. The priority of a move increases (a) if the material is in the wrong class for multiple periods, (b) if the change in pick frequency is large or (c) if the variation in demand is small.

We adapted this dynamic assignment approach in a way that allows the use of alternative storage paradigms beyond A-B-C classification. The only requirement is that the storage locations in the warehouse have to be grouped into *zones*. These can be “virtual” zones and do not need to correspond to the actual zoning strategy within the warehouse. The algorithm is described in more detail in Sect. 29.3.2.

29.2 Case Study

We obtained 13 monthly snapshots of real-world order planning and assignment data from an Austrian manufacturer in the automotive industry. The data was anonymised and simplified to generate this multi-period benchmark problem.³ In the following paragraphs we briefly describe the warehouse layout, zones, material handling equipments, routing strategy and constraints.

29.2.1 Layout and Assignment Constraints

To increase reproducibility of the experiments, a simplified warehouse layout is used instead of the real rack configuration. Our model warehouse is rectangular with 12 parallel, two-sided aisles. Each rack is 27 pallets deep and 9 levels high, which amounts to a maximum storage capacity of 5,832 pallets for the whole warehouse. The distance between aisles is 5 m, the horizontal distance between locations within an aisle 0.5 m. We assume that each material occupies only one storage location and that a sufficient amount is in stock to fulfil the orders (no stock-outs). Materials can be stored in any location in the warehouse. In total, almost 12,000 distinct products are stored in the warehouse over the observation period but only 4,867–5,585 products are in stock during any given period.

29.2.1.1 Zones

We define twelve zones of equal size, each of them denoting one two-sided aisle in the warehouse. This simple zone definition allows us to track how often a material is assigned to a different aisle in our multi-period benchmark. It is completely independent of the slotting approach, which can be a simple construction heuristic such as COI, a clustering-based approach or a meta-heuristic.

³ The multi-period SLAP benchmark instance is available for download via <http://dev.heuristiclab.com/trac/hl/core/wiki/AdditionalMaterial>.

29.2.1.2 Material Handling Equipment and Picker Routing

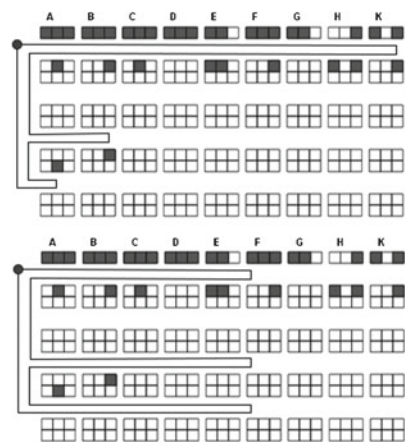
All storage, re-location and picking tasks are performed by two fork lifts. To process an order, pick locations are sorted in ascending order by aisle and then by rack column. The fork lifts can only enter the aisle from the front and have to reverse to back out because there is insufficient space to turn around. This strategy is known as *return routing*. Figure 29.4 (top) shows the exact path of a picker to retrieve an order. The fork lift traverses each aisle until the depth of the last part and then backs out. If we assume that parts are randomly located within an aisle, the picker has to travel to mid-point on average as depicted in Fig. 29.4 (bottom). With the approximate routing algorithm the evaluation produces the same result, as long as an assignment strategy places the parts of the order in these three aisles.

We evaluate assignments with the approximate routing algorithm to avoid that optimisation algorithms greedily fill the most desirable locations in the warehouse as depicted in Fig. 29.5. While this greedy behaviour is acceptable for a static SLAP instances, it leaves only less desirable storage locations for incoming materials in a multi-period formulation as discussed in [11]. Therefore we chose to evaluate slottings with the approximate routing algorithm, which ignores placement variations as long as parts are assigned to the same aisles.

29.3 Algorithms

We employ three different algorithms to generate target warehouse assignments, namely random, COI and a novel affinity based slotting algorithm, which is discussed in more detail in the following subsection.

Fig. 29.4 Path for retrieving an order with return routing. Locations with parts that need to be picked are coloured *dark grey*. The real route is depicted on the *left*, an approximation that assumes that each aisle is traversed until mid-point is shown on the *right*. Only the first four out of twelve two-sided aisles of the benchmark layout are displayed



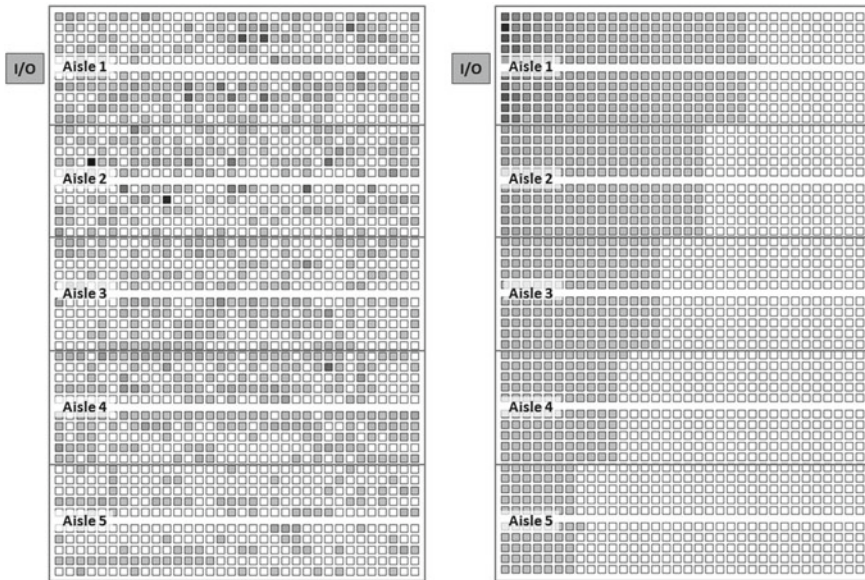


Fig. 29.5 An initial random slotting (*left*) is compared to a turnover optimised slotting (*right*). In the optimised assignment all the closest locations to the shipping dock in the *upper left* are occupied. Reproduced with permission from [11] ©Springer 2014

29.3.1 Density Based Clustering

Various hierarchical and non-hierarchical clustering approaches have been used to group parts by affinity [1, 4, 7, 9, 18]. We decided to use the OPTICS algorithm [3] to identify clusters of similar density—which translates to similar affinity—in the data. We define the affinity between products a and b as

$$affinity(a, b) = \begin{matrix} \text{relative number of orders in which} \\ \text{products } a \text{ and } b \text{ occur together.} \end{matrix} \tag{29.1}$$

Note that affinities are not transitive, that is if products a and b are frequently ordered together and products b and c are also frequently ordered together that does not allow us to infer an affinity between a and c . The pairwise affinities per period are then normalised and transformed via a distance function

$$distance(a, b) = 1 - \frac{affinity(a, b)}{\max affinity}, \tag{29.2}$$

which can subsequently be used to cluster materials. The OPTICS algorithm creates a linear ordering of data elements in such a way that those with the closest distance become neighbours. If Eq. (29.2) is used as a distance function this means that parts

with high affinity are close to each other in the ordering. In addition, OPTICS calculates the so-called *reachability distance* for all pairs of elements. The reachability distance denotes the minimum density that needs to be covered to result in both elements becoming part of the same cluster. If the ordering and reachability distances are plotted in a *reachability plot* (cf. Figure 29.6), clusters can be identified visually as valleys.

In addition to manual, visual identification of clusters from the reachability plot, a variety of automated extraction methods can be used. The simplest is to cut the reachability plot at a certain level, similar to cutting a dendrogram. We use ξ -extraction [3], which has the added benefit of extracting a hierarchy of clusters of varying density. The ξ parameter is used to set a minimum relative density jump (upwards or downwards) that indicates a density cluster border. We empirically determined a minimum cluster size (MinPoints) = 5 and $\xi = 0.02$. In the first period this leads to the extraction of four top-level clusters as depicted in Fig. 29.6. This result illustrates that only a small number of parts have strong affinities. The solid grey areas in the plot are of low density and considered noise. The clusters are then ranked by mean pick frequency and assigned to the aisles in sequence to generate an assignment. Within an aisle the placement of parts is randomised. Materials that were categorised as “noise” are assigned to a randomly selected aisle in the warehouse.

29.3.2 Selection of Re-locations to Perform

As discussed in the introduction performing re-locations in a dynamic, multi-period SLAP is only worthwhile if the efficiency gains are higher than the re-location efforts. Implementing a completely new slotting in each period is often neither necessary or feasible. We therefore compare a greedy approach, which ranks moves by estimated picking performance improvements, with an approach that ranks moves by “robustness”. For the robust slotting approach, we first generate a target warehouse slotting with a suitable slotting strategy. Depending on the order profiles, one could for example use COI, A-B-C slotting or the affinity based clustering approach described in Sect. 29.3.1.

As described in Sect. 29.1.2 the history of suggested aisle assignments per part is logged and used to gauge robustness. Parts are not necessarily in stock dur-



Fig. 29.6 Reachability plot generated with ELKI for period 1. The ordering is plotted on the x-axis and the reachability distance on the y-axis. Four top-level clusters are identified with ξ -extraction (ξ set to 0.02)

ing all periods. Let us revisit the material from Fig. 29.3, which was classified as CBBCBABAACACA via A-B-C slotting. By using OPTICS clustering instead, we obtain an assignment to aisles of $x-1-1-8-1-1-2-1-1-x-1-x-1$. x indicates that the part is not in stock during that particular period. The algorithm exhaustively generates moves within the warehouse for all parts that are currently outside their target zone as indicated in Fig. 29.7.

To rank the suggested moves with respect to robustness, the history of suggested placements and the current location of the material have to be considered. Based on [17] we employ the following three measures to calculate a combined robustness rank of each move:

- *Stability*: If the optimisation algorithm suggests a different target aisle during each period, moving that part is less desirable because the new placement may become rapidly outdated again. The aisle variance is calculated from historical placement suggestions. The variance is scaled between [0, 1] and inverted. Therefore a part with a stable pattern is assigned a high stability score.
- *Importance*: If the suggested target storage location changes significantly between two periods, these moves should be favoured as opposed to moves between neighbouring aisles. We therefore calculate the absolute distance between “ideal” and current location and divide it by the number of aisles to obtain a number in the range [0, 1]. A larger value indicates that the move is more important.

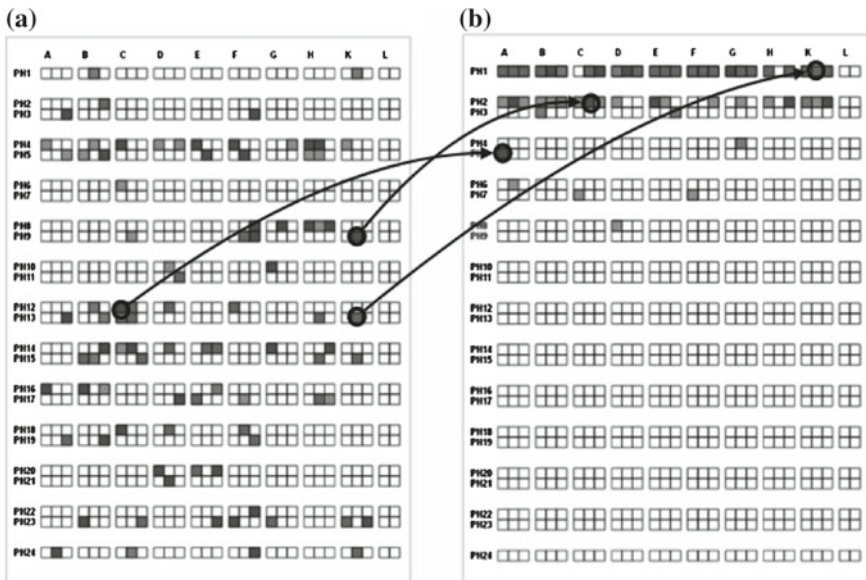


Fig. 29.7 Based on a source and target assignment, we generate transformation moves that allow us to iteratively convert one into the other. The moves can either be sampled or enumerated exhaustively. **a** Current slotting. **b** Target slotting

- *Urgency*: Counts how many periods a part has been in the “wrong” aisle already. The score is reset to zero if the part is moved or if the optimisation algorithm suggests a different aisle than in the last period. The value is divided by the number of observed periods to scale between [0, 1].

We will explore how to best combine and weight the three measures in the experiments section.

29.4 Experiments and Results

All experiments start with the same randomly generated initial assignment of the warehouse in period one. For the subsequent re-location and storage steps, the algorithms discussed in the previous section are used. The OPTICS clusterings were created with ELKI [2]. All other algorithms, the transformation of the clustering to an assignment, the evaluation and plots were implemented in HeuristicLab [20], a framework for heuristic and evolutionary algorithms.⁴ In the re-location phase of a period each algorithm is allowed to move a maximum of fifty materials.

29.4.1 Parameterising the Robustness Score

First, we generated an “ideal” target assignment with OPTICS and HeuristicLab for each period. In the actual optimisation run, all transformation moves that re-locate materials to the “correct” aisles were generated. The top 50 moves were selected in each period based on various combined robustness score variants. Pierre et al. [17] originally used the product of *Stability*, *Importance* and *Urgency* to derive a combined robustness measure. We compared the impact of using a product versus a weighted sum on the overall picking performance and sampled weights in the range [0, 5] for each of the three measures. Figure 29.8 displays a schematic of the results. Most of the parameter combinations have a comparable performance, but the ranking by *Urgency* only (= solid black line in the graph) clearly outperformed the other settings. By comparison, the product used in [17] is marked with a dotted black line. For reference, we will plot both the results for the *Urgency* only and the Pierre et al. robustness score in the comparison charts with the greedy heuristics.

⁴ HeuristicLab and ELKI are free software and available via <http://dev.heuristiclab.com/> and <http://elki.dbs.ifi.lmu.de/>.

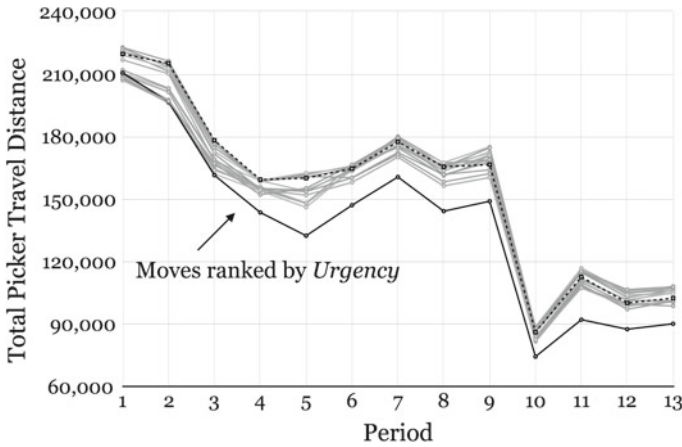


Fig. 29.8 Result of different weights for the robustness calculation. Selecting parts by their *Urgency* only (=solid black lines) provided the best results. The dotted line denotes the setting used in [17]

29.4.2 Greedy Versus Robust Re-Locations

We compare our robust re-locations with two greedy approaches. The greedy algorithms do not rely on a target slotting but instead sample 10,000 moves per period and rank them by (a) pick frequency or (b) affinity and apply the top 50 re-locations. Figure 29.9 shows the results of using a greedy local search versus robust move selection.

If we use the product robustness measure, it takes a couple of periods until the historical data collected on each part is sufficient to get a good robustness ranking. Therefore the robust algorithm initially performs a bit worse than the greedy turnover or affinity algorithms. Starting from period 7, the algorithm with the product robustness ranking delivers better results than the greedy approaches. Overall, the average improvement is merely around 1 %, though. This small improvement does not really merit the use of robust slotting.

By contrast, the Urgency only ranking performs roughly equivalent to the greedy approaches in the first three periods but is consistently better from the fourth period onwards, which results in a respectable average improvement of 9.6 %. This illustrates the large impact of parameter tuning on algorithmic performance.

In addition, robust slotting has another major advantage over greedy approaches: Greedy approaches tend to slot material in the best locations in the warehouse as illustrated in Fig. 29.5. This leads to higher storage costs for incoming material, for which no good locations are available. We tried to alleviate this problem by using a different distance calculation for the return routing as outlined in Fig. 29.2. However, in this case the greedy approaches tend to fill up the “more favourable” aisles, which are closer to the shipping dock. The re-location and storage costs were similar for

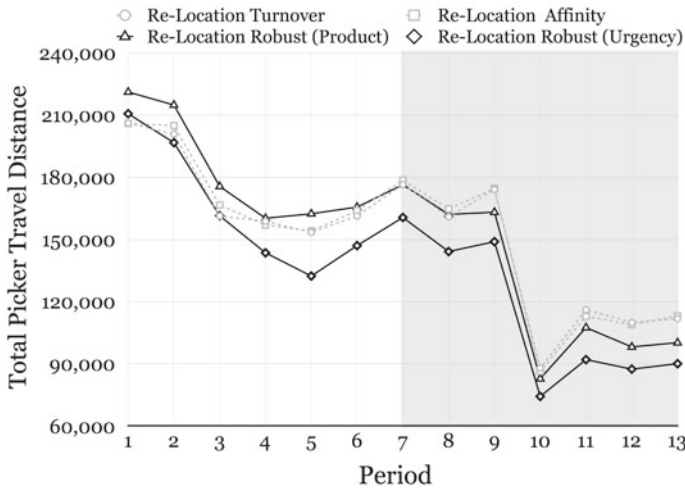


Fig. 29.9 Comparison of total order picker travel costs over all 13 periods with greedy (Turnover, Affinity) versus robust (Product, Urgency) re-location approaches

the greedy and robust approaches, but the pick costs *increased after storage* because incoming material was slotted in a sub-optimal way.

If we plot the picking costs in each period after re-location and after storage we can observe large jumps in picking effort after storage (as shown in Fig. 29.10). In the subsequent re-location stage the just slotted materials are re-arranged and the quality improves—but it is very inefficient. On the contrary, Fig. 29.10 also shows that the peaks are far less pronounced in robust slotting. This means that we can reduce the number of re-locations per period or potentially only re-locate every second or third period without much impact on quality. This also holds for the Pierre et al. parameterisation and shows that robust slotting can be beneficial even without parameter tuning.

29.5 Conclusions and Further Research

In this chapter we proposed a simple extension of the dynamic A-B-C slotting approach developed by Pierre et al. As long as (virtual) warehouse zones are defined this re-location approach can now be combined with the full spectrum of storage assignment strategies proposed in the literature. In addition we first employed a density-based clustering approach to generate an affinity based target slotting. We found that robust slotting outperformed greedy re-location approaches by up to 9.6%, although this required parameter tuning for the robustness measures Stability, Importance and Urgency. On a benchmark dataset based on real-world order profiles the Urgency measure alone yielded the best results. It would be interesting to see how

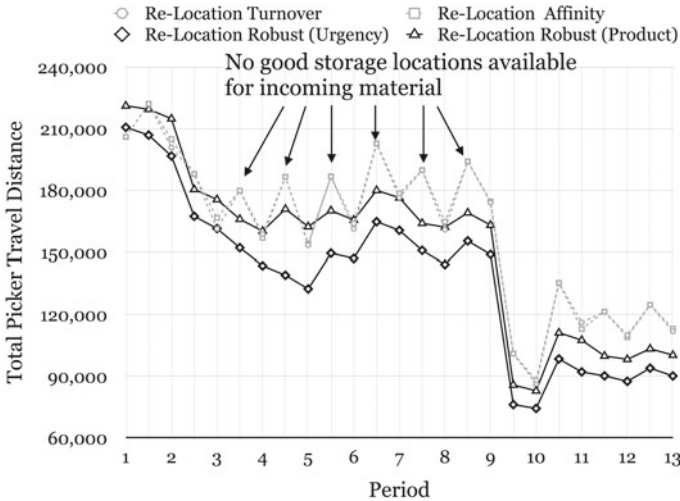


Fig. 29.10 This chart displays the same data as Fig. 29.9 plus the pick costs after storage of incoming materials (plotted on the half-point between periods). The peaks indicate that the greedy approaches do not leave enough good storage locations for incoming material. The picker travel effort rises and has to be corrected in the subsequent period via re-locations

well robust slotting performs on other SLAP instances and whether the parameterisation has to be tuned instance-by-instance or whether the Urgency measure is sufficiently generic. Furthermore, in this chapter the number of re-locations per period was still determined empirically. A self-adaptive algorithm that increases and reduced the number of re-locations on demand would be highly desirable to achieve a better over-all performance and robustness. We believe that robust, multi-period slotting is a crucial step towards a more wide-spread implementation of novel storage assignment approaches in practice.

Acknowledgments This paper is an updated and extended version of [12] and was first presented at the APCASE 2014 conference. The work described in this chapter was done within the Josef Ressel-Centre HEUREKA! for Heuristic Optimization sponsored by the Austrian Research Promotion Agency (FFG).

References

1. Accorsi, R., Manzini, R., Bortolini, M.: A hierarchical procedure for storage allocation and assignment within an order-picking system. A case study. *Int. J. Logist. Res. Appl.* **15**(6), 351–364 (2012)
2. Achtert, E., Kriegel, H.-P., Zimek, A.: ELKI: a software system for evaluation of subspace clustering algorithms. In: Ludscher, B., Mamoulis, N.: (eds.) *Scientific and Statistical Database Management. LNCS*, vol. 5069, pp. 580–585. Springer, Berlin Heidelberg (2008)

3. Ankerst, M., Breunig, M., Kriegel, H.-P. Sander, J.: OPTICS: ordering points to identify the clustering structure. In: Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data (SIGMOD'99), vol. **28**(2), pp. 49-60 (1999)
4. Bindi, F., Manzini, R., Pareschi, A., Regattieri, A.: Similarity-based storage allocation rules in an order picking system: an application to the food service industry. *Int. J. Logist. Res. Appl.* **12**(4), 233–247 (2009)
5. Carlo, H., Giraldo, G.: Optimizing the rearrangement process in a dedicated warehouse. In: Ellis, K., Gue, K., de Koster, R., Meller, R., Montreuil, B., Ogle, M. (eds.) *Progress in Material Handling Research*, pp. 39–48. Material Handling Institute, Charlotte, North Carolina (2010)
6. de Ruijter, H., Schuur, P., Mantel, R., Heragu, S.: Order oriented slotting and the effect of order batching for the practical case of a book distribution center. In: Proceedings of the 2009 International Conference on Value Chain Sustainability (2009)
7. Frazelle, E., Sharp, G.: Correlated assignment strategy can improve order-picking operation. *Ind. Eng.* **4**, 33–37 (1989)
8. Garfinkel, M.: *Minimizing Multi-zone Orders in the Correlated Storage Assignment Problem*. PhD thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, January 2005
9. Gu, J., Goetschalckx, M., McGinnis, L.: Research on warehouse operation: A comprehensive review. *Eur. J. Oper. Res.* **177**(1), 1–21 (2007)
10. Heskett, J.: Cube-per-order index—a key to warehouse stock location. *Trans. Distrib. Manag.* **31**, 27 (1963)
11. Kofler, M., Beham, A., Wagner, S., Affenzeller, M.: Affinity based slotting in warehouses with dynamic order patterns. In: Klempous, R., Nikodem, J., Jacak, W., Chaczko, Z.: (eds.) *Advanced Methods and Applications in Computational Intelligence. Topics in Intelligent Engineering and Informatics*, vol. 6, pp. 123–143. Springer (2014)
12. Kofler, M., Beham, A., Wagner, S., Affenzeller, M.: Robust solutions for a multi-period storage location assignment problem with correlated demand. In: Chaczko, Z., Gaol, F., Pichler, F., Chiu, C.: (eds.) *Proceedings of the 2nd Asia-Pacific Conference on Computer Aided System Engineering (APCASE)*, South Kuta, Indonesia, February 2014, pp. 99–100 (2014)
13. Kofler, M., Beham, A., Wagner, S., Affenzeller, M., Achleitner, W.: Re-warehousing vs. healing: strategies for warehouse storage location assignment. In: Proceedings of the IEEE 3rd International Symposium on Logistics and Industrial Informatics (Lindi 2011), Budapest, Hungary, 25–27 August 2011, pp. 77–82 (2011)
14. Mantel, R., Schuur, P., Heragu, S.: Order oriented slotting: a new assignment strategy for warehouses. *Eur. J. Ind. Eng.* **1**(3), 301–316 (2007)
15. Muralidharan, B., Linn, R., Pandit, R.: Shuffling heuristics for the storage location assignment in an AS/RS. *Int. J. Prod. Res.* **33**(6), 1661–1672 (1995)
16. Neuhäuser, D.: *Ein Ansatz zur simulationsgestützten Planung und Bewertung von Lagerreorganisationsmaßnahmen am Beispiel des Lebensmittelhandels*. PhD thesis, Institut für Fördertechnik und Logistik, Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik, Universität Stuttgart, Germany (2013)
17. Pierre, B., Vannieuwenhuysse, B., Dominanta, D., Van Dessel, H.: Dynamic ABC storage policy in erratic demand environments. *Jurnal Teknik Industri* **5**(1), 1–12 (2003)
18. Sadiq, M.: *A Hybrid Clustering Algorithm for Reconfiguration of Dynamic Order Picking Systems*. PhD thesis, University of Arkansas (1993)
19. Tompkins, J., White, J., Bozer, Y., Tanchoco, J.: *Facilities Planning*, 4th edn. Wiley, New York (2010)
20. Wagner, S.: *Heuristic Optimization Software Systems—Modeling of Heuristic Optimization Algorithms in the HeuristicLab Software Environment*. PhD thesis, Johannes Kepler University, Linz, Austria (2009)

Chapter 30

Concise Supply-Chain Simulation Optimization for Large Scale Logistic Networks

Erik Pitzer and Gabriel Kronberger

Abstract Optimization of supply chains and logistic networks have been largely addressed by simulation-based optimization. The logistics for low-energy biological residues poses a great challenge for logistics that has to be tackled at an international scale. For this application, a new simplified model of logistic networks was created that allows not only an accelerated evaluation through precalculation and aggregation but incorporates on-the-fly optimizations to greatly reduce the solution space. Moreover, an implementation is provided as plugin for the open-source optimization framework HeuristicLab.

30.1 Introduction

Optimization of logistics networks is a challenging problem. Several abstracted problem formulations, such as the traveling salesman problem [15], the vehicle routing problem [2] and different incarnations of routing problems have been created. These are solvable via either A* [9], heuristics as employed in many navigation systems, or via contraction hierarchies. However, no standard abstraction exist for a holistic problem formulation that involves not only transport but also planning of facility construction. For this reason, when first tackling large complex problems, simulation-based optimization [8] is often the method of choice. In this work, a first attempt towards one of many possible standard abstractions for supply chain network optimization is proposed. While stemming from an simulation-based approach the formulation resembles a classic optimization approach comparable in complexity to, e.g., an extension of the capacitated vehicle routing problem [23].

E. Pitzer (✉) · G. Kronberger
School of Informatics, Communication and Media,
University of Applied Sciences Upper Austria,
Research Center Hagenberg, Softwarepark 11,
4232 Hagenberg, Austria
e-mail: erik.pitzer@fh-hagenberg.at

G. Kronberger
e-mail: gabriel.kronberger@fh-hagenberg.at

In particular, the idea stems from an application to an optimization problem, investigated in a European-wide supply chain optimization problem for biological residue recycling.¹ A problem of handling many biological residues is their low energy density. So, while they still possess an usable residual energy, the effort for transportation and handling is deemed to high to profitably extract it [4, 10]. Therefore, the underlying idea is to include relatively cheap preprocessing steps for increasing the energy density. To efficiently perform these intermediate preprocessing steps a good balance between transport and construction cost of the facilities has to be found, which is the motivation for creating an integrated supply chain network for multi-stage biomass transport and conversion on an European scale that has been previously introduced in [19].

One major disadvantage is the size of the solution space. The complexity of this supply chain network is enormous, even when limiting the granularity to regional planning instead of selecting exact locations. First, for each region, the amount of purchased feedstock has to be determined, followed by the selection of appropriate transport modes and means, which are in turn dependent on the selected transport targets that are embodied by decentralized and central conversion facilities. To complicate matters even further, several different conversion technologies are available with different efficiency factors for different kinds of biomass feedstock. Therefore, it is necessary to reduce the complexity and the solution space to facilitate rapid evaluation and optimization of scenarios [14].

This chapter is structured as follows: First, the supply chain optimization problem is detailed in Sect. 30.2, followed by an explanation of the techniques for simplification and solution space reduction in Sect. 30.3. In Sect. 30.4, the mathematical model for the evaluation of scenarios is given. Moreover, some details on the routing approximations are given in Sect. 30.6, followed by a few remarks on the implementation as part of HeuristicLab [24, 25]. Finally, the benefits and shortcomings of this level of abstraction are discussed in Sect. 30.8.

30.2 Modeling

For the analysis of our European-wide simulation model, several stages or steps have to be considered. Figure 30.1 summarizes the steps in the supply chain for each combination of feedstock and region. Several different types of feedstock as well as several different decentralized and central conversion technologies have to be considered. In addition to financial optimization, the environmental impact should also be considered.

The granularity of this simulation model is on a regional level. Specifically, the geographical encoding scheme called “Nomenclature of Units for Territorial

¹ EU FP7 Project BioBoost, 282873: Biomass based energy intermediates boosting biofuel production, <http://www.bioboost.eu>.

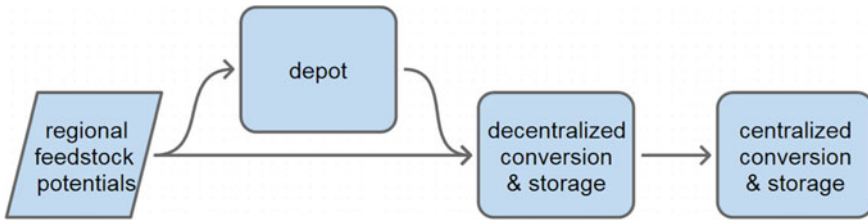


Fig. 30.1 Supply chain template: For each combination of feedstock and region, a supply chain might be initiated, each comprising feedstock acquisition, transport and handling, decentralized conversion, further transport and handling and final centralized conversion. Every *arrow* indicates transport and handling between the individual processing entities

Statistics” (NUTS) by the European Union [3] is used at level 3, which is the finest-grained standardized subdivision of regions in the European Union.

The considered decentralized conversion technologies can each use several different types of biological residues as feedstock such as straw, wood chips, malt spent grains and municipal waste. The decentralized conversion technologies employed in this study are fast pyrolysis [22], catalytic pyrolysis [16] and hydrothermal carbonization [12].

For all considered regions, the potentials of different types of feedstock have been investigated [20] broken down to NUTS level 3 and collected in a database [18].

In the following, the different stages of the simulation are outlined:

- *Feedstock acquisition* is the first stage in the simulation where different amounts of different types of feedstock are acquired in each region. The percentage of acquired feedstock for each region is a free parameter in the simulation model that is adjusted by the optimization process.
- The second stage in the simulation model is *transport and handling of the feedstock* to the decentralized facilities where they are further processed or, alternatively, to decentralized depots where feedstock is further aggregated to reduce transport costs. The optimization task in this stage is the selection of the transport targets as well as the selection of appropriate transport means for different types of feedstock and different distances.
- In the first *conversion and storage* stage of the simulation model, the feedstock is converted into *intermediate* energy carriers with increased energy density. For maintaining continuous operation of the conversion plants a certain amount of feedstock has to be stored locally. This is not only dependent upon the capacity of the plant but also depends on the feedstock type as well as seasonal availability and harvesting times. The optimization task in this step is the selection of appropriate conversion and storage facility types and capacities. During the simulation, the costs for construction, maintenance and operation of facilities have to be calculated.
- In a second iteration of the *transport and handling* stage, the intermediate energy carriers are now transported to central plants for conversion to fuels or energy. Similar to the first transport and handling stage, the transport targets as well as appropriate transport means have to be selected.

- Similarly, in the second *conversion and storage* stage, the intermediate products are transformed into end-products such as fuel or energy. Again appropriate types and sizes for the facilities and storages have to be selected.
- Depending on the number of simulated echelons, the succession of transport and handling followed by conversion and storage could continue several times. Moreover, in different regions, different combinations of echelons might be possible. For example, some regions directly deliver to decentralized conversion facilities, whereas other regions first deliver to a decentralized depot. This choice in the number of echelons depending on the transport infrastructure and converted amounts of feedstock has to be optimized as well.
- Finally, in a stage we called *sink evaluation* all produced products, emissions, wastes and costs have to be accumulated into a single fitness value that can be used to drive the optimization process.

30.3 Accelerations

As the simulation model harbors many free variables and consists of potentially complex simulation steps, several attempts were made to reduce to complexity, both in number of possible scenarios and in evaluation time. In the following, several attempts are presented to reduce calculation time for the simulation model.

The overall simulation contains many free variables that have to be optimized:

- feedstock acquisition percentages ($f \cdot r$ real numbers)
- transport routes for: (1) feedstock, (2) from depot and (3) for intermediate products ($3 \cdot f \cdot r$)
- conversion facility types and sizes for each region per feedstock (up to $f \cdot r$)

This gives on the one hand a real-valued optimization problem with $3 \cdot f \cdot r$ dimensions for utilization, plant size and storage size, where f is the number of feedstock types and r is the number of regions. For example, with nine feedstock types and 1,305 regions this gives a problem with more than 35,000 dimensions for this part. Additionally, the transport and plant choice is a combinatorial optimization problem that has a solution space size of $f \cdot r^3 \cdot t \approx 0.5 \times 10^{12}$ per echelon, which gives 2×10^{32} possibilities for the combinatorial part, where t is the number of technologies. For this reason, an important modeling consideration is the intelligent reduction of the solution space size.

An instructive insight is the interdependence between variables and the anticipated solution quality. By eliminating some of the free variables and letting the simulation itself calculate optimal values, the perceived solution space can be reduced without limiting the available choices. This is done by eliminating variables that would have suboptimal values in combination with other choices:

- Transport means can be automatically optimized according to the product type as well as source and target locations. If a certain transport means is cheaper only for shorter distances, it will automatically be used and, therefore, does not need to be selected by the optimization algorithm.

- The type and especially the size of both conversion facilities and storage can be inferred by the amount of transported feedstock to a potential plant location. This means, that plant type and size are indirectly defined by the amount of feedstock acquired and transported to a certain region. If no feedstock is brought, no plant is build and if feedstock is delivered a plant of the optimal size is built. This eliminates the need to evaluate obviously inferior solutions that contain plants of inappropriate size.

These implicit variables can now be removed from the practical solution space which limits the number of choices to just the feedstock acquisition and transport target selection per feedstock and echelon and reduces the size to $f \cdot r = 11,754$ dimensions for the real-valued part and $f \cdot r^2 \approx 12 \times 10^6$ combinations per echelon for the combinatorial part.

For all types of facilities and cost factors different scaling factors can be used to account for the economy of scale [21], where larger facilities become cheaper to build, operate and maintain compared to a linear scaling.

Another important simplification is the yearly aggregation of all costs. Instead of simulating the scenario progress step by step, all simulation steps are abstracted and calculated on a yearly basis. So, it does not matter when exactly a certain amount of feedstock is bought, transported or converted. Also for construction costs only yearly depreciation cost and investment costs need to be taken into account. This also makes scenarios comparable regardless of the actual depreciation time and drastically reduces computation times as the evaluation of a certain scenario reduces to a carefully selected sum of different yearly factors.

To further accelerate the simulation of individual scenarios, it also pays of to prepare the results of costly calculations in advance. For this reason, direct routing between sources and targets have not been calculated during simulation, but were simply replaced with precalculated expected values for route lengths between each pair of regions and also for intra-regional transport in case a facility is built in the same region as the feedstock is acquired.

30.4 Formulas

All intermediate and final results of the simulation model depend on the data loaded into the model, as well as two parameters:

- a matrix $T = (t_{p,r})_{p \in \mathcal{P}, r \in \mathcal{R}}$ of transport targets for each product and source region, where \mathcal{P} is the set of all products, i.e. feedstock, intermediates or final products, and \mathcal{R} is the set of all regions.
- and a matrix $U = (u_{p,r})$ of feedstock utilizations for each product and region.

The results of the simulation are a set of cost matrices and revenues for each product and region, as well as the total aggregated annual costs over all regions and products $C(T, U)$. The optimization objective is to minimize this total annual cost by tuning the transport targets T and utilizations U for all products and regions [13].

For each region and product, a valid transport target $r \in \mathcal{R}$ and the feedstock utilization $u \in [0..1]$ must be selected. Feedstock types in one region are independent, i.e. straw and wood chips from one region can be transported to mutually different destinations. There are no constraints on the target region such as restriction of transport to direct neighbors or within a certain distance. The full definition of the optimization problem is given in Table 30.1, a description of all identifiers and their units are given in Table 30.2 (cf. [13]).

This optimization problem is a mixed-integer nonlinear programming problem (MINLP) as it contains real-valued as well as integer-valued variables and the optimized function depends non-linearly on the variables [17].

Several optimization techniques including exact as well as heuristic algorithms can be used to solve this optimization problem. Here, we propose to use an evolutionary

Table 30.1 Formulas for calculating individual steps of the simulation, an explanation of the identifiers can be found in Table 30.2

quantities at source	$Q_{p,r}^s = P_{p,r}^f \cdot \mathbf{u}(\mathbf{p}, \mathbf{r})$
Feedstock acquisition cost	$C_{p,r}^f = P(p, r) \cdot Q_{p,r}^s$
Transport cost	$C_{p,r}^t = Q_{p,r}^s \cdot D_{r, \mathbf{t}(\mathbf{p}, \mathbf{r})} \cdot (P_{p,r}^t + P_{p, \mathbf{t}(\mathbf{p}, \mathbf{r})}^t) / 2$
Handling cost	$C_{p,r}^h = Q(p, s)^s \cdot (P_{p,r}^h + P_{p, \mathbf{t}(p, r)}^h)$
Quantities at target ^a	$Q_{p,r}^t = \sum_{s \in \mathcal{R}} Q_{p,s}^s \cdot \delta(\mathbf{t}(\mathbf{p}, \mathbf{s}), r)$
Optimized plant capacities	$CP_{p,r}^c = Q(p, r)^t / P_p^{\text{uf}}$
Variable conversion cost	$C_{p,r}^v = Q_{p,r}^t \cdot P_{p,r}^{\text{cv}}$
Scaled maintenance cost	$C_{p,r}^m = (CP_{p,r}^c / P_p^{\text{ds}})^{\sigma^{\text{mc}}} \cdot P_{p,r}^{\text{mc}}$
Scaled construction cost	$C_{p,r}^c = (CP_{p,r}^c / P_p^{\text{ds}})^{\sigma^{\text{cc}}} \cdot P_{p,r}^{\text{cc}}$
Financing cost	$C_{p,r}^i = \lambda \cdot C_{p,r}^c$
Storage capacity	$CP_{p,r}^s = CP_{p,r}^c \cdot P_p^{\text{ss}} / 365$
Storage cost	$C_{p,r}^s = CP_{p,r}^s \cdot P_p^{\text{sc}}$
Produced quantities ^b	$Q_{p',r}^s = Q_{p,r}^t \cdot \gamma_{p,p'}^{\text{cy}}$
Revenue	$R_{p,r} = Q_{p,r}^s \cdot S_p^s + Q_{p,r}^t \cdot S_p^t$
Total cost ^c	$C = \sum_{p \in \mathcal{P}} \sum_{r \in \mathcal{R}} \left[\sum_{f \in \mathcal{F}} C_{p,r}^f \right] - R_{p,r}$
Purchase price	$P(p, r) = P_p^b \left(\frac{\exp(\sigma_p^b) - P_p^m}{\exp(\sigma_p^b) - 1} + \frac{\exp(\sigma_p^b \mathbf{u}(\mathbf{p}, \mathbf{r})) (P_p^m - 1)}{\exp(\sigma_p^b) - 1} \right)$

^aTo select only those amounts for which a transport is scheduled, the Kronecker delta function $\delta(i, j)$ is used. Its value is 1 when $i = j$ and zero otherwise

^bfor next logistics echelon

^c \mathcal{F} is the set of all cost factors

Table 30.2 Identifiers used in the supply chain optimization problem

Identifier	Index variables	Unit
p	Product	[product]
r	Region	[region]
<i>Identifier</i>	<i>Facts</i>	<i>Unit</i>
\mathcal{R}	Set of regions	[{region}]
\mathcal{P}	Set of products	[{product}]
S_p^s, S_p^t	Product p 's sales price (at source s or target t)	[€]
$D_{src,dest}$	Distance matrix (route lengths)	[km]
$P_{p,r}^f$	Feedstock potentials for product p and region r	[t/a]
$P_{p,r}^{cv}$	Product conversion cost (per ton of feedstock)	[€/t]
$P_{p,r}^{mc}$	Plant maintenance cost	[€/a]
$P_{p,r}^{cc}$	Plant construction cost	[€/a]
$P_{p,r}^{sc}$	Product / plant storage cost	[€/t/a]
P_p^{ds}	Plant design size	[t/a]
P_p^{uf}	Plant design utilization factor	[t/a]
P_p^{ss}	Product/plant safety stock duration	[d/a]
$P_{p,r}^t$	Regional product transport costs	[€/t/km]
$P_{p,r}^h$	Regional product handling costs	[€/t]
$\gamma_{p,p'}^{cy}$	Mass rate for conversion of product p into p'	[t/t]
σ^{cc}	Size-dependent construction cost scaling factor	[1]
σ^{mc}	Size-dependent maintenance cost scaling factor	[1]
λ	interest rate	[%/a]
<i>Identifier</i>	<i>Variables (optimization targets)</i>	<i>Unit</i>
$U = (u_{p,r})$	Feedstock utilizations	[1/a]
$T = (t_{p,r})$	Product transport targets	[region]
<i>Identifier</i>	<i>Intermediate results</i>	<i>Unit</i>
$Q_{p,r}^s, Q_{p,r}^t$	Quantities or amounts at source and at target	[t/a]
$CP_{p,r}^c$	Converter capacities	[t/a]
$CP_{p,r}^s$	Storage capacities	[t/a]
$C_{p,r}^c$	Construction cost	[€/a]
$C_{p,r}^f$	Feedstock purchase cost	[€/a]
$C_{p,r}^i$	Investment cost	[€/a]
$C_{p,r}^m$	Maintenance cost	[€/a]

(continued)

Table 30.2 (continued)

<i>Identifier</i>	<i>Intermediate results</i>	<i>Unit</i>
$C_{p,r}^s$	Storage cost	[€/a]
$C_{p,r}^t$	Transportation cost	[€/a]
$C_{p,r}^h$	Handling cost	[€/a]
$C_{p,r}^v$	Variable conversion cost	[€/a]
$P(p, r)$	purchase price	[€]
P_p^b	Base price per ton of product	[€/t]
P_p^m	Price multiplier for full market saturation	[1]
σ_p^p	Price curve scaling factor	[1]
$R_{p,r}$	Revenue	[€/a]
<i>Identifier</i>	<i>Result</i>	<i>Unit</i>
C	Total cost	[€/a]

algorithm for solving this problem [11]. We are not primarily concerned about finding a globally optimal solution as such a solution would be optimal only with respect to the collected input data which could be inaccurate or outdated. Therefore, it is possible to use meta-heuristics such as evolutionary algorithms, as we are satisfied with feasible nearly-optimal solutions [6]. One advantage of this approach is that near-optimal solutions can be found with less effort, and therefore, optimization runtimes are much shorter. This facilitates experimentation with different plausible scenarios such as different feedstock or product prices.

For solving the above supply chain optimization problem, a specialized evolutionary algorithm has been formulated that is a combination of an evolutionary strategy [1] and a genetic algorithm [7]. This is motivated by the fact that solution candidates have two components (1) the transport targets that modeled as an integer matrix and (2) the feedstock utilizations that are modeled as a real-valued matrix. In the proposed algorithm, different kinds of specialized evolutionary operators (crossover and mutation) are used for these two components of the solution candidates. Figure 30.2 shows the two components of solution candidates for an example with only one feedstock type (straw) and two products (straw and biosyncrude).

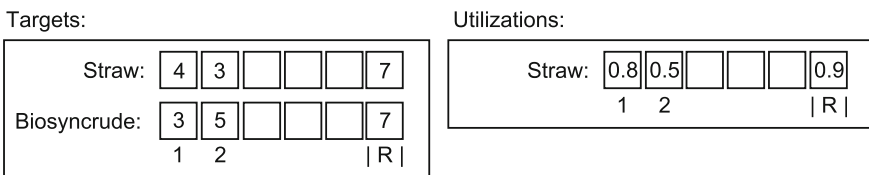


Fig. 30.2 A solution candidate consists of two components: the matrix of transportation targets (*left*) and the matrix of feedstock utilizations (*right*)

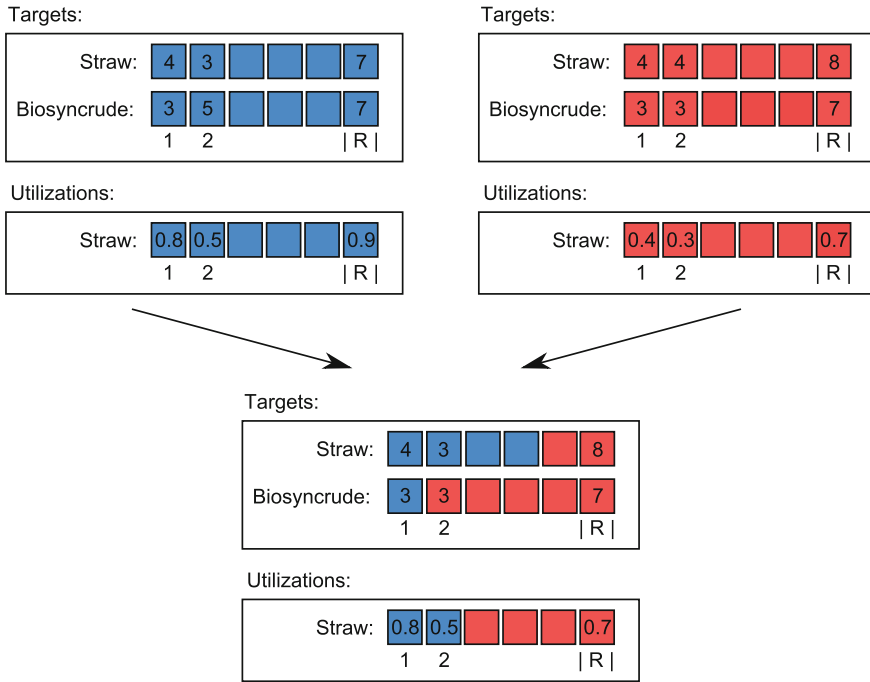


Fig. 30.3 The crossover operator creates a new solution candidate taking information from both parent solution candidates. Each component of the solution encoding and each row vector is crossed independently

Figure 30.3 shows an example of applying a crossover operation independently on the components of two solution candidates. The two matrices for feedstock utilizations and transport targets are crossed over independently using specific crossover operators. Each row in the matrix represents utilizations or transportation targets for one particular type of product. The products which are assumed to be independent in the optimization model, therefore, potentially different crossover points are chosen independently for each product type.

30.5 Solution Space Discovery

The formulation of supply chains can be a complex task, especially in the light of different or competing options, where it should rather be called a supply graph. For this reason, the formulation of a supply chain is broken into individual components. In our model these components are the following:

- *Products* are defined by their name and can have an attached buying and/or selling price.
- *Locations* comprise feedstock potential and can accommodate conversion facilities.

- *Logistics* describe the cost and environmental impact of handling and transporting products depending on their amount and distance. Moreover, transports need to be calculated over a distance matrix.
- *Conversions* describe stationary facilities that consume products, energy and money and, in return, produce other products at a certain location. For this purpose they also need to store their input products for a certain time to accommodate for supply variances.

The question remains how these components are connected to each other. Conceptually, for humans, the answer to this question is obvious. Transports are defined for certain products that are used in conversions and the final products, possibly after several logistic echelons are sold. In a naive implementation, all possibilities are tried in every step. Although certainly this would work, it would be rather inefficient as many infeasible configurations are tried. Therefore, we have implemented a simple algorithm, that discovers the solution space. For this purpose, it tries to convert and transport all products, however, only a single time. In every step, the algorithm examines how a possible supply chain could continue. On the one hand, an output product could be directly sold, on the other hand, it could be the input to another conversion process. However, this conversion process must then produce valuable products which can then be sold. In this way, possible parts of the supply chain are discovered that will ultimately lead to a profit and other options will be excluded which reduces the size of the solution space to only feasible supply chains. Moreover, in certain situations there might be a choice between several possibilities, for example if a product is available as input to two different conversion process. In this case, this choice has to be included as part of the solutions space.

Many frequently-used crossover and mutation operators in evolutionary algorithms are undirected and do not use semantic information when creating or manipulating solution candidates. As described in Sect. 30.4, transport targets are encoded as integer vectors. However, in this particular problem setting, different transport targets are not independent, because there is an underlying semantic meaning for the elements representing NUTS-3 regions. In particular, the distance between two regions, or the information whether two regions are direct neighbors, is strongly related to the effect on the solution quality when changing a target region. If mutation changes a target region and replaces it by a direct neighbor, the effect on the solution quality is smaller than than using a distant region.

The semantic information can be used in the algorithm to bias the search in a way to increase the likelihood of sampling similar solutions, which, in a certain sense, makes the algorithm more greedy leading to faster convergence. We implemented several such semantic mutation operators:

- The *distance-based mutator* uses the distance matrix to determine the likelihood for all possible replacement regions. The likelihood for a replacement region is inverse proportional to the distance to the replaced region.
- The *neighbor-based mutator* uses an adjacency matrix which encodes direct neighborhood relations between all regions and selects the replacement region randomly

from the set of all neighbors of the replaced region with a bias towards closer regions according to neighborhood class.

- The *common target mutator* changes the transport targets for a whole set of regions that currently have the same target for a certain feedstock to a different (but common) target region.
- *common target swapping mutator* chooses two sets of regions where each set shares the same target and mutually exchanges the target for some of the regions contained in each set.
- The *target merging mutator* is similar to the *distance-based mutator* and uses the distance matrix to increase the likelihood of replacing a target region with a nearby region. First the source regions are grouped by the target region. For two randomly selected groups a new target region is calculated that has minimal distance to both groups' target regions. Finally, the target region of a random subset of source regions in both groups is set to the newly selected target region.

30.6 Route Length Estimation

To further accelerate the simulation, direct route length estimation during the simulation is by far too slow on the network the size of the European road network. Therefore, road lengths have been calculated in advance. For this purpose, 30 random starting points have been selected in each region and used to create routes between each pair of regions. Figure 30.4 shows the mapping of random points inside the regions to the nearest road section.

In total, more than 12,000,000 routes have been calculated. The underlying road network has been extracted from open street maps data² and was fed into the software “monav” a routing engine that uses contraction hierarchies [5] to come up with optimal routes in a very short time. For every region to every other region, the average of 30 routes has been used as a measure of the transport distance that is likely to occur between these regions. Contraction hierarchies are much faster than traditional algorithms like the A* algorithm and also more accurate than typical heuristics used inside satellite navigation devices.

30.7 Implementation

The supply chain optimization problem as described above has been implemented as an additional plugin for HeuristicLab,³ an environment for heuristic optimization that provides a graphical user interface and ready-to-use implementations for many

² <http://openstreetmap.org>.

³ <http://dev.heuristiclab.com>.

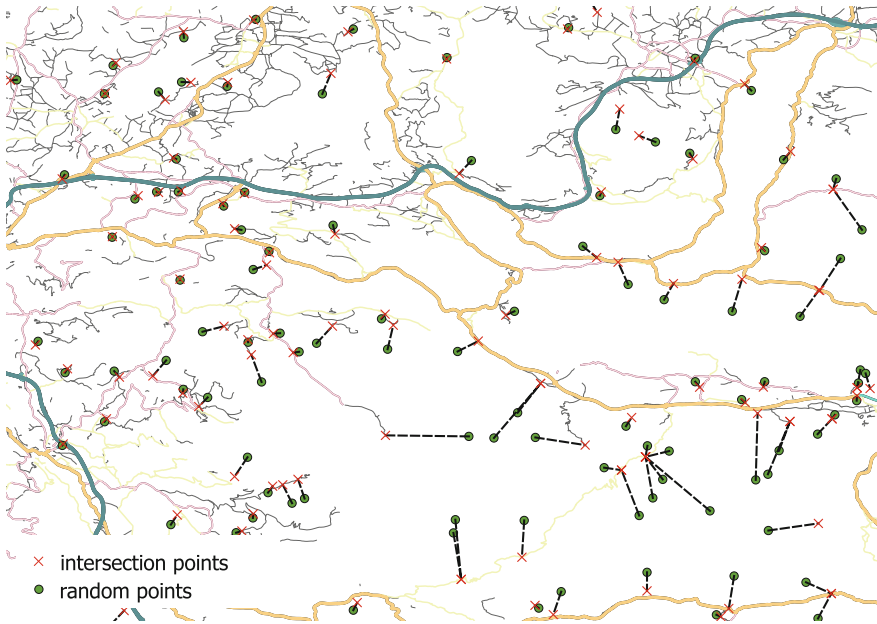


Fig. 30.4 Mapping random points to road sections

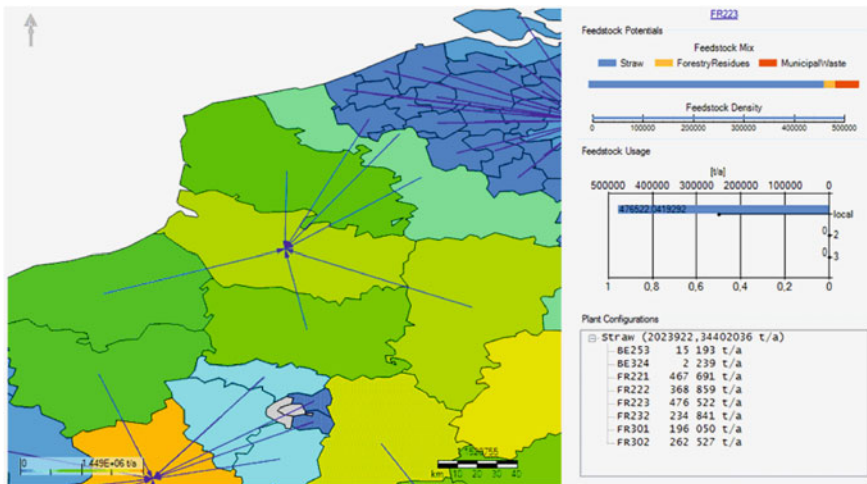


Fig. 30.5 Interactive visualization of an optimization supply chain configuration

well-known heuristic optimization algorithms [25]. The new plugin provides specialized classes and user interface components for the configuration of the supply chain problem and for loading the necessary data files as well as user interface components for the visualization and analysis of optimization results and solutions.

Simulation parameters can be imported from a YAML-file⁴ or edited directly via the graphical user interface. After the configuration of the problem, any supported optimization algorithm which is available in HeuristicLab can be used to find optimized solutions. This facilitates experimentation with different kinds of algorithms and different algorithm configurations. Feedstock utilizations, transport targets, and plant capacities of optimized scenarios are visualized on an interactive map which can be used to explore and analyze the optimization results. Figure 30.5 shows a screen shot of an exemplary result of an optimization run.

30.8 Discussion

In this chapter we described a concise supply chain optimization model that has been tested on a large European-wide decentralized biomass logistics network. Optimization of a European-wide logistics model is only possible if the model is expressed in a concise manner so that different scenarios can be evaluated efficiently. The concise simulation model for biomass logistics has been developed in the EU-funded project BioBoost⁵ for biomass based energy intermediates boosting biofuel production. In this model several notable shortcuts are used to improve efficiency of evaluations. First, we pre-calculated a distance matrix for each pair of NUTS-3 regions based on OpenStreetMap data. Second, the number of free parameters in the supply chain model has been reduced as much as possible to reduce the complexity of the solution space of the optimization problem. The only remaining parameters are the feedstock utilizations for each feedstock type and region and the transport targets for each feedstock or product and region. All other central parameters of the model such as the plant capacities are determined solely by these input values and thus are intermediate results in the model. Third, the model has been implemented using a generic and flexible design which allows to add or remove feedstock, intermediate products as well as final products. Additionally the set of considered regions can be filtered to restrict the optimization to a smaller set of regions.

The supply chain simulation model and its operators for evolutionary algorithms have been implemented in HeuristicLab, an open-source environment for heuristic optimization that provides implementations of well-known optimization algorithms.

An interactive graphical user interface for the visualization and analysis of optimization solutions has been implemented to support exploration of different scenarios and analysis of effects of external factors such as feedstock prices or energy prices.

⁴ <http://yaml.org/spec/1.2/spec.html>.

⁵ This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 282873.

References

1. Beyer, H.-G., Schwefel, H.-P.: *Nat. Comput.* **1**(1), 3–52 (2002)
2. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. *Manag. Sci.* **6**(1), 80–91 (1959)
3. European Parliament, Council of the European Union. On the establishment of a common classification of territorial units for statistics (NUTS). Official Journal of the European Union. Commission Regulation (EC) No 105/2007 (amending Regulation (EC) 1059/2003), (2008)
4. Evans, G.: International biofuels strategy project. liquid transport biofuels—technology status report, NNFCC 08-017. Technical report, National Non-Food Crops Centre, (2008)
5. Geisberger, R.: Contraction hierarchies: Faster and simpler hierarchical routing in road networks. Master's thesis, Institut for Theoretical Computer Science, University of Karlsruhe, (2010)
6. Glover, F., Kochenberger, G.A.: *Handbook in Metaheuristics*. Springer (2003)
7. Goldberg, D.E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley (1989)
8. Gosavi, A.: *Simulation-Based Optimization*. Springer (2003)
9. Hart, P.E., Nilsson, N.J., Raphael, B.: A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern.* **4**(2), 100–107 (1968)
10. Inderwildi, O. R., King, D. A.: Quo vadis biofuels? *Energy Environ. Sci.* **2**, 343–346 (2009)
11. De Jong, K.A.: *Evolutionary Computation*. MIT Press, Cambridge (2006)
12. Kruse, A., Funke, A., Titirici, M.-M.: Hydrothermal conversion of biomass to fuels and energetic materials. *Current Opin. Chem. Biol.* **17**(3), 515–521 (2013). Next generation therapeutics energy
13. Kronberger, G., Pitzer, E., Brunner, S., Hutterer, S.: Protocol On Milestone 4, Deliverable 4.2, EU FP7 Project BioBoost (282873), BioBoost Consortium. <http://bioboost.eu/results.php> (2014)
14. Law, A.M.: *Simulation Modeling and Analysis*. McGraw-Hill, New York (2007)
15. Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G., Shmoys, D.B.: (Eds.) *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. Wiley (1985)
16. Lappas, A.A., Samolada, M.C., Iatridis, D.K., Voutetakis, S.S., Vasalos, I.A.: Biomass pyrolysis in a circulating fluid bed reactor for the production of fuels and chemicals. *Fuel* **81**(16), 2087–2095 (2002)
17. Nemhauser, G.L., Wolsey, L.A.: *Integer and Combinatorial Optimization*, vol 18. Wiley, (1988)
18. Pudelko, R., Brozecka-Walker, M., Faber, A.: The Feedstock Potential Assessment for EU-27, Deliverable 1.2, EU FP7 Project BioBoost (282873), BioBoost Consortium. <http://bioboost.eu/results.php> (2014)
19. Pitzer, E., Kronberger, G.: Accelerated simulation-based optimisation using pre-calculation, aggregation, implicit solution spaces and dynamic solution space reduction. In: APCASE 2014, Extended Abstracts, pp. 85–86 (2014)
20. Rozakis, S., Kremmydas, D., Pudelko, R., Borzecka-Walker, M., Faber, A.: Straw potential for energy purposes in Poland and optimal allocation to major co-firing power plants. *Biomass Bioenergy* **58**, 275–285 (2013)
21. Sullivan, A., Sheffrin, S.M.: *Economics: Principles in Action*. Prentice Hall (2006)
22. Trippe, F., Fröhling, M., Schulterman, F., Stahl, R., Henrich, E.: Techno-economic analysis of fast pyrolysis as a process step within biomass-to-liquid fuel production. *Waste Biomass Valorization* **1**(4), 415–430 (2010)
23. Toth, P., Vigo, D.: *The Vehicle Routing Problem (Monographs on Discrete Mathematics and Applications, vol. 9)* Society for Industrial and Applied Mathematics (2001)
24. Wagner, S.: *Heuristic Optimization Software Systems—Modeling of Heuristic Optimization Algorithms in the HeuristicLab Software Environment*. PhD thesis, Johannes Kepler University, Linz, Austria (2009)
25. Wagner, S., Kronberger G., Beham A., Kommenda M., Scheibenpflug A., Pitzer E., Vonolfen S., Kofler M., Winkler S., Dorfer V., Affenzeller, M.: Architecture and Design of the HeuristicLab Optimization Environment. *Topics in Intelligent Engineering and Informatics*, chapter 10, vol. 6, pp 197–261. Springer (2014)