James J. Park
Laurence T. Yang
Changhoon Lee (Eds.)

Communications in Computer and Information Science 185

Future Information Technology

6th International Conference, FutureTech 2011
Loutraki, Greece, June 2011
Proceedings, Part II

Part 2

Springer

Communications
in Computer and Information Science 185

James J. Park Laurence T. Yang
Changhoon Lee (Eds.)

Future Information Technology

6th International Conference, FutureTech 2011
Loutraki, Greece, June 28-30, 2011
Proceedings, Part II

Springer

Volume Editors

James J. Park
Seoul National University of Science and Technology
Department of Computer Science and Engineering
172 Gongreung 2-dong, Nowon-gu, Seoul, 139-743, Korea
E-mail: parkjonghyuk1@hotmail.com

Laurence T. Yang
St. Francis Xavier University
Department of Computer Science
Antigonish, NS, B2G 2W5, Canada
E-mail: ltyang@stfx.ca

Changhoon Lee
Hanshin University
411 Yangsandong, Osan-si, 447-791, Korea
E-mail: chlee@hs.ac.kr

ISSN 1865-0929 e-ISSN 1865-0937
ISBN 978-3-642-22308-2 e-ISBN 978-3-642-22309-9
DOI 10.1007/978-3-642-22309-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011930821

CR Subject Classification (1998): C.2, H.4, H.3, I.2, D.2, H.5

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Message from the General Chairs

On behalf of the Organizing Committees, it is our pleasure to welcome you to the 6th International Conference on Future Information Technology (FutureTech 2011), which was held in Loutraki, Greece, June 28–30, 2011.

In past 25 years or so, information technology (IT) as influenced and changed every aspect of our lives and our cultures. Without various IT-based applications, we would find it difficult to keep information stored securely, to process information efficiently, and to communicate information conveniently. In the future world, IT will play a very important role in the convergence of computing, communication, and all other computational sciences and applications. IT will also influence the future world's various areas, including science, engineering, industry, business, law, politics, culture, medicine, and so on.

FutureTech 2011 added to the series of the highly successful International Conference on Future Information Technology (FutureTech 2010, Busan, Korea, 2010), previously held as International Symposium on Ubiquitous Applications & Security Services (UASS 2009, USA, 2009), UASS 2008 (Okinawa, Japan, 2008), UASS 2007 (Kuala Lumpur, Malaysia, 2007), and UASS 2006 (Glasgow, Scotland, UK, 2006).

This year we were proud not only to have a distinguished set of the following invited speakers, but also the distinguished support of James J. (Jong Hyuk) Park as the FTRA president.

- Habib F. Rashvand, PhD, School of Engineering, University of Warwick, UK
- J.C. (Hans) van Vliet, PhD, Department of Information Management and Software Engineering (IMSE), Vrije Universiteit, The Netherlands
- Peter Sloot, PhD, Full Professor of Computational Science, Scientific Director of the Informatics Institute, The Netherlands

We are sure you will find many exciting and stimulating papers in the proceedings, and we would like to thank the General-Vice Chair, Changhoon Lee, and the Program Chairs, Serge Chaumette, Jin Kwak, and Konstantinos Markantonakis, for collecting an extremely strong set of papers across a wide range of topics.

Finally, we want to extend our thanks to all our Chairs and Committees who are truly responsible for making this conference a success through their generous contributions of time and energy and are listed in the following pages. We also sincerely thank the EAPIC-IST 2011, ISCC 2011, ITSA 2011, F2GC 2011, MUST 2011, and SocialComNet 2011 Workshop Chairs for their kind cooperation.

Damien Sauveron
Sang-Soo Yeo

Message from the Program Chairs

Welcome to the proceedings of the 6th International Conference on Future Information Technology (FutureTech 2011), held in Loutraki, Greece, June 28–30, 2011.

Future information technology would stand for all of continuously evolving and converging information technologies, including digital convergence, multimedia convergence, intelligent applications, embedded systems, mobile and wireless communications, bio-inspired computing, grid and cloud computing, Semantic Web, user experience and HCI, security and trust computing and so on, for satisfying our ever-changing needs.

FutureTech 2011 intended to foster the dissemination of state-of-the-art research in all future information technology areas, including their models, services, and novel applications associated with their utilization.

We received 274 papers with the following topics from around the globe and after a rigorous review process, we accepted 123 papers, for a competitive acceptance rate of 45%.

- Track 1. Hybrid Information Technology
- Track 2. High Performance Computing
- Track 3. Cloud Computing
- Track 4. Ubiquitous Networks and Wireless Communications
- Track 5. Digital Convergence
- Track 6. Multimedia Convergence
- Track 7. Intelligent and Pervasive Applications
- Track 8. Security and Trust Computing
- Track 9. IT Management and Service
- Track 10. Bioinformatics and Bio-Inspired Computing

We believe that the selected papers provide novel ideas, new results, work in progress and state-of-the-art techniques in this field, stimulating future research activities.

We would like to express our appreciation to all the contributors and authors for the submissions to FutureTech 2011. Special thanks are due to the members of the Technical Program Committee and all the external reviewers for their invaluable help with the review process. Finally, we would like to thank everyone who participated in the conference, and acknowledge the members of the Organizing Committee for their contributions and in particular Leonie Kunz at springer for the great help and hard work.

Serge Chaumette
Jin Kwak
Konstantinos Markantonakis

FutureTech 2011 Organization

Steering Chairs

James J. (Jong Hyuk) Park	Seoul National University of Science and Technology, Korea
Laurence T. Yang	St. Francis Xavier University, Canada
Hamid R. Arabnia	The University of Georgia, USA

General Chairs

Damien Sauveron	University of Limoges, France
Sang-Soo Yeo	Mokwon University, Korea

General Vice-Chair

Changhoon Lee	Hanshin University, Korea

Program Chairs

Serge Chaumette	LaBRI, University of Bordeaux, France
Jin Kwak	Soonchunhyang University, Korea
Konstantinos Markantonakis	Royal Holloway University of London, UK

Steering Committee

Han-Chieh Chao	National Ilan University, Taiwan
Shu-Ching Chen	Florida International University, USA
Stefanos Gritzalis	University of the Aegean, Greece
Vincenzo Loia	University of Salerno, Italy
Yi Mu	University of Wollongong, Australia
Witold Pedrycz	University of Alberta, Canada
Wanlei Zhou	Deakin University, Australia
Young-Sik Jeong	Wonkwang University, Korea

Advisory Committee

Ioannis G. Askoxylakis	FORTH-ICS, Greece
Hsiao-Hwa Chen	National Sun Yat-Sen University, Taiwan
Jack J. Dongarra	University of Tennessee, USA
Javier Lopez	University of Malaga, Spain
Bart Preneel	Katholieke Universiteit Leuven, Belgium
Harry Rudin	IBM Zurich Research Laboratory, Switzerland

Waleed W. Smari	University of Dayton, USA
Doug Tygar	University of California Berkeley, USA
Shiuh-Jeng Wang	Central Police University, Taiwan
Doo-soon Park	SoonChunHyang University, Korea

Workshop Co-chairs

Sabrina De Capitani di Vimercati	Università degli Studi di Milano, Italy
Naveen Chilamkurti	La Trobe University, Australia

Publicity Chairs

Deok-Gyu Lee	ETRI, Korea
Theo Tryfonas	University of Bristol, UK
Karim El Defrawy	University of California, Irvine, USA
Sang Yep Nam	Kookje College, Korea
Young-June Choi	Ajou University, Korea

International Liaison Chairs

Jiankun Hu	RMIT University, Australia
Louis Marinos	ENISA, EU
Qishi Wu	University of Memphis, USA

Publication Chair

Jose Maria Sierra	Universidad Carlos III de Madrid, Spain

Local Arrangements Co-chairs

Pierre-Francois Bonnefoi	University of Limoges, France
Pierre Dusart	University of Limoges, France

Web and System Management Chair

Yang Sun Lee	Chosun University, Korea

Track Co-chairs

Track 1. Hybrid Information Technology

Umberto Straccia	ISTI - C.N.R., Italy
Malka N. Halgamuge	The University of Melbourne, Australia
Andrew Kusiak	The University of Iowa, USA

Track 2. High-Performance Computing

Waleed W. Smari	University of Dayton, Ohio, USA
Omer Rana	Cardiff University, UK
Satoshi Matsuoka	Tokyo Institute of Technology, Japan

Track 3. Cloud Computing

Geoffrey Fox	Indiana University, USA
Pascal Bouvry	University of Luxembourg, Luxembourg
Andrzej M. Goscinski	Deakin University, Australia

Track 4. Ubiquitous Networks and Wireless Communications

Ivan Stojmenovic	University of Ottawa, Canada
Tuomo Tuikka	Research Centre, Finland
Jiannong Cao	Hong Kong Polytechnic University, Hong Kong

Track 5. Digital Convergence

Gabriel Jakobson	Altusys Corp., USA
Vladimir Zaborovsky	Polytechnic University of Saint-Petersburg, Russia
Johnnes Arreymbi	University of East London, UK

Track 6. Multimedia Convergence

Jonathan C.L. Liu	University of Florida, USA
Sofia Tsekeridou	Athens Information Technology, Greece
Chia-Chen Lin	Providence University, Taiwan

Track 7. Intelligent and Pervasive Applications

Hongli Luo	Indiana University-Purdue University Fort Wayne, USA
Isabelle Demeure	TELECOM ParisTech, France
Kuan-Ta Chen	Institute of Information Science, Academia Sinica, Taiwan

Track 8. Security and Trust Computing

Luis F.G. Sarmenta	MIT Media Lab, USA
Jose Maria Sierra	Universidad Carlos III de Madrid, Spain
Chan Yeob Yeun	Khalifa University of Science, Technology & Research (KUSTAR), UAE

Track 9. IT Management and Service

Youakim Badr	INSA-Lyon, France
Ying Jiang	Wuhan University, China
Hang-bae Chang	Daejin University, Korea

Track 10. Bioinformatics and Bio-Inspired Computing

Frederic Guinand	University of Le Havre, France
Sun-Yuan Hsieh	National Cheng Kung University, Taiwan
Sheng-Lung Peng	National Dong Hwa University, Taiwan

Table of Contents – Part II

Social Computing, Network, and Services

Forensics for Future Generation Communication Environments

Intelligent Transportation Systems and Applications

Multimedia and Semantic Technologies

Information Science and Technology

Table of Contents – Part I

Future Information Technology (24 Papers)

IT Service and Cloud Computing (43 Papers)

A User Supporting Personal Video Recorder Based on a Generic Bayesian Classifier and Social Network Recommendations

Benedikt Engelbert, Malte Blanken, Ralf Kruthoff-Brüwer, and Karsten Morisse

Research Group Next Generation PVR, Faculty of Engineering and Computer Science
University of Applied Sciences Osnabrück, Albrechtstr. 30,
49076 Osnabrück, Germany
{B.Engelbert,M.Blanken,R.Kruthoff-Bruewer,
K.Morisse}@hs-osnabrueck.de

Abstract. The handling of the enormous offer of TV content is a challenge for every TV user since the number of receivable channels highly increased in recent years. Regular TV guides or newspapers present just a limited number of channel timetables and the user won't spend additional time to discover the rest of timetables by using other media e.g. the Internet. Hence, the user only focuses on favored channels and interesting content on others won't be recognized. The result is that the user might not select the most appropriate content regarding his or her interests. Assistive systems and tools are desirable to counteract this problem. We extend a Personal Video Recorder (PVR) with a recommendation system based on a Bayesian classifier and a collaborative approach using social networks like Facebook or Twitter. In case of the Bayesian classifier, the system is analyzing the user's watching behavior to generate personalized TV program recommendations. With the social network component the user receives recommendations from acquaintances and friends. The recommendations are automatically stored on the system's internal hard disc drive for the user to watch. This paper presents the current state of development by introducing the system's architecture and implemented recommendation mechanisms.

Keywords: personal video recorder, generic recommendation system, Bayesian classifier, Social network recommendations.

1 Introduction

The number of receivable TV channels highly increased in recent years. The satellite provider Astra holds up to 1700 TV channels just for the region of Germany [1]. It is therefore not surprising that users are often overstrained in handling the enormous offer of content. To figure out the most interesting content regarding to the user's interests is one of the biggest challenges. TV guides and newspapers present just a limited number of channel timetables, that the user is forced to discover the rest of channels using other media e.g. the Internet. This takes time the user will not invest. So the user focuses on favored or popular TV channels, which means that appropriate

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 1–8, 2011.

content on other channels will be ignored. To counteract this problem, the user needs assistive systems and tools especially for the duty of filtering to most appropriate content regarding the user's interests. Recommendation systems face the task to provide personalized suggestions for any type of content on the basis of a user's likes and dislikes. E.g. Amazon.com [2] provides a recommendation system to offer users new items in their online store like books, CDs etc. We consider options of implementing several recommendation techniques on a commercial PVR. It is the aim to generate personalized TV content recommendations, where the recommended content will be stored on an internal hard disc drive. For the recommendation process two different recommendation techniques are used. A Bayesian classifier was implemented as a machine-learning component. The classifier provides machine-generated recommendations on the basis of the user's consumption behavior. Furthermore a social network component was implemented. Users can take advantage of the social networks Facebook [3] and Twitter [4] to receive recommendations from acquaintances and friends.

This paper presents the current results of development of the NG-PVR system (Next Generation – Personal Video Recorder) and is organized as follows: section 2 presents related work on recommendation systems in the context of multimedia and TV. Work about social networks within recommendation systems will be considered as well. Section 3 presents the system architecture and the connection between the recommendation components. Section 4 describes both components in detail. The paper ends with a conclusion and some future work.

2 Related Work

Recommendation systems exist in several application areas, where various content is recommended. As mentioned before, Amazon.com recommends items like books or CDs in their online store. But also in the context of multimedia content recommendation systems are available. For instance the video broadcaster YouTube [6] analyzes user activities on the YouTube website to generate recommendations for related video content [5]. In [7], a mobile recommendation application is presented to make the huge video content of YouTube also manageable on devices with limited resources e.g. smart phones. The application implements a Bayesian classifier, which is one of the main recommendation mechanisms also in our work. Already in 1997 a Bayesian classifier based recommendation system was suggested in the work of Pazzani and Billsus [14]. In this approach web pages were classified as relevant or irrelevant on the basis of user rated web pages. Also for TV content several approaches exist. Already in 2002, an overview of personalized TV program was released [8]. General ideas and techniques of development for personalized TV recommendation are described. Some of the presented ideas are also considered in our work. AVATAR [12] is an implemented recommendation system of personalized TV content which has been conceived as an application conform to the Multimedia Home Platform (MHP) standard. The MHP standard was developed by the Digital Video Broadcasting (DVB) Project and is the collective name for a set of specifications and aims to work across all DVB transmission technologies [13]. With MHP it is possible to integrate java-based applications into the native system of a PVR. To suggest

personalized TV content, the AVATAR system implements a Bayesian classifier as well. TiVo is a commercial Set-Top-Box system with an integrated recommendation service. The system uses a combination of a content-based Bayesian algorithm and a collaborative-filtering approach [18]. Unlike the work described in this paper, recommendations are mainly generated from the collaborative approach, where item-to- item (e.g. show-to-show) relations are calculated from the user's feedback for TV content. For this, TiVo uses a data base with 100 million of user ratings for up to 30.000 TV shows and movies. A large number of active users is needed for the collaborative approach. In the TiVo system the Bayesian algorithm is just used to overcome the problem for new shows without any user feedback.

Social networks open up new areas of application since they got more and more popular. The huge number of users in better-known social networks might be a suitable data base for a collaborative recommendation system. Several approaches for social network based recommendation systems exist. In [9] a model of a trust-based recommendation system on a social network is presented. The systems filters information for agents based on the agents' social network and trust relationship in an automated and distributed fashion. FilmTrust is a semantic web-based social network, which creates predictive movie recommendations [10][11]. The similarity in this system is also calculated on trust ratings within the social network. The aim of the research project MovieLens is also to generate personalized movie recommendations [16][17]. For this, a collaborative filtering approach was implemented as well. For FilmTrust and for MovieLens applies that an initial indication of favored content is needed. On a web-based interface it is possible to rate several movies for creation of a required data base.

The difficulty of an initial user profile is also discussed in this paper, whereas the creation is optional for the PVR system (cp. Section 3). In case of using social networks in recommendation systems, this paper presents new approaches, where the use of existing social networks is considered (cp. Section 4.2). As the basis mechanism for machine-generated recommendations a Bayesian classifier is used as an established method in recommendation systems (cp. Section 4.1).

3 System Architecture

The recommendation system suffices a generic architecture, so it is possible to adapt the system for the use in other contexts. A suitable data model is needed to keep the system generically. For that, we defined Domain Objects as a generic data model. Domain Objects represent content objects, which can be described by metadata in a unique way. In case of the PVR System, a Domain Object is a single TV content (e.g. show, movie, etc.). The metadata of the PVR system is derived from an Electronic Program Guide (EPG), which might be broadcasted concurrently to the TV content (e.g. DVB-SI) or is a dedicated system service. Our system is based on a dedicated service and uses the following attributes: *TV Channel, Title, Subtitle, Category (e.g. Movie, show), Genre (e.g. Comedy), Actors, Description, Year.*

The recommendation system recommends instances of a Domain Object. The recommendation process is based on User Profiles. We differ between an initial user profile and an adaptive user profile. By using a simple weight function a Domain Objects gets a weight like or dislike and fall in one of the two classes (like or dislike)

within the user profile. A Domain Object is rated with a certain probability how it is liked or disliked. The probability is calculated within the recommendation process, where attributes of already rated Domain Objects and the new Domain Objects are compared against each other. The classification is done by a Bayesian classifier, which is described in section 4.1. Both user profiles serve as the data basis for new recommendations. As mentioned before, a set of Domain Objects within the profiles is needed to generate accurate recommendations. The growth of the profiles in form of an increasing number of Domain Objects occurs in two separated phases. The initial profile is handled as a static profile, where the user defines content he or she is interested in. To setup the initial profile, the user can choose TV content from the last days on a web interface. This step is optional for the user and just needed, if the user expects to get accurate recommendation from day one. Thus, the system is able to work without any data of the user's watching behavior. The evolution of the adaptive profile is a continuously process, where the system evaluates the user's watching behavior and the relative duration that the user has watched a certain TV content. It is necessary to assign the content in one of the two classes within the profile. If the determined value exceeds a certain threshold it is assumed that the user likes the viewed content. Under this assumption the user rates content by implicit feedback. There is also the option to give an explicite feedback. For this, a user rates or records a specific content. In case of the recommendations via social networks the user gets an explicite feedback by another user using direct messages. The functionality of the social network approach is described in section 4.2. In a long term usage of the system, the aim is to obtain a smooth transition between both of the profiles that the influence of the initial user profile decreases and the recommendation process rather basis on the data of the adaptive profile.

Figure 1. shows an architectural overview of the system, where the interaction and connection between the different components of the system is given. Besides the described interaction between user profiles and the recommendation engine, Figure 1. shows the connection to the PVR. The PVR works on a Set-Top-Box, which is connected to the recommendation engine. The PVR provides the data about the user's viewing behavior. Furthermore, the recommendation engine generates record lists with recommended content and sends it to the PVR system.

This architectural concept offers two possibilities for the implementation: the recommendation service can be implemented within a set-top-box (PVR) but it can be implemented as an external service too.

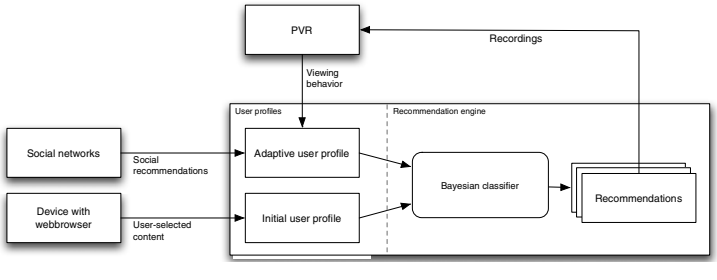


Fig. 1. Architectural overview

4 Recommendations Process

The previous section presented the system architecture. Within this section both of the recommendations components are described in detail. The Bayesian classifier is described in subsection 4.1 and the social network approach in subsection 4.2.

4.1 Machine Generated Recommendations

A Bayesian classifier is the core for our machine generated recommendations and has been proved as an efficient classification mechanism in several areas (cp. Section 2). In principal the classifier bases on a 2-class decision model, where an item – in our case a Domain Object – belongs to the class like or dislike [15]. As described in section 3 the initial and adaptive user profiles enriched with Domain Objects are the basis for mechanically generated recommendations. For the following, we define C+ to be the class like and let C- the class dislike. To determine to which class an item belongs, the Bayesian classifier calculates the conditional probability for an unrated item, where the attributes – also described as features – of an item will be used for calculation (attributes also have been defined in section 3). For this, the count of how many times each feature occurs in the set of the classes C+ and C- is relevant. So, the classifier is moreover a statistical learning technique, which learns from the underlying data. In our case, the underlying data basis consists of the Domain Objects from the initial and adaptive user profile. Let the features of a Domain Object be $f_1,...,f_n$ based on the number of defined attributes within a Domain Object.

We define $k(C+)$ to be the count of items in class C+ and $k(f_i \mid C+)$ the count of items with the features f_i in class C+. Then we have the conditional probability that a given feature occurs in the class C+ as:

$$P(f_i \mid C+) = k(f_i \mid C+) / k(C+) . \tag{1}$$

We can use the same notions for class C- and have the equivalent probability for the class C-:

$$P(f_i \mid C-) = k(f_i \mid C-) / k(C-) . \tag{2}$$

It is possible to estimate the probability for a single feature f_i. To estimate this for an item, the whole feature set needs to be respected. The features are summarized as a vector $\vec{x}$. The probability that an item with the feature set $\vec{x}$ fits into the class C+ can be estimated by using Bayes Theorem with

$$P(C+ \mid \vec{x}) = P(\vec{x} \mid C+) * P(C+) / P(\vec{x}) \tag{3}$$

where

$$\begin{aligned} P(\vec{x} \mid C+) &= P(f_1 \cap ... \cap f_n \mid C+) \\ &= P(f_1 \mid C+)...P(f_n \mid C+) \end{aligned} . \tag{4}$$

$P(\vec{x})$ is the a priori probability, which can be neglected in the calculation. It is assumed for the equation above that the features used for calculation don't depend on each other. The probability that an item belongs to the class C+ can be estimated with a count from the available data:

$$P(C+) = \frac{\mathrm{k(C+)}}{\mathrm{k(C+)} + \mathrm{k(C-)}} \tag{5}$$

$k(C+)$ stands again for the count of items in class C+. The divisor stands for the count of items in both classes. The calculations are analogous for the class C-. Finally, a total dependency ratio Q can be defined

$$Q = \frac{P(C+ \mid \vec{x})}{P(C- \mid \vec{x})}. \tag{6}$$

If the ratio exceeds a predefined threshold value, the item is assigned to class C+ and will be recommended to the user. Otherwise the item will be dropped. The threshold depends on the capacity of the PVR storage device. If there is less space for content, the threshold value increases and less content will be recommended. In the end a minimal threshold is needed to ensure a certain quality of recommendations. If the threshold value is set too low, too much content would be recommended and the user would get a limited personalized content. To evaluate a well-working threshold is a part of future work.

4.2 Recommendations via Social Networks

A fundamentally different approach of generating recommendations is to use social networks and in this context human-to-human interaction. We use the assumption that acquaintances and friends are able to suggest TV content in a more appropriate way than the Bayesian classifier does. The classifier uses meta- and statistical information to generate recommendations, whereas a human being interprets the substance of e.g. a movie to weight if it's the right content to recommend. The system provides interfaces to the social networks Facebook and Twitter. As shown in Figure 2. the PVR isn't connected to a network directly. Both networks applying different approaches for human-to-human interaction, why a data aggregation on a separate server is needed. Within the Facebook network every user provides a wall, where the user self or other users can publish messages. Facebook provides an Application Programming Interface (API) for Facebook application development including the usage of Facebook functionality. In the Facebook application developed for the NG-PVR system a user can choose TV content from a digital TV guide to suggest it to other users. The suggestion will be published to the respective user wall and will be sent to the recommendation server for data aggregation. The server keeps the recommendation ready for the user's PVR. A relation between a Facebook user and the matching PVR is made by the Facebook login information.

Twitter however is a microblogging service, where users can publish short status messages of 140 characters. The service is based on the principle of followers, that a user receives messages just from those users he or she is following. It is possible to integrate hashtags within the status messages. The Twitter network can be discovered

for hashtags to find recommendation messages for the NG-PVR system. If a message is marked with the hashtag #ng-pvr and has the needed TV content information, all followers using the system get a recommendation for the specific content. As an alternative, recommendations can be sent to an individual user only by including a user hashtag (e.g. @username). Unlike Facebook a crawler service is needed to forward the recommendations to the PVR. The service is running on the recommendation server and is looking for the described hashtags within the Twitter network. The relation between a Twitter user and the PVR is made by the Twitter login information. For a comfortable use of the Twitter service, a Smartphone application has been developed, where users can publish status messages from a mobile TV guide. The application integrates the needed TV content data and hashtags by itself.

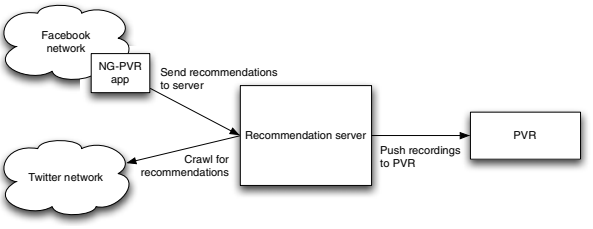


Fig. 2. Social network interaction

5 Conclusion and Future Work

This paper presents an approach for a user supporting PVR system by implementing two different recommendation mechanisms - machine generated recommendations using a Bayesian classifier and recommendations via social networks. By now both mechanisms are fully implemented on the recommendation server as described in section 3 and 4. The next step is to implement an automatic push service, which provides a connection for the exchange of data and record lists between the recommendation server and the PVR system. There are plans for a web service interface to realize the connection. To give a user the ability to manage the recommendations and automatic records, the possibility of implementing an HbbTV user interface within the PVR system will be prospected. Furthermore the recommendation process needs to be evaluated. Especially the quality of the machine-generated recommendations according to the user's interests will be focused.

References

1. SES Astra (October 6, 2010), http://www.astra.de/2117/de
2. Amazon.com (October 11, 2010), http://www.amazon.com
3. Facebook (November 16, 2010), http://www.facebook.com
4. Twitter.: (November 16, 2010) http://www.twitter.com
5. Davidson, J., Liebald, B., Liu, J., Nandy, P., Vleet, T.V.: The YouTube Video Recommendation System. In: ACM Conference On Recommender Systems 2010, Barcelona, Spain, pp. 293–296 (September 2010)

6. YouTube – Broadcasting yourself (February 2005) http://www.youtube.com (October 13, 2010)
7. De Pessemier, T., Deryckere, T., Martens, L.: Extending the Bayesian Classifier to a Context-Aware Recommender System for Mobile Devices. In: Fifth International Conference on Internet an Web Applications and Services 2010, Barcelona, Spain (2010)
8. Xu, J., Lu, H., Li, Y.: The Development and Prospect of Personalized TV Program – Recommendation Systems. In: IEEE Fourth International Symposium on Multimedia Software Engineering 2002, Newport Beach, California, USA (2002)
9. Walter, E.F., Battiston, S., Schweitzer, F.: A Model of a Trust-based Recommendation System on a Social Network. In: Autonomous Agents and Multi-Agent Systems, October 18, vol. 16. Springer Science and Business LLC (2007)
10. Goldbeck, J., Hendler, J.: FilmTrust: Movie Recommendations using Trust in Web-based Social Networks. In: IEEE Consumer Communication and Network Conference 2006, Las Vegas, USA (2006)
11. FilmTrust.: http://trust.mindswap.org/FilmTrust/ (November 25, 2010)
12. Fernandez, Y., Pazos-Arias, J.J., Gil-Solla, A., Ramos-Cabrer, M., Lopez-Nores, M., Barragans-Martinez, B.: AVATAR: A Multi-Agent TV Recommender System using MHP Applications
13. DVB Project Office.: DVB Fact Sheet – Multimedia Home Platform (April 2009) http://www.mhp.org/docs/dvb-mhp-fact-sheet.0409.pdf (October 29, 2010)
14. Pazzani, M., Billsus, D.: Learning and revising user profiles: The identification of interesting web sites. Machine Learning 27, 313–331 (1997)
15. Kurapati, K., Gutta, S., Martino, J., Zimmerman, J.: A Multi-Agent TV Recommender. In: Proceedings of the UM 2001 workshop Personalization in Future TV, Sonthofen, Germany, pp. 13–14 (2001)
16. Miller, B.N., Albert, I., Lam, S.K., Konstan, J.A., Riedl, J.: MovieLens Unplugged: Experiences with an Occasionally Connected Recommender System. In: IUI 2003, Miami, Florida, USA, January 12-15 (2003)
17. MovieLens.: http://movielens.umn.edu (November 25, 2010)
18. Ali, K., van Stam, W.: TiVo: Maiking Show Recommendations Using a Distributed Collaborative Filtering Architecture. In: KDD 2004, Seattle Washington, USA (August 2004)

Local Community Finding Using Synthetic Coordinates

Harris Papadakis[1,*], Costas Panagiotakis[2,*], and Paraskevi Fragopoulou[1,**]

[1] Department of Applied Informatics and Multimedia,
Technological Educational Institute of Crete,
71500 Heraklion, Crete, Greece
[2] Department of Marketing, Technological Educational Institute of Crete,
72200 Ierapetra, Crete, Greece
adanar@epp.teicrete.gr, cpanag@csd.uoc.gr,
fragopou@ics.forth.gr

Abstract. A fundamental problem in social networking and computing is the community finding problem that can be used in a lot of social networks' applications. In this paper, we propose an algorithm that finds the entire community structure of a network, based on interactions between neighboring nodes (distributed method) and on an unsupervised centralized clustering algorithm. Experimental results and comparisons with another method found in the literature are presented for a variety of benchmark graphs with known community structure, derived by varying a number of graph parameters. The experimental results demonstrate the high performance of the proposed algorithm to detect communities.

Keywords: Community finding, network coordinates, local community algorithm.

1 Introduction

During the last years, social networking has been driving a dramatic evolution due to increasing use of Web 2.0 elements such as blogs, twitter, facebook, LinkedIn, and wikis, etc. One of the fundamental problems in social networking with a lot of applications is to detect the communities that are created by the users' interaction. Several attempts have been made to provide a formal definition for this generally described "community finding" concept in networks. A *strong community* was defined as a group of nodes for which each node of the community has more edges to other nodes of the same community than to nodes outside the community [2]. This is a relatively strict definition, in the sense that it does not allow for overlapping communities and creates a hierarchical community structure since the entire graph can be a community itself. A *weak community* was later defined as a subgraph in which the sum of all node degrees within the community is larger than the sum of all node degrees toward the rest of the graph [4].

---

[*] Harris Papadakis and Costas Panagiotakis are also with the Department of Computer Science, University of Crete, 71409 Heraklion, Crete, Greece.

[**] Paraskevi Fragopoulou is also with the Foundation for Research and Technology-Hellas, Institute of Computer Science, 70013 Heraklion, Crete, Greece.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 9–15, 2011.

Variations also appear in the method used to identify communities: Most of the algorithms that appear in the literature follow an iterative approach starting by characterizing either the entire network, or each individual node as community, and splitting [3], [9] or merging [4], respectively. These methods produce a hierarchy of partitions. There is an entire hierarchy of communities, because communities are nested: small communities compose larger ones, which in turn are put together to form even larger ones. By merging or splitting communities one can build a hierarchical tree of community partitions called *dendrogram*. The modularity criterion defined in [3] is a measure of the quality of a partition, and can be used to identify a single optimal partition, i.e. the one corresponding to the largest modularity value.

Apart from the variations in definitions, there are many variations in the final desired solution. As an example, several researcher aim to find the entire hierarchical community dendrogram [3], [9] while others wish to identify only the optimal community partition [2]. More recently used approaches aim to identify the community surrounding one or more seed nodes [10]. Some researchers aim at discovering distinct (non-overlapping) communities, while others allow for overlaps [6]. The variations to the problem are non-exhaustive.

Algorithms that follow a global approach, not only require the entire network to work but they often require the entire network to be manipulated at each iteration [3], [9]. However, real world network, like for example Web graphs, Social Networks, Communication and autonomous system graphs, Peer-to-peer systems, Blogs, collaborative networks, citation graphs, Database relations, etc [7], count millions of nodes whose manipulation using global algorithms is prohibitive by any means. Recently, the emerging and demanding need to analyze this type of networks led the research towards the development of local community detection algorithms [10], [5]. These algorithms are base exclusively on local interaction between pairs of neighboring nodes for the identification of the community structure, thus global knowledge of the network is not required.

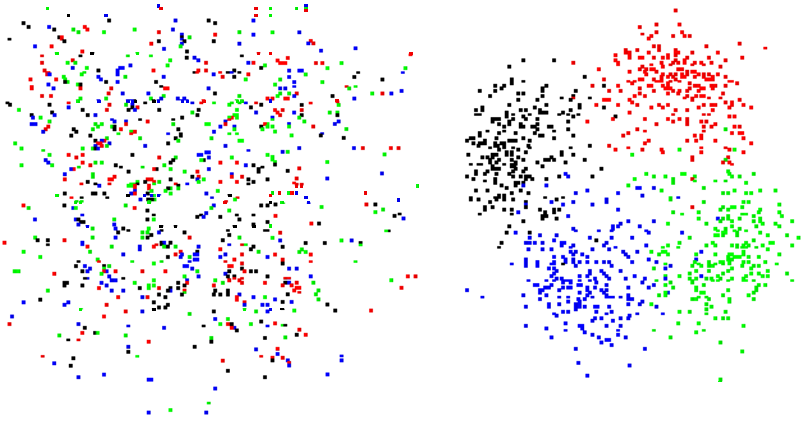


Fig. 1. Snapshots of the execution of the first phase of our algorithm for a graph with known community structure. (a) initialization, (b) after 400 iterations**.**

In this paper, we propose an algorithm to identify the entire community structure of a network based on interactions between neighboring nodes, which is inspired the Vivaldi synthetic network coordinate algorithm [1]. Our algorithm is based on the idea that by providing our algorithm with our own, appropriate, definition of distance between nodes, we can use Vivaldi to position the nodes in a manner as to reflect community membership, i.e. nodes in the same community will be placed closer in space than nodes of different communities. In the second phase of the algorithm, *K-means* is used to identify the natural communities formed in space. Experimental on several benchmark graphs with known community structure indicate that our algorithm is highly accurate in identifying community membership of nodes.

The remaining of the paper is organized as follows: The proposed community detection algorithm is presented and analyzed in Section 2. Section 3 describes the experimental framework and comparison results with another known algorithm on a number of benchmark graphs. Finally, we conclude in Section 4.

2 Local Community Finding

2.1 The Position Estimation Algorithm

The position estimation algorithm is inspired by the Vivaldi algorithm [1]. Initially, each node is placed at a random position in the n – dimensional Euclidian space. Each node calculates a "local" set containing all its neighbors, and a "foreign" set containing an equal number of randomly selected nodes. The size of local set as well as the size of foreign set of a node corresponds to the degree of node. As the algorithm evolves, and the springs connecting local and foreign nodes are tightened and relaxed respectively (according to a distance function we describe below), nodes of the same community pull each other close together, while nodes of different communities push each other further away. The success of the algorithm is based on the fact that, even though not all nodes in the "local" set belong to the same community, the majority of them do. Thus, the cumulative force of the neighbors in the same community is the larger part of the sum of all forces and so, will be able to pull themselves closer to one another. This means that the nodes of the same community will be placed closer in space than nodes of different communities.

Let C(x), C(y) denote the communities' sets of two nodes x ∈ V, y ∈ V, respectively, of a given graph. Since two nodes either belong to the same community (C(x) = C(y)) or not, we define the initial node distance between two nodes x and y as d(x, y):

$$d(x,y) = \begin{cases} 1, & C(x) = C(y) \\ 100, & C(x) \neq C(y) \end{cases} \quad (1)$$

By Equation 1, we can execute the Vivaldi algorithm to position the nodes appropriately in the n – dimensional Euclidian space. Vivaldi will position nodes in the same community close-by in space, while place nodes of different communities away from each other.

Let L(x) denote the "neighbor set" of a node x ∈ V, the neighbors of node x. It holds that the distance from node x to nodes in L(x) is set to 1 according to equation (1). Let F(x) denote the "foreign set" of a node x ∈ V, random nodes that are not connected with x. It holds that the distance from node x to the nodes in F(x) is set to 100 according to equation (1).

Iteratively, each node x ∈ V performs the following algorithm that is described hereafter. Initially each node is placed at a random position in the n−dimensional Euclidian space. Having defined the two sets L(x) and F(x) for each node x (the neighbor set and the foreign set) and randomly selecting, at each iteration, a node y either from set L(x) or from F(x), we use the Vivaldi algorithm with parameters the current position of node x (p(x)), the current position of node y (p(y)) and their distance d(x, y) to acquire a new position estimate for node x. The iterations stop when all the nodes' positions have converged.

Fig. 1 shows a small time-line of the execution of our algorithm on a graph with 1024 nodes, degree 20 and known community structure comprising four communities. We have used different colors for the nodes of each different community. Initially, nodes were randomly placed in the 2-dimensional Euclidian space. Finally, we see that nodes of the same color, belonging to the same community, gravitate to the same part, forming distinct clusters in space.

2.2 Community Detection Using K-means Clustering

Having estimated the coordinates (position) for each node of the graph, we use the K-means clustering algorithm [8] in order to identify the communities. K-means is a centralized clustering algorithm [8] and it is one of the simplest unsupervised learning algorithms that solves the well known clustering problem under the assumption that the number of clusters (number of communities), Kis provided.

K-means has linear computation cost O(K · N · n), where N denotes the number of graph nodes. This makes K-means an efficient choice even for large graphs. The main idea of K-means is to define Kcentroids, one for each cluster. These centroids will eventually be placed at the centers of the natural clusters. The goal of K-means is to minimize the sum, over all clusters, of the within-cluster sums of node-to-cluster-centroid distances.

It makes sense to apply K-means clustering to our data (the positions of nodes), since according to Vivaldi, the nodes of the same cluster (community) are placed close to each other in the Euclidean space and nodes of different communities place themselves away from each other.

Table 1. The different values for the used parameters

The number of nodes *N* of the graph	1024, 4096
The number of communities (Comm) of the graph	2, 4, 8, 16, 32, 64
The ratio of local links to degree (local/degree)	0.55, 0.65, 0.75, 0.85
The degree per node (degree)	5, 10, 20, 30, 40

3 Experimental Results

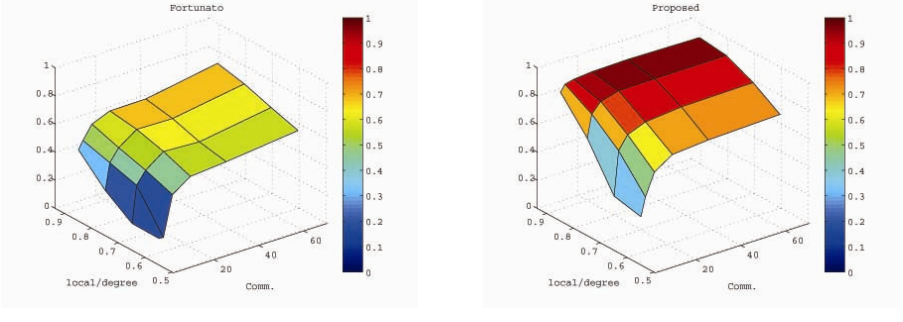
3.1 Benchmark Graphs

We have created our own benchmark graphs with known community structure to test the accuracy of our algorithm. Our benchmark graphs were generated randomly given the following set of parameters (see Table 1). The parameters used by the algorithm and their corresponding values are shown in Table 1.

The total number of nodes is divided among the various communities and then the wiring of the links is performed. The local links are selected randomly among the nodes of the same community (excluding, of course, nodes which are already neighbors, nodes with complete neighbor set and the connecting node itself). The foreign links are selected randomly among all nodes in the graph, excluding the nodes in the same community as the connecting node. Concerning, the dimensions (n) of Euclidian space that is used by Vivaldi algorithm, we have used n = 20 dimensional Euclidian space getting high performance results.

3.2 Community Detection Results

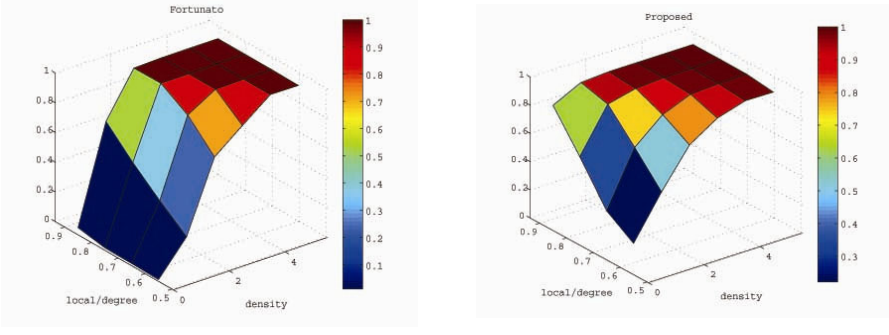
In this subsection, the results of the proposed algorithm are presented on the benchmark graphs. In addition, in order to evaluate the accuracy of the proposed scheme, we compare it to an efficient distributed community detection algorithm found in the literature [6] (Fortunato). In our experiments, the parameter α of Fortunato algorithm has been set to one, which is the default value of α[6]. In order to measure the performance of the community detection algorithms, the measure of accuracy has been used. Let S_i, $i \in \{1, ...,K\}$ be the estimated and $\dot{S}_i$, $i \in \{1, ...,K\}$ the corresponding natural communities. The accuracy is given by the average of number of nodes that belong to the intersection of $S_i \cap \dot{S}_i$ divided by the number of nodes that belongs to the union $S_i \cup \dot{S}_i$. It holds that accuracy $\in$ [0, 1], the higher the accuracy the better the results. Let density be the product of degree^2 and Comm, that measures how dense the graph is.



(a) (b)

Fig. 2. The mean value of accuracy under different ratios of total degree to local links (*local/degree*) and number of communities (*Comm*) for (a) the Fortunato and (b) proposed algorithms

As we mentioned before, the benchmark graphs used in the experiments were constructed based on four parameters of Table 1. In the presentation of the experimental results below, the accuracy is plotted against one or two of those parameters. This means that each value of the bound variable is an average of more than one experimental result. Fig. 2 illustrates the mean value of accuracy under different ratios of local to degree (local/degree) and number of communities (Comm) for the proposed and Fortunato algorithms. Fig. 3 illustrates the mean value of accuracy under different ratios of local to degree (local/degree) and densities for the proposed and Fortunato algorithms. According to these figures, it holds that the proposed method outperforms the Fortunato, since the accuracy of the proposed method is higher than the corresponding accuracy of Fortunato under any case. The most important reason for this lies in the fact that the Fortunato algorithm relies on the existence of triangles in the graph (i.e.: two nodes with common neighbor). The possibility of existence of those triangles relies on the density of the graph (i.e: the average degree of the nodes in conjunction with the total number of nodes). Fortunato fails on graphs with low density, while our proposal is un-affected by the degree (i.e. amount) of density. Our algorithm is mostly affected by the ratio of local links to all links of a node, a metric which defines the clarity of the community structure in the graph and thus also affects Fortunato.



(a) (b)

Fig. 3. The mean value of accuracy under different ratios of total degree to local links (*local/degree*) and densities for (a) the Fortunato and (b) proposed algorithms

4 Conclusions

We presented a community finding algorithm that solves a fundamental problem in social networking and computing and it can be used in a lot of social networks' applications. The proposed algorithm is based on the Vivaldi network coordinate system (distributed method) and on an unsupervised centralized clustering algorithm. The proposed method has been tested on a large number of benchmark graphs with known community structure. Furthermore, it has been compared to the local community detection algorithm found in [6]. Its effectiveness in terms of community

detection has been demonstrated with a large set of experiments. In addition the proposed algorithm demonstrates that network coordinates could be used to derive efficient solutions to a variety of problems.

References

1. Dabek, F., Cox, R., Kaashoek, F., Morris, R.: Vivaldi: A decentralized network coordinate system. In: Proceedings of the ACM SIGCOMM 2004 Conference (August 2004)
2. Flake, G.W., Lawrence, S., Giles, C.L., Coetzee, F.M.: Self organization and identification of web communities. IEEE Computer 35, 66–71 (2002)
3. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America 99(12), 7821–7826 (2002)
4. Katsaros, D., Pallis, G., Stamos, K., Vakali, A., Sidiropoulos, A., Manolopoulos, Y.: CDNs content outsourcing via generalized communities. IEEE Transactions on Knowledge and Data Engineering 21, 137–151 (2009)
5. Lancichinetti, A., Fortunato, S.: Community detection algorithms: a comparative analysis. Physical Review E 80 (5 Pt 2), 56117 (2009)
6. Lancichinetti, A., Fortunato, S., Kertész, J.: Detecting the overlapping and hierarchical community structure in complex networks. New Journal of Physics 11(3), 33015 (2009)
7. Leskovec, J.: Snap: Stanford network analysis package
8. MacQueen, J.B.: Some methods for classification and analysis of multivariate observations. In: Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297 (1967)
9. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Physical Review E 69(2), 26113 (2004)
10. Papadopoulos, S., Skusa, A., Vakali, A., Kompatsiaris, Y., Wagner, N.: Bridge bounding: A local approach for efficient community discovery in complex networks. Technical Report arXiv:0902.0871 (February 2009)

Virtual B-Tree Topology: A Self-Routing Topology for a Wireless Sensor Network

Shan-Yang Wang[1,2], Tzu-Chiang Chiang[3], and Chu-Sing Yang[4]

[1] Department of Computer Science and Information Engineering, Far East University
Tainan County, Taiwan, R.O.C.
[2] Department of Computer Science and Engineering, National Sun Yat-Sen University
Kaohsiung, Taiwan, R.O.C.
[3] Department of Information Management, Tunghai University
Taichung, Taiwan, R.O.C.
[4] Department of Electrical Engineering, National Cheng Kung University
Tainan, Taiwan, R.O.C.
{Shan-Yang Wang,Tzu-Chiang Chiang,
Chu-Sing Yang}mango@ms1.hinet.net

Abstract. A wireless sensor network consists of many wireless sensor nodes that work together. Since communication between nodes depends on routing protocol, a portion of the required energy is consumed by periodic routing protocol or flooding. This research shows that the routing overhead can be reduced by making a virtual B-Tree topology out of these nodes through utilizing the characteristics of both a wireless network and a B-Tree, and thus the network can perform a self-routing characteristic. As soon as the topology is organized, paths between peers come out, and there is no need to wait to find path. During forwarding packets, the nodes can transmit packets through a better path. The results indicate that 35 unit times are required to organize 100 randomly deployed nodes and do not need periodic routing messages.

Keywords: Self-Routing, Virtual B-Tree, self-routing topology, ad-hoc networks, b-tree embedding.

1 Introduction

The applications of WSNs have gradually grown more diverse. They include military affairs, environmental monitors, ecology maintenance, and so on. [1] [2] Most WSNs build randomly, and lack a central infrastructure. Thus, it is difficult to enable nodes to communicate with each other. Some research has emphasized routing protocol to route packets between nodes. Most routing protocols need extra messages to exchange the routing table. This research adopts another method through utilizing the characteristic wireless network and a B-Tree to propose a topology named *Virtual B-Tree*. By grouping the WSN to a Virtual B-Tree topology, the network can be self-routing. After it finishes construction in a short time, all the nodes can forward packets to each other. In this paper, we proposed a Virtual B-Tree topology and apply it to a randomly deployed WSN.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 16–23, 2011.

2 Related Works

Since WSNs can be considered to be a special type of ad hoc networks, all general ad hoc network routing protocols, including DSDV and AODV (Ad Hoc On-Demand Distance Vector) [3-5], can be adopted for WSN.[3, 4] Some use Content-Based Routing (CBR), which determinates how to route packets according to the senders' interests.[6] A self-routing topology for Bluetooth scatternets were introduced by Min-Te, Sun, et al. [7] They proposed the operations of the Blue-Tree to organize the scatternets. Wen-Zhan Song and Xiang-Yang Li use hierarchical structures and clusters to group the scatternets. [8] They use the connected dominating set (CDS) as the backbone of scatternets to perform self-routing.

The Ad Hoc routing protocol can be divided into two categories. One is table-driven, and the other is an on-demand routing protocol. More than one routing table might be stored in the nodes. The first ones need to constantly exchange their routing information and to relearn the routing entries, and DSDV[3] is one of them. Frequent system-wide broadcasts limit the size of ad-hoc networks that can effectively use DSDV because the control message over-head grows as $O(N^2)$ [4]. On-demanding routing protocol is the other one[4]. Before it begins to send a packet, it must flood requests to the entire network in order to get knowledge of the path. The routing protocol has the disadvantages of delaying sending and flooding packets.

The DSDV and AODV listed above can be modified to WSNs, but they all need a period of time, proportional to N^2, to learn routes. They originate from a wire network, and are modified to a wireless network. WSN characteristics include multiple nodes and less energy. Besides these characteristics, broadcasting to neighbor nodes is an important feature. Neighbor nodes can receive packets regardless of whether or not they are the destination. This feature might be utilized in network communication. The alternative characteristics listed below are challenges in WNS communication. Synchronizing all the nodes is difficult. Routing protocol ultimately results in problems of scalability. Unneeded broadcasts exhaust the energy.

Thus, in this paper, we propose a Virtual B-tree topology to alleviate problems caused by scalability and too much of routing traffic. WSNs organized into a Virtual B-tree can perform self-routing, and can organize them in a distributed and parallel way without advance synchronizing.

3 Virtual B-Tree Topology

This section is divided into two parts to induce a virtual B-Tree Topology. The first part explains how the topology was constructed, and the second part explains how a Virtual B-Tree Topology forwards routed packets.

Virtual B-Tree topology is a logical topology imposed on any peer to peer networks. In this research, we impose it on WSNs. Because the topology of randomly deployed nodes is a disorderly connection, the study imposes a virtual B-Tree topology on it to use the characteristics of a B-Tree. The scenery is a wireless sensor network randomly placed and based on connected graph. Each node of the WSN is location-aware, and it can communicate with other nodes within its transmission range. All the links are symmetrical. The codeword is a string of binary characters,

assigned to each node to imply its position in the topology, and we assume that each node can identify and receive the packets with its node number or codeword in the virtual B-Tree topology. The purpose of constructing a virtual B-Tree topology is to assign each node a codeword, by which to help routing.

3.1 Codewords for the Virtual B-Tree

The codeword is 0{0,1}*. Such as {0,00,01,001,…} can be treated as the codeword of the node, including the inner node or leaf node. The codeword, or root node, is "0", the path to the left is "0", and the path to right is "1". The codeword for each node is "0" followed by the path coding to the root. Figure 1 is a codeword example for a virtual B-tree. A vrtual B-tree is designed to make a logic B-Tree out of a random point. Based on the logical B-Tree, it outputs the codeword for each node.

Node	Label
A	0
B	00
C	01
D	000
E	001
F	010
G	011

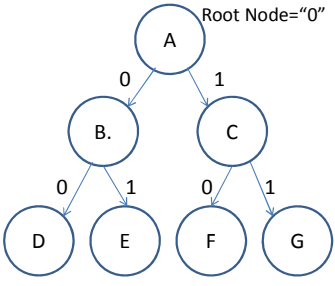


Fig. 1. The Codeword of the Virtual B-Tree

3.2 Algorithm to Construct a Virtual B-Tree Topology

The algorithm to build a virtual B-Tree is distributed and parallel. Suppose the topology is one kind of undirected graphics. It randomly chooses a node as the root node and also the first parent node. The sink node is the fittest node to be the root node. In Figure 2, root node first send requests to all its neighbors to start constructing the virtual B-Tree. All the procedures are described below. Assuming two nodes, A and B, A is the parent node and B is the child node. Node A wants node B to be one of its descendants.

Node A sends a request message to its neighbors to start constructing the virtual B-tree. After B receives the packet, B replies to A to be the descendant of node A. As A receives B's message, it designates B as a child node, calculates B's codeword and sends it to B. Node B latches A as its parent node and sends its node number and codeword to the root node. As the packet is forwarding B to the root node, Node B can also send the packet to its neighbor as the first step, concurrently. Now, node B takes the role of a parent node. The 4th step is a key step. It sends the node number and codeword to the root so that all the nodes forwarding their packets can learn the map of B's node number and codeword. This utilizes a characteristic of the tree that helps routing.

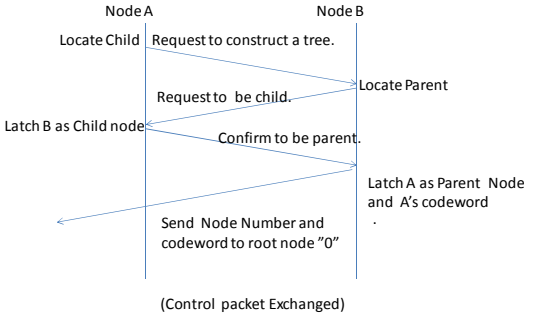


Fig. 2. Protocol to construct a Virtual B-tree

Packets to the root node must pass its ancestral node, than all the ancestral nodes can learn the map of the node number and codeword from the packet and store them to the Label Table a map of the node number and codeword.

3.3 Operations of Parent Node in Constructing a Virtual B-Tree

Apparently, parent nodes play an important role in constructing a virtual B-Tree. After a parent node receives child node's message that it is the child, it must arrange all the child nodes to be a virtual tree, so as to generate the codeword for each node. The parent node arranges the child nodes to be the left child or the right tree according to the child node's Y-position. By and large, cases have more than two nodes to joining the same side of the child tree. Parent nodes must have knowledge to control the situation by virtually adding on a loopback node. A loopback node is a node structure with only with three pointers. One pointer is assigned to the parent node, and the others are pointed to the child nodes. The pointers are the same as physical nodes pointers. The loopback nodes functionally maintain a virtual B-Tree. A virtual B-Tree is built mainly by adopting loopback nodes when the branch degree is more than 2.

Node	Codeword
A	0
B	000
C	0010
D	011
E	0101

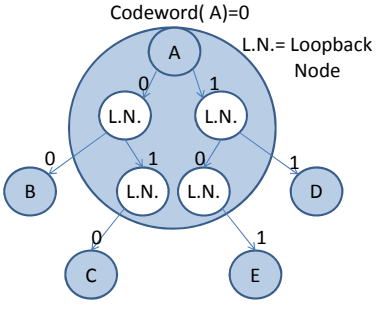


Fig. 3. Multiple nodes are added to the same parent node

Figure 3 shows how the parent node adds multiple children nodes, and gives them the codeword to present a virtual B-Tree. Although A has degrees greater than 2, it still can maintain the child nodes in a B-Tree codeword.

3.4 Routing

There are three tables in each physical node, including the Label table, the Link table, and the Pointer table. These three tables are all the maps of the codewords and node numbers from the different methods. The Label table is from the 4th step in the construction of a virtual B-tree where the child node sends its codeword and node number to the root node. The Label table is also from the packets sent by neighbor nodes.[9] The Link table is node's neighbors those can direct send packets to the node. The Pointer table has three tuples storing the node's physical upstream and two children node numbers. The Pointer table mainly constructs the virtual B-Tree topology.

Table 1. Algorithm for routing

```
On_receve_data_packet(){
if(destination's codeword is unknown)
 {
   Search destination's codeword in the Label table;
   if(Destination's codeword is found )
    Routing the packet to the neighbor by codeword;
   else
    Routing the packet to parent node;
 }
 else
    Routing the packet to the neighbor by codeword;
```

The essential component of virtual B-Tree routing is that it can route dual node numbers and codewords. If a destination's codeword is known, packets can be routed to neighbors according to the Link table or the Point table. If the codeword is unknown, packets can be routed to the parent of the Point table by default. As these packets are routed in the nodes, the nodes can query the Label table to locate the packets' codewords. The nodes can route the packets to short paths by comparing the prefix of the codeword of the packet destinations with the prefix of the codeword of the nodes in the Link table. In extreme cases, a destination's codeword can be found at root node, and be routed by the codeword. Table 1 shows the algorithm for routing.

4 Example of the Virtual B-Tree

Figure 4 illustrates a randomly deployed WSN and its virtual B-Tree topology is Figure 5. Figure 5 shows that node 10 is a root node, the dark nodes indicate physical node, and empty nodes mean loopback nodes. Each node is assigned a codeword as soon as the virtual B-Tree is constructed. Their codewords are listed in Table 2. Each node receives the maps of node number and codeword for all of its downstream nodes. The node can then begin to route packets by node number or codeword

immediately. Node 2 can route packets to Node 0 through Node 10 by default without Node 10's codeword. Node 2 can route to Node 0 through Node 11 in a shorter path by codeword, also.

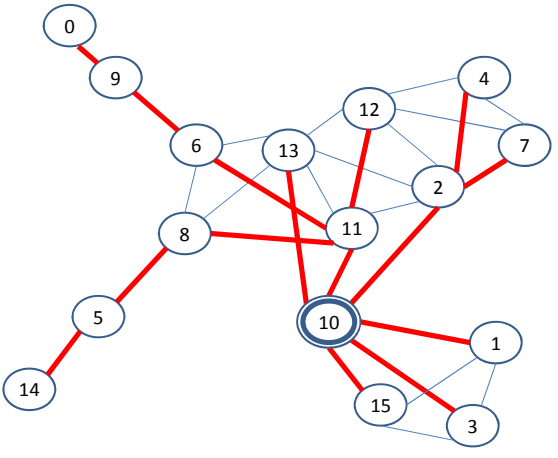


Fig. 4. A randomly deployed WSN network

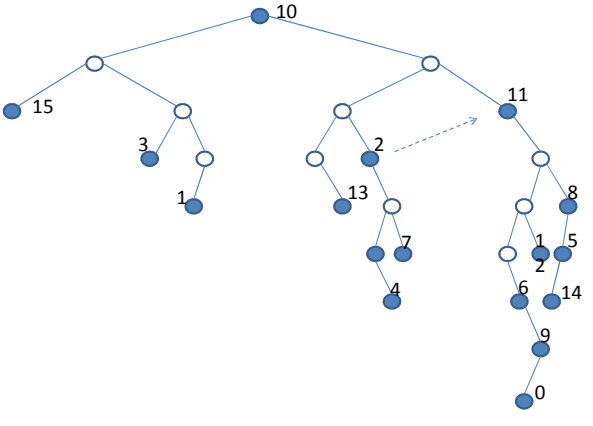


Fig. 5. A Virtual B-Tree Topology

Table 2. Codewords

Node	Codeword	Node	Codeword	Node	Codeword
0	011100110	6	0111001	12	011101
1	00110	7	010111	13	01001
2	0101	8	01111	14	0111100
3	0010	9	01110011	15	000
4	0101101	10	0		
5	011110	11	011		

5 Experiment

In order to experiment the wide practical use of a virtual B-Tree, we applied the Virtual B-tree on both sensor networks, in a grid deployed. The thin lines stand for nodes' links and the thick lines stand for the virtual B-Tree topology.

Figure 6 shows a Virtual B-Tree topology in a 10x10 grid network. In this experiment, node 10 is randomly chosen as a root node, and its codeword is "0", node 99 being the farthest node from node 10. From node 10 to 99 needs just 10 hop counts. The path between nodes 92 to 95 is obtained according to the codeword as soon as the virtual B-tree is created. The first packet from node 95 to 92 is routed by the virtual B-Tree by default and the second packet from node 92 to 95 is routed by the codeword in a shorter path.

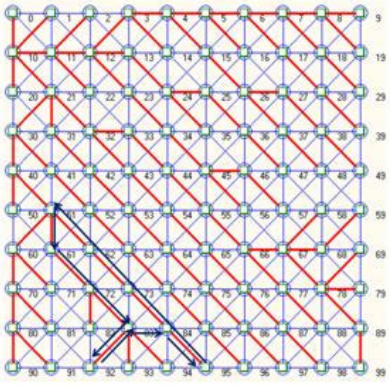


Fig. 6. A virtual B-Tree in a grid network and the packets routed between node 95 and 92

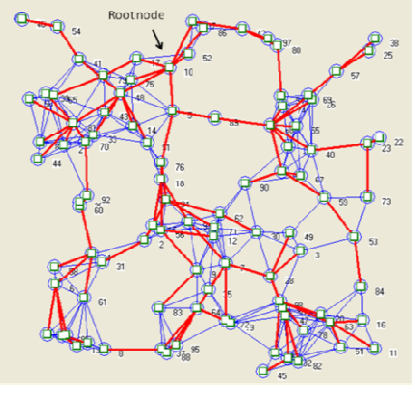


Fig. 7. A virtual B-Tree in a Randomly Deployed WSN

In order to observe its scalability, we do several experiments on the relation between consumption time and node counts. In this experiment, we assume there is no collision and that the bandwidth is enough, then started to measure the construction

time under different node counts and transmit distances. Figure 7 is a randomly deployed WSN. Figure 8and Figure 9 illustrate that transmit distance has more effect on construction time than nodes counts have.

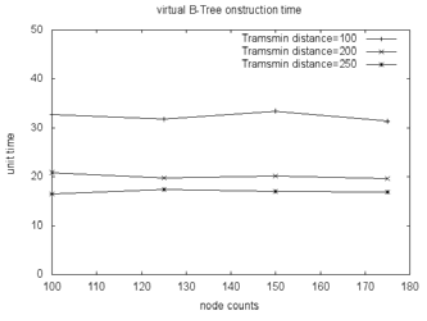


Fig. 8. Construction time under different node counts

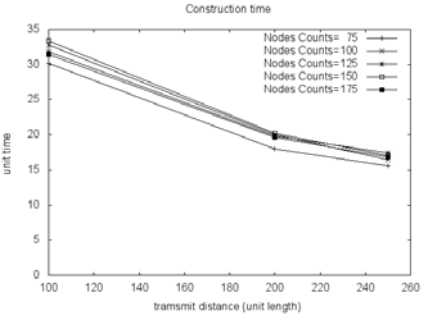


Fig. 9. Construction time under different transmit distance

6 Conclusion

In this research, we present a virtual B-Tree topology and how the virtual B-Tree is used in WSNs. The experiment show the virtual B-Tree topology don't need extra periodic routing messages to maintain its routing table, and don't query the paths before sending packets. As soon as the topology is build, paths from peer to peer are obtained, and the construction time is dependent on nodes' transmits distance.

References

1. Akyildiz, I.F., Weilian, S., Sankarasubramaniam, Y., Cayirci, E.: A survey on sensor networks. IEEE Communications Magazine 40(8), 102–114 (2002)
2. Chien-Chung, S., Srisathapornphat, C., Jaikaeo, C.: Sensor information networking architecture and applications. City (2001)
3. Perkins, C., Bhagwat, P.: Highly-dynamic destination-sequenced distance-vector routing for mobile computers. In: Proc. ACM SIGCOMM, pp. 234–244 (September 1994)
4. Perkins, C.E., Royer, E.M.: Ad-hoc on-demand distance vector routing. City (1999)
5. Zhang, L., Sun, H., Sun, Q., Hu, X.Y., Guo, C.J.: Improvement of Routing Protocols AODV in Wireless Ad Hoc Network. City (2009)
6. Mottola, L., Cugola, G., Picco, G.P.: A Self-Repairing Tree Topology Enabling Content-Based Routing in Mobile Ad Hoc Networks. IEEE Transactions on Mobile Computing 7(8), 946–960 (2008)
7. Min-Te, S., Chung-Kuo, C., Ten-Hwang, L.: A self-routing topology for Bluetooth scatternets. City (2002)
8. Wen-Zhan, S., Xiang-Yang, L.: Hierarchical self-routing scatternet for multihop bluetooth networks. City (2006)
9. Chong, L., Wei, P., Chang Wen, C., Jun, S., Feng, W.: Forepressure Transmission Control for Wireless Video Sensor Networks. City (2009)

A Peer-to-Peer Social Network Overlay for Efficient Information Retrieval and Diffusion

Salma Ktari and Artur Hecker

TELECOM ParisTech, Paris France
{ktari,hecker}@enst.fr

Abstract. In this paper we explore new ways in P2P based Online Social Network (OSN) design and present a novel methodology for efficient information retrieval and diffusion in OSN. We propose a new P2P topology model that matches real OSN graph. The new topology decreases the graph diameter and thus the route length by rendering the graph power law. Our proposal can be adapted to existing P2P based OSN systems and thus can fit diverse OSN application requirements. We present the architecture of our adaptable overlay network model as well as the proposed search techniques employed to provide efficient data search and information diffusion. Through simulations, we show the effectiveness of our approach in term of search route length and signaling cost.

Keywords: complex search, KBR, flooding, scale free graph, DHT.

1 Introduction

In both professional and personal life, humans naturally form groups based on affinities and expertise. We gravitate to those with whom we share interests to meet other people and to gather and share first hand information and experiences, like developing friendships and professional alliances. With the ongoing technology development, these networks rapidly migrated to the online world and formed Online Social Networks (OSNs). OSNs form online communities among people with common interests, activities, backgrounds, and allow users to upload profiles and documents (text, images, and video) and interact with others in numerous ways.

The nature of today's OSN services raises major concerns about user privacy and autonomy. In particular, the predominant services concentrate personal data for millions of people under a single centralized administrative domain giving the respective service providers too much power regarding the possible analysis of the social links and data of the participants. A decentralized approach to online social networking could avoid the above problems, and give the users fine control over what personal data they share and with whom, in a distributed environment.

For this purpose, several researchers propose new OSN system based on a P2P distributed architecture like Vis à Vis [1], PeerSoN [2] and Safebook [3]. However, in these P2P based OSN systems, some common OSN operations are expensive and infeasible to implement in the new distributed environment. One such feature is an "Advanced Search" or "Complex Search" interface, which allows users to filter users

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 24–33, 2011.

across a wide range of attributes or compute the intersection of multiple typed groups. In fact, unlike the Web, no explicit links exist between the content stored in social networks (typically photos, videos, and blog postings). Instead, explicit links between users, who generate or publish the content, serve as the primary structuring tool.

The predominant method of finding information in online social networks is to navigate through the social network, browsing content introduced or recommended by other users. Some centralized OSN sites provide keyword-based search for textual or tagged content. Additionally, other sites have 'top 10' lists showing the most popular content, where the popularity is determined according to how often users have accessed the content or based on explicit recommendations provided by users. Since the content can be non-textual, obscure, or short-lived, it may be hard to find through classical Web search engines.

In our paper, we explore new ways in P2P OSN design using social network graph studies, and present a novel methodology for efficient information retrieval and diffusion in OSN. Our proposal can be adapted to existing P2P based OSN systems (Vis à Vis, PeerSoN, Safebook…) and thus can fit diverse OSN services requirements.

The rest of this paper is organized as follows. Related work is detailed in section 2. We describe our P2P based OSN model in section 3. Section 4 details our proposed search techniques for information retrieval and diffusion. In section 5, we detail the simulation environment and provide performance results. Finally, we conclude and discuss possible future work in section 6.

2 Related Work

To address the privacy concerns in the centralized online social networking platforms, alternative P2P-based distributed approaches were proposed in OSN. Authors in [1] propose the Vis à Vis approach, a privacy preserving framework for online social networking based on the novel concept of Virtual Individual Servers. VISs are personal virtual machines running in a cloud computing facility. In Vis à Vis, each person manages personal data such as friend lists, photographs, and messages through her own VIS. Each user owns a VIS and maintains control over the data, software, and access-control policies on her VIS. VISs form overlay networks, one overlay for every OSN group that its owner joins. Vis à Vis uses Pastry distributed hash tables (DHTs) to achieve scalable and efficient OSN operations (joining, leaving, and searching) on this large federation of machines. Another approach proposed in [2] is PeerSoN. To address the privacy problem, PeerSoN uses encryption and access control coupled with a peer to peer approach to replace the centralized authority of classical OSNs. Decentralization makes it easier to integrate direct data exchange between users' devices into the system. Direct exchange allows users to use the system without constant Internet connectivity, leveraging real life social networking and locality. The Safebook approach [3] adopts a decentralized architecture relying on the cooperation among users of the online social network. Safebook is based on trust relationships that are part of social networks in real life in order to cope with the problem of building trusted and privacy-preserving mechanisms as part of the OSNs. The addressing and routing, both for P2P lookup and for data retrieval are based on the pseudonyms of nodes. The mapping between the pseudonym of a node and its

identifier is only known to the TIS (Trusted Identification Service [3]) and direct connections (friends), which are trusted by the node, no private information can be derived from it either.

In addition, commercial initiatives to develop decentralized infrastructures for OSNs are available. Tribler uses a P2P infrastructure for video on demand application among friends. Wuala [6] is a partially decentralized encrypted file system based on Chord DHT, allowing users to share files freely and securely over the Internet. Other P2P based OSN approaches were proposed like Social VPN [4] and NOYB [5].

3 Exploiting Social Network Information Graph for Topology Design

Unlike the Web, which is largely organized around content, online social networks are organized around users. Participating users join an OSN network, publish their profile and any content they like, and create links to other users typically reflecting their real life contacts, domains of interests, etc. Understanding the graph structure of online social networks is essential to evaluate the current systems, to design future online social network based systems. OSNs present a behavior consistent with a power-law network [9] [10], which is characterized by a strong local clustering and a small diameter ($d \sim log\ log\ N$ [7] [8]). The majority of nodes have a small degree, and few nodes have a significantly higher degree. This suggests that high-degree nodes in the core are critical for the connectivity and the flow of information in these networks. Moreover, studies in [11] observed a high degree of reciprocity in user contacts, leading to a strong correlation between user indegree and outdegree in the social graph. In fact, most contacts are reciprocal (i.e. the corresponding links in the graph are bidirectional). This implies that in online social networks, the distribution of outgoing links is similar to that of incoming links. The symmetric nature of the social links affects the network structure: symmetry increases the overall connectivity of the network and typically reduces its diameter.

Based on these observations, we propose to build a bidirectional power law (PL) OSN graph topology from a classic DHT based OSN graph (Wuala or Vis à Vis). We develop a distributed rewiring algorithm using a heuristic inspired by the Barabasi Albert model [12] that transforms an existing DHT graph to a bidirectional power law graph. Our mechanism can thus fit existing P2P DHT based OSN networks like Vis à Vis (based on Pastry DHT) or SPROUT [13] (based on Chord DHT). In what follows, we show how our idea can be applied to Chord [14], producing its power law DHT variant. Note that the general idea is applicable to other structured P2P systems.

3.1 Topology Construction

To construct a scale-free DHT topology, the first step is to modify the neighbor selection. Instead of selecting the calculated neighbor according to the respective DHT algorithm, we choose one of the node's sequential neighbors with probability p_i proportional to their full degrees (outdegree plus indegree). In Chord for example, when a node is selected to be a finger, instead of using this node directly like in Chord, we choose one of the finger's successor $succ_i$ with probability p_i proportional to their degree. This results in the Power Chord graph (figure 1).

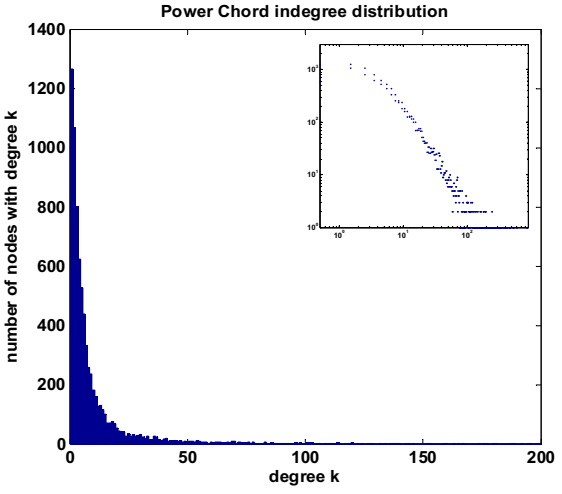


Fig. 1. Power Chord degree distribution

Then, to make this changed new Power Chord graph bidirectional, for each usual neighbor link, we add the *reverse link* linking the same two nodes but in the opposite direction. We use this additional knowledge for routing and object search in the resulting topology.

Therefore, using social network analysis, we can identify important nodes in the network and exploit the advantage of their structural significance. Routing optimization in such networks is a considerable aspect, as it will enable the network to act faster. However, this result in a trade-offs that exist when using central nodes in the network for routing optimization: while the routing efficiency increases, the unfairness of the protocol on the central nodes is affecting their resource usage. In fact, the number of the links varies considerably between nodes. For Power Chord for example, 80% of nodes have a maximum of 22 links, while only 5% of nodes maintain a more than 55 links. As a result, only 5% of nodes would have to maintain large node state information, which represent a small tiny proportion of the network.

In what follow, we present our novel methodology for efficient information retrieval and diffusion in P2P based OSN. We exploit the proprieties of the resulting power law OSN graph to optimize DHT routing and information search, to allow efficient advanced search and information diffusion in a distributed environment.

4 OSN Information Retrieval and Diffusion

4.1 Exact Information Retrieval

The DHT based approaches structure the node graph according to the structure of the stored information, and bind addresses and (virtual) placement of the participating nodes to the content they store. DHTs then employ different additional structuring methods to produce an interconnection graph with small-world properties (routing in

log N hops). Theoretically, this allows deterministic calculation of optimal routes to the content and the respective destination nodes: it therefore implements the exact or key-based search.

DHT routing is a well known efficient way to find responses to exact queries like friend profiles or groups. In the proposed new power law graph, we propose to optimize the classic DHT routing by using the additional bidirectional links as follows: when receiving a query, the intermediate node first examines its reverse links to find the final destination if it exists: if so, the query is transmitted directly to the corresponding reverse destination, as it is the owner of the searched content. Thus, the lookup search is resolved. If not, the node forwards the query to its closest preceding neighbor or reverse neighbor (using both forward and reverse links). In this case, high connected nodes will accelerate the routing process since they have a large knowledge of the network graph.

4.2 Advanced Search

In an OSN network, the complete content identifier is often not known at the moment of the search issue, or when the search targets a group of objects by specifying conditions on different dimensions of content criteria. In such situations, which we refer to as "advanced search", the key-based search method and generally, the DHT-typical key-based routing (KBR) cannot be easily used. The support for advanced search would require additional costly structures like reverse-indices, unite-tables, etc., difficult to implement and maintain in practice.

An interesting idea would be to employ flooding, but using the resulting DHT structure to optimize the search cost. In search by flooding, the source node sends the query to all its neighbors in the forward table and reverse routing tables. If the neighbors do not have the requested item, they send on to their respective neighbors excluding the source node. In this case, the query is very likely to reach a node more than once, known to cause an exponentially increasing network load. To render such flooding more efficient, we adapt and implement to power-law DHT graph our constrained flooding scheme described in [15]. The latter performs floods on structured topologies but exploits structural constraints to ensure that nodes are visited only once during a query. By assigning a disjoint search interval for each node, the number of visited nodes is reduced and the number of redundant forwarding is controlled accurately.

Our constrained flooding scheme can be applied to SPROUT social network (based on Chord DHT). Moreover, in can be adapted to other DHTs and consequently to other DHT based OSNs.

4.3 Information Diffusion

Our proposed constraint flooding scheme would avoid redundant forwarding and thus reduce the search cost. The same scheme can be also employed to efficiently diffuse information in OSN groups or between OSN friends. In Vis à Vis for example, users form overlay networks, one DHT overlay (Pastry) for every OSN group. To employ the constrained flooding in Pastry and so in Vis à Vis, a node S floods a query by sending the query to all nodes x in the routing table of the adequate row r_j (row with

the longest prefix). When an intermediate node receives the query, it recursively forwards the query to nodes of adequate row r_i where $i > j$.

In this case, using the constrained flooding scheme, but applied to Pastry for Vis à Vis, the user can diffuse information (photos or videos) to targeted users of a given group in a decentralized way without causing exponential load increase.

5 Simulation Results

To evaluate our P2P-based OSN implementation proposal and the proposed search and diffusion techniques applied to Chord (Power Chord) and Pastry (Power Pastry) as examples, we have performed numerous experiments based on simulation. We run simulations using Oversim, a flexible overlay network simulation framework based on OMNeT++ [17]. We evaluated the proposed search techniques (constraint flooding/diffusion and KBR routing) using the following criteria:

Search cost: the total number of signaling messages sent per second in the network produced by the classic full flood scheme (like in Gnutella) and our constraint flooding scheme proposed for Power DHT. Note that the results of the constraint flooding scheme are equivalent to those of the diffusion scheme.
Constraint flood path length: number of hops for constraint flooding search.
Request success rate for KBR routing: The total fraction of requests for which the source node receives an answer.

Figure 2 presents the results for constraint flooding search and figure 3 present the simulation results of KBR routing.

First, we measured the search path length as a function of network size N. We simulated a network with $N = 2^k$ nodes and varied k from 9 to 14. During every experiment, each node picked up a random set of keys to lookup employing constraint flooding in Power Chord and Power Pastry graphs respectively. Figure 2 (a, c) plots the average constraint flood path length for both Power DHTs. We note that Power DHT constraint flood search achieves shorter routes for both algorithms compared to the classic full flood scheme. In fact, for the largest network, constraint flood in Power Pastry (fig. 2 c) and Power Chord (fig. 2 a) achieves 41% and 39% less hops respectively. The fact that the query traverses essentially highly connected nodes expands the search region and accelerates the flooding search.

Moreover, constraint flooding in Power Chord and Power Pastry results in paths of about *log log N* hops (fig. 2 a, c) as would be expected in scale free networks. As a result, when diffusing information in an OSN group, the query reaches all targeted users in *log log N* hops. That means that a user can reach any friend neighbor in *log log N* hops.

We also measured the total signalling cost during simulation time for constraint flooding. From figures 2 (b, d), we note that for both search algorithms, the number of sent messages is increasing with the network size. For comparison, we plotted the signaling cost generated by classic full flooding in Power Chord and Power Pastry. We notice that constraint flooding scheme generates less signaling cost than classic full flooding (about 30% less messages in average), since it eliminates redundant forwarding and consequently produces less flooded messages.

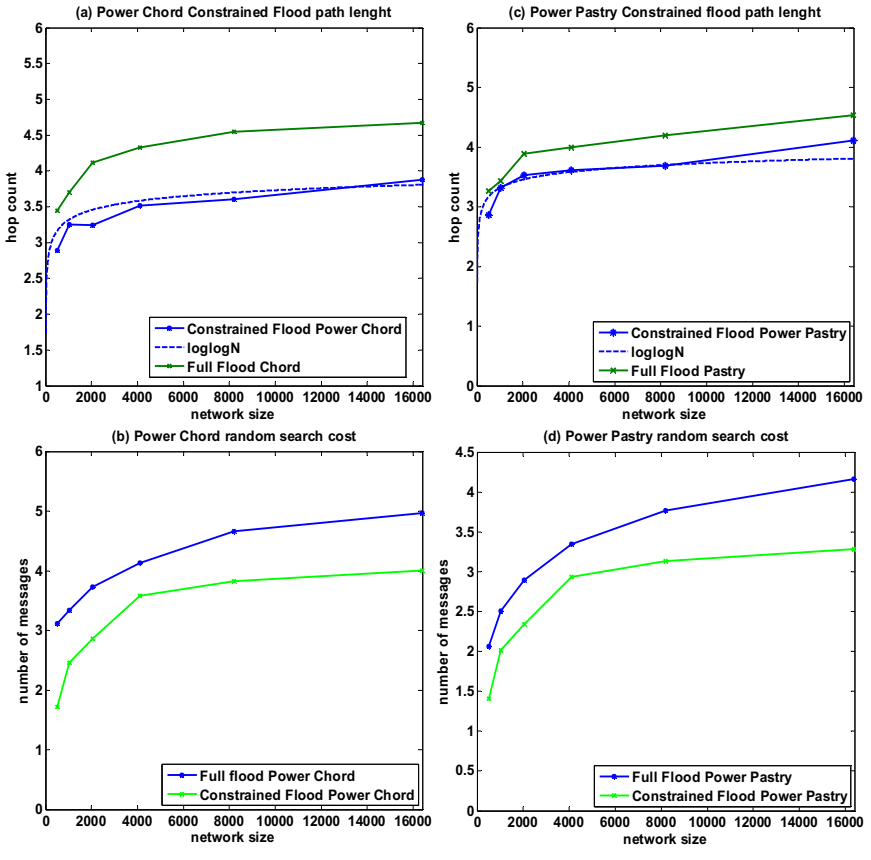


Fig. 2. Advanced search simulation results

To evaluate KBR routing, we measured the KBR path length as a function of network size N. We simulated a network with $N = 2^k$ nodes and varied k from 9 to 14. During every experiment, each node picked up a random set of keys to look up using Chord, Pastry, Power Chord and Power Pastry KBR lookup algorithm respectively.

Figures 3 (a, d) plot the average query path length as a function of N. We note that Power Chord and Power Pastry outperform classic Chord and Pastry respectively in term of average search path length. The fact that the query traverses essentially highly connected nodes expands the search region at each intermediate hop (to more than *log N* neighbors), and accelerates considerably the query process.

We measured also the request success rate as a function of node churn. We simulated a network of 2^{14} nodes and varied the mean lifetime (the node lifetime follows a Poisson distribution). This metric reflects the protocol's ability to deliver packets and to resolve multiple-hop queries in face of churn, i.e. the protocol's resilience to failure. From figures 3 (b,e), we notice that the request success rate is reduced with node churn. Power DHT variants outperform classic DHTs especially Power Pastry. In fact Power Pastry features over 95% success for all life time values,

while for Power Chord, the success ratio degrades to less than 70%. Since in Pastry a node has a larger routing table than in Chord, and so a larger neighborhood, there are more alternative next hop nodes and consequently node failure affects less the routing process.

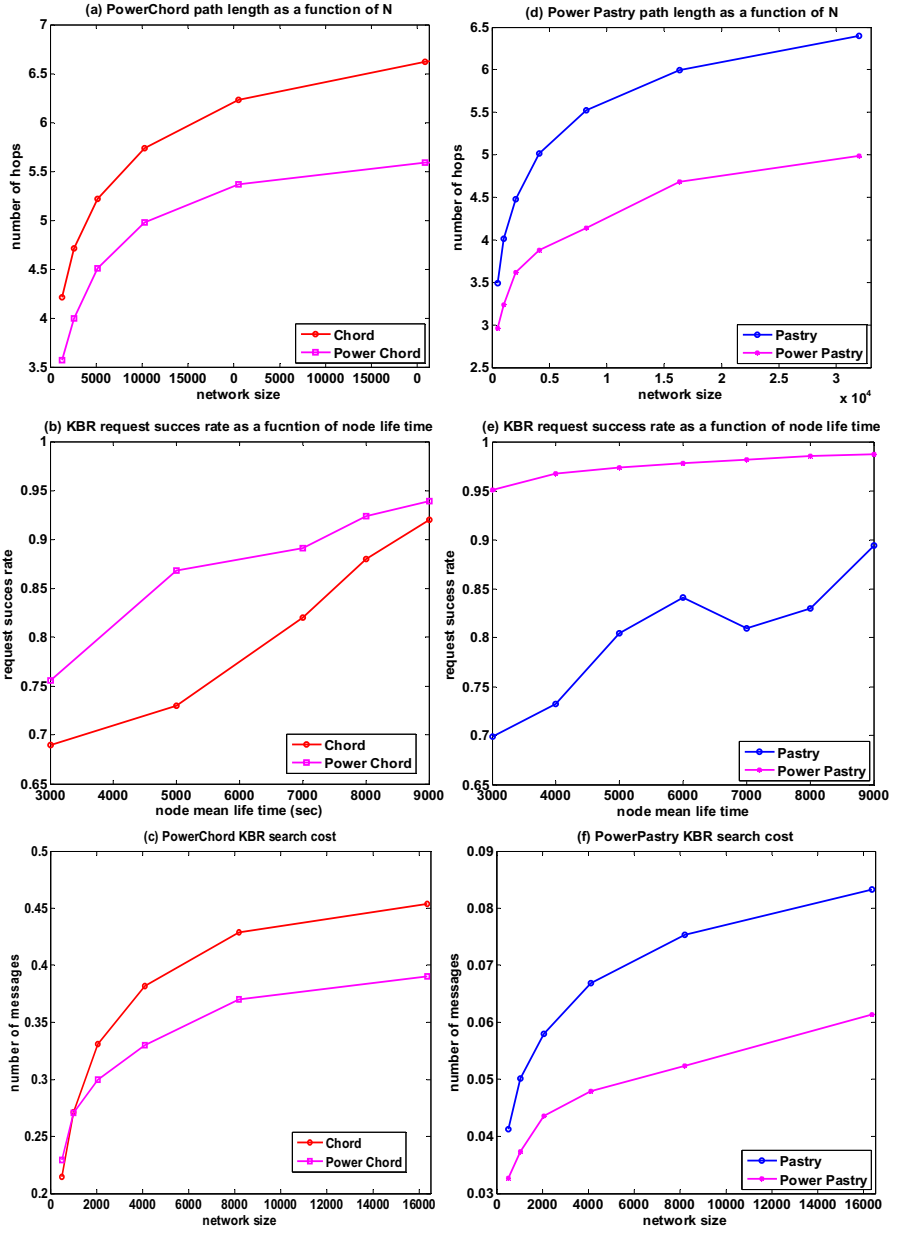


Fig. 3. KBR search simulation results

Finally, we measured the KBR search cost during simulation time. From figures 2 (c,f), we note that for all algorithms, the number of sent messages is increasing with the network size. For larger networks, however, the network load generated by Power DHT variant is considerably smaller. Since power law DHTs produce shorter routes, the number of forwarded query messages is smaller, and consequently, the total load is reduced.

5 Conclusion

This paper investigates the realization of advanced search and information diffusion techniques in P2P based OSN systems, which are often required in real applications.

First, we propose a new P2P topology model named Power DHT that matches real OSN graph. Power DHT can be adapted to different existing algorithm by transforming an existing DHT into a scale free like graph, by rewiring the overlay connections. The new topology decreases the graph diameter and thus the route length by rendering the graph power law. Second, we propose new methodologies for exact search through KBR routing and information diffusion using constraint flooding algorithm.

Simulation results show at the examples of two distinct DHTs that the proposed new graph results in a more efficient advanced search, achieving improvements over classic DHTs. For instance, the average path length is shortened to about *log log N* for constraint flooding. Moreover, the signaling cost is reduced.

The resulting power-law structure might amplify the non-uniformity of node usage. Yet, given the node heterogeneity of real-world networks, it is interesting to use the more resourceful nodes more extensively. Besides, it remains possible for a node to reject assignment. This issue is not treated within this paper. An additional proposal protocol is subject of a future work.

References

1. Shakimov, A., Lim, H., Cox, L.P., Caceres, R.: Vis-à-Vis:Online Social Networking via Virtual Individual Servers. In: Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds (MobiHeld 2009), ACM, New York (2009)
2. Buchegger, S., et al.: Implementing a P2P Social Network - Early Experiences and Insights from PeerSon. In: Second ACM Workshop on Social Network Systems (Co-located with EuroSys 2009), Nuremberg, Germany, March 31 (2009)
3. Cutillo, L.-A., Molva, R., Struf, T.: Safebook: A Privacy-preserving Online Social Network Leveraging on Real-life Trust. Published in the IEEE Communications Magazine
4. Figueiredo, R., Boykin, O., Juste, P.S., Wolinsky, D.: Social VPNs: Integrating overlay and social networks for seamless P2P networking. In: Proc. of COPS (2008)
5. Guha, S., Tang, K., Francis, P.: NOYB: Privacy in online social networks. In: Proc. of WOSN (2008)
6. `http://www.wuala.com/`
7. Guclu, Y.: Scale-Free Overlay Topologies with Hard Cutoffs for Unstructured Peer-to-Peer Networks. In: 27th ICDCS Conference (2007)

8. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J.: Graphstructure in the web. In: 9th International WWW Conference, Amsterdam, May 15-19 (2000)
9. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: Proceedings of the 7th ACM SIGCOMM Conference on internet Measurement, San Diego, California, USA, October 24 - 26 (2007)
10. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the Internet topology. In: Proceedings of the 1999 Conference on Applications, technologies, architectures, and protocols for computer communications, pp. 251–262 (1999)
11. Ripeanu, M., Foster, I., Iamnitchi, A.: Mapping thegnutella network: Properties of large-scale peer-to-peer systems and implications for system design. IEEE Internet Computing Journal 6(1) (2002)
12. Adamic, L.A., Lukose, R.M., Puniyani, A.R., Huberman, B.A.: Search in power-law networks. Physical Review E 64(4), 46135 (2001)
13. Marti, Sergio and Ganesan, Prasanna and Garcia-Molina, Hector. SPROUT: P2P Routing with Social Networks. In: First International Workshop on Peer-to-Peer Computing and Databases (P2P&DB (2004)
14. Gummadi, K., Gummadi, R., Gribble, S., Ratnasamy, S., Shenker, S., Stoica, I.: The impact of DHT routing geometry on resilience and proximity. In: Proceedings of the, Conference on Applications, Technologies, Architectures, and Protocols For Computer Communications (2003)
15. Ktari, S., Zoubert, M., Hecker, A., Labiod, H.: 'Symmetric replication for efficient flooding in DHTs'. In: Mobihoc, Honkong China (2008)
16. Crainiceanu, A., Linga, P., Machanavajjhala, A., Gehrke, J.: Querying Peer-to-Peer Networks Using PTrees. In: Proc. 7th Int. Workshop on the Web and Databases (WebDB), Paris, France, pp. 25–30 (July 2004)
17. OverSIM, `http://www.oversim.org/wiki/OMNeT%2B%2B`

Good Friends, Bad News - Affect and Virality in Twitter

Lars Kai Hansen, Adam Arvidsson, Finn Aarup Nielsen,
Elanor Colleoni, and Michael Etter

DTU Informatics, Technical University of Denmark,
DK-2800 Lyngby, Denmark
lkh@imm.dtu.dk fn@imm.dtu.dk
University of Milan, via Conservatorio Milan, Italy
adam.arvidsson@unimi.it
Copenhagen Business School, DK-2000 Frederiksberg, Denmark
elc.ikl@cbs.dk me.ikl@cbs.dk

Abstract. The link between affect, defined as the capacity for sentimental arousal on the part of a message, and virality, defined as the probability that it be sent along, is of significant theoretical and practical importance, e.g. for viral marketing. The basic measure of virality in Twitter is the probability of retweet and we are interested in which dimensions of the content of a tweet leads to retweeting. We hypothesize that negative news content is more likely to be retweeted, while for non-news tweets positive sentiments support virality. To test the hypothesis we analyze three corpora: A complete sample of tweets about the COP15 climate summit, a random sample of tweets, and a general text corpus including news. The latter allows us to train a classifier that can distinguish tweets that carry news and non-news information. We present evidence that negative sentiment enhances virality in the news segment, but not in the non-news segment. Our findings may be summarized 'If you want to be cited: Sweet talk your friends or serve bad news to the public'.

1 Introduction

Viral communication is already of significant practical importance and the scientific interest is increasing. The scientific interest derives in part from the light shed by viral communication on meme diffusion and opinion formation. In the growing viral communication literature there has been an interest in affect and emotion as crucial factors behind successful viral diffusion. The idea has been that people are more likely to send on material that like 'pets, sex and the absurd' is able to actualize a common experience of affective arousal, be this laughter, compassion, anger or surprise. This idea that affectively charged viral messages are more likely to spread than affectively neutral ones has been present within the more anecdotic marketing literature [1, 9], as well as in the more systematic aggregations of qualitative studies that have arisen in recent years. For example, in their study of word of mouth (WOM) marketing, Kozinets et al. [14] argue that such messages are more likely to be taken seriously and further diffused by consumers once they are inserted within a network of affectively significant relations that is able to transform messages from 'persuasion oriented 'hype' to relevant, useful, communally desirable social information that

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 34–43, 2011.

builds individual reputations and group relationships'. It is in this 'transformation of a market narrative into a social one' that 'the WOM communicator performs [the] services [that] are valuable to the marketers'[14]. Similarly, based on a qualitative analysis of nine viral marketing campaigns, Dodele et al. [5] conclude that the key to success is the ability to stir up an emotional arousal among the people who pass along the message. Indeed, as Vincent Miller [20] argues, communication on social media can be primarily understood as 'phatic', that is, geared towards the creation and consolidation of affectively significant relations, rather than towards the transfer of information. Only recently however has the hypothesis of the effect of affective charge on viral diffusion been tested in a large-scale quantitative study. In their recent work, Berger and Milkman [2] use a sample of 6; 956 articles from the New York Times articles published between August 30th and Nov 30th 2008. The authors conclude that there is a strong link between affect, as measured by a sentiment analysis of article content, and whether content is highly shared; as defined as whether the particular article made the New York Times' list of 'most emailed' articles. They also conclude that positive content is more viral than negative content, but that articles with some negative content, like anger or anxiety are both more likely to make the paper's most emailed list. These results hold controlling for how surprising, interesting, or practically useful content is, as well as external drivers of attention, like how prominently articles were featured. Berger and Milkman's suggestion that affective charge has a discernable impact on viral diffusion is corroborated by a host of quantitative studies. It is also supported by common sense. In a social media environment where social relations have effectively become a medium of communication, content that is more likely to activate such relations is also more likely to spread. Interestingly a similar relationship has been suggested, if implicitly, by a long range of sociological research on the diffusion of ideas and innovations, from the work of Tarde [32] via that of Lazarsfeldt [16] and Rogers [25] without the terms 'virality' or 'affect ever being used. However, Berger and Milkman also claim that their conclusions about the link between positive affect and viral diffusion has a general validity, and they found this link on psychological theories, like the claim that 'consumers often share things to self-enhance [...] or communicate identity, and consequently positive things may be shared more because they reflect positively on the self '[2]. In other words, they suggest that it is part of human nature to share positive content more often than negative content. However, their conclusion appears to be in contrast with classic theory of selection and diffusion in news media [8], which emphasizes negative affect as promoting propagation. This applies in particular to the theory of news factors. News factors are 'relevance indicators that not only serve as selection criteria in journalism but also guide information processing by audience' [6]. News factors research can be traced back to 1922, when Lippman [18] introduced the first model of news values. In part inspired by cognitive psychology it has developed into several different models, see e.g., [8, 29, 26] that aim to explain why certain events obtain more media attention or readers' awareness then others. The notion of news factor negativity has been introduced by Galtung and Ruge in 1965 [8] and since then has been taken into account by other models [6, 26, 29] partly splitting up negativity in subcategories as for example conflict, damage, aggression or failure. Several studies show that the news factor negativity has a significant impact on readers' awareness as well as on journalists' selection [10, 22,

28, 30]. Straughan's study [30] shows that news items containing conflict had a significant positive impact on reader interest in news stories. Similar, Schulz [27] found that events with the news factor aggression obtained high awareness. Further, the role of negativity in news selection has been highlighted by Peterson [22], showing that conflictual events are rather published than cooperative events. In a comparison of different news factors Harcup and O'Neill [10] range negativity among the highest news factors. The most recent study on news selection by Schwarz [28] provides evidence that the news factor negativity correlates significantly with assigned text space in news media. The contrast between Berger and Milkman's [2] findings and established research on news factors can be resolved by distinguishing between different communication media. We suggest that while Berger and Milkman's findings might apply to email networks that are marked by strong reciprocity, not all networks of information diffusion in social media are reciprocal. Such platforms may support diffusion dynamics that are more similar to those of conventional news media.

In recent years, microblogging, where users can describe their current status in short posts, has emerged as a new important communication platform [13]. The most popular microblogging platform is Twitter. Twitter connects both friends like in the symmetric Facebook network and users with shared interest like one-way structures like news groups or conventional blogs. In-links in Twitter are called followers, while out-links are called friends or 'followings'. Business entrepeneurs Naval Ravikant and Adam Rifkin [23] suggest that Twitter's value is increased by the fact that Twitter is in part an interest graph, thus revealing more of the users behaviors than a purely friends-based social graph. Their notion of an interest graph describes a network that differs from a social graph in four important respects: One-way following rather than two-way reciprocal relationship; it is organized around shared interests, not personal relationships; it is public by default, not private by default; it is 'aspirational': not who you were in the past or even who you are, but who you want to be [23].

Without making reference to Ravikant and Rifkin, Weng et al. challenges the notion of Twitter as an interest graph. Based on a quantitative study they reveal many similarities between Twitter and conventional social networks [33]. Weng et al. study a corpus of tweets created by the most followed twitterers in Singapore. The total number of tweets collected from these accounts was 1,021,039, with the majority dated between April 2008 to April 2009. Weng et al. find that a large fraction (>74%) of their population share symmetric links with the majority of their friends (>80%), furthermore they find notable assortative mixing also referred to as homophily, i.e., the tendency to follow twitterers with the same number of followers as yourself, and also a powerlaw degree distributions; both are characteristics of social networks. Yet, in a topic model of the tweet texts, Weng et al. report evidence that Twitter friends share interests (topics), hence, connect with the idea that Twitter relations reveal similar interest and behaviors. Thus while Twitter users might entertain reciprocal relations with some followers, information diffusion through retweets tends to expand far beyond this circle of reciprocity, and proliferate through the 'interest graph', among followers with whom no reciprocal relations are entertained [12]. A large scale study by Kwak et al. [15] supports the notion that Twitter is not simply social network. They crawled the entire 'Twitter sphere' to find no sign of a global power

law link distribution and they report low reciprocity in contrast to the results of the more geographically localized study of Weng et al. [33].

Retweeting, is the mechanism by which a user quotes another user and is one of the important means of meme propagation and opinion formation in Twitter. Meme propagation was studied recently in context of blogs and news media by Lescovec al. [17]. In contrast to the use of quotes ("..") used to track memes in the work of Lescovec et al., in Twitter it is explicitly indicated which user is cited by use of either the 'RT @user' or 'via @user' notations. Retweet practice is a topic of significant current interest. Honeycutt and Herring [11] and Boyd et al. [3] both discuss retweet syntax, and the role of retweeting in Twitter discourse. Boyd et al. note three mechanisms supporting meme propagation, including the use of topical tags '#topic', mentions of given user '@user', and finally the use of shortened URL's to allow receivers to access background sources of a given meme. Boyd et al. explicitly asked a Twitter community what and why they retweeted. The feedback revealed a quite complex set of retweeting mechanisms, including (self-) branding, general news interest, and encouraging social activity amongst many other reasons. A more quantitative study of meme propagation in Twitter is reported in Suh et al. who build simple models of the probability of being retweeted [31]. Suh et al. collected 10,000 tweets and subsequently found that 291 of these had been retweeted (mere 3%). They build a model of retweet probability based on context variables mentioned above and further includes basic aspects of the graph structure, i.e., the number of followers and friends of the tweeter. The study confirms that inclusion of URL and hash tags both improve the probability of retweet, while explicit mentions of another user seem to reduce retweet probability, although this finding is only a trend in the data ($p \sim 0.07$). Thus Twitter is an interesting and complex communication platform serving both as a social network and as a new medium of information sharing. Thus when asking 'what are the determining factors for my message to be retweeted' it may depend on both on the type of content and whether the communication is intended for a broader audience or for a more closed community of friends. In case of addressing a broader audience with news content the message sentiment may be an important determinant. Following this analysis we set out answer the following research questions

- Q1 How accurately can text be characterized as 'news'?
- Q2 How big a fraction of Twitter is news?
- Q3 If Twitter is a news medium, does negative sentiment influence virality?
- Q4 Does sentiment influence retweet probability diffentially in news and social messages?

The paper is organized as follows: In the next section we define the methodology, including Twitter samples and the statistical models. In the following section we present the results regarding news classification and sentiment detection, and in a final section we discuss the findings, provide conclusions and ideas for further study.

2 Methods

To address the research questions Q1-Q2 we will use a simple machine learning method - a Naive Bayes classifier - to detect whether a message is 'news'. This

classifier is trained and tested from a labeled text corpus, the Brown Corpus. The trained classifier is applied to two medium large Twitter samples containing a 'complete sample' of a discourse driven by a news event and a sample of randomly selected tweets, respectively. In both Twitter samples we estimate sentiments and we build generalized linear models to investigate whether sentiment influences virality in terms of the retweet probability.

We use three different text corpora to test the above questions,

- C1 BROWN, a general text corpus with a known mixture of news / non-news documents [7]. The corpus consist of N_B = 47.134 sentences of which 4.623 are categorized as news.
- C2 COP15 a Twitter data set that is designed to comprise a complete set of tweets for a specific news driven vent. COP15 refers to the The 2009 United Nations Climate Change Conference that took place in Copenhagen, Denmark, between December 7. and December 18. The conference included the 15th Conference of the Parties (COP15) to the United Nations Framework Convention on Climate Change. The conference and the international activism that took place during the conference were extensively covered in news media and in the blogosphere. A total of N_C = 207.782 tweets were downloaded during the month of December 2009 by querying the Twitter Search API with the term 'cop15'.
- C3 RANDOM a random sample of tweets. N_R = 348.862 tweets were downloaded with the Twitter streaming API during the time interval from September 9. To September 14., 2010. The Tweets were randomly sampled following the 'Spritzer' protocol.

The sentiment analysis is designed to work on English tweets. Downloaded tweets are provided with an indication of the language. However, for many non-English tweets the field is set to its default value 'English'. We therefore first constructed a language filter: After download a language detector scored each tweet for *englishness*. The language detector used a list of words manually scored from -3 to +3 indicating an estimate on how English each word is. The englishness of the words in a tweet was accumulated, and tweets with a positive sum were regarded as English. For the streaming Twitter data 106,719 remained after extracting tweets where the user had set the language to English and the language detector also detected the tweet as English. The language filter eliminated only a few percent of the COP15 tweets, while more than 60% of the RANDOM data was removed as non-English, see table 1 for details.

To determine the 'newsness' of a tweet we trained a classifier on the sentences of the Brown corpus with the NLTK toolkit [19]. The news category of the Brown corpus has 4.623 sentences, while other categories have a total of 42.511 sentences (we excluded the 'editorial' category). The classifier is trained on a bag-of-terms representation. A stop word list of 571 words excluded common words and the 10,000 most frequent words were extracted from the corpus and used as terms. For the RANDOM data 'lol', 'love', and 'good' were the most frequent terms, while for the COP15 data the most frequent terms were 'obama', 'world', and 'deal', reflecting

the latter is driven by a news event. A sentence is represented in the classifier as D = 10.000-dimensional term vector $\boldsymbol{w}$ where the d'th entry, w_d is either 1 or 0 depending on whether the d'th term is present or not in thegiven sentence. We use the so-called Naive Bayes classifier based on univariate discrete distributions [24]. Splitting the corpus in 75% for training and the rest for testing, the NLTK naive Bayes classifier will report the test error which is an unbiased measure of performance. We also apply the trained classifier on the tweet data getting a probability of 'newsness' for each tweet in RANDOM and COP15. Based on the set of probabilities for all tweets we calculate the rate of news tweets, i.e., fraction of tweets with $p(\text{news}|\boldsymbol{w}) > 0.5$.

For English tweets sentiment was estimated via an English word list manually curated for Twitter. Thus, we follow the classical approach used for sentiment scoring in conventional English text [4], but with a dedicated wordlist. The present list associates 1.446 words with a valence between -5 and +5. Sentiment estimation results in a *valence* and an *arousal* score for each tweet. The valence of a tweet n, v_n, is computed as the sum of the valences of the individual words in the tweet $v_{n;i}$, while the arousal a_n the sum of the absolute value of the valences.

The extracted features were included in a generalized linear model (GLM) assuming a binomial distribution with the standard link function [21]. Denoting the probability of retweet by $p(RT|\boldsymbol{f})$, where $\boldsymbol{f}$ is a set of F features derived from the tweet, the GLM estimates the posterior probability of retweet, $p(RT \mid \boldsymbol{f}) = (1 + exp(\Sigma^{F}_{i=0} f_i\beta_i))^{-1}$, where the coefficient β_0 of the 0^{th} feature $f_0=1$ ensures proper normalization. The coefficients β_i are estimated using iterative likelihood maximization. The difference in log-likelihood of two nested models is approximately χ^2 distributed for large data, hence, can be used to test hypotheses about the relevance of individual features. In particular we estimate a model with all features and F sub-models with a single feature removed. The t-statistic values in table 1 express the relevance of the given feature to the retweet model.

We formulate the modeling problem as in [31]: Imagine you are a tweeter who wants to be retweeted, how should you formulate your tweet? However, relative to [31] we introduce three modifications. First, we focus on features that actually can be manipulated at the time of tweeting, this includes the presence of '#', '@', and 'URL', but not the variables related to the graph structure, since these variables cannot be manipulated at the time of tweeting. The second modification is that we flag a tweet as a retweet if the text has a pattern with 'RT' or 'via' followed by a user name ('@user'). In [31] the authors located retweets of an initial pool of tweets. Our approach leads to a larger sample of retweets, but may suffer from an unknown bias related to text modifications introduced by the retweeting party. We expect this bias to be limited and for retweets based on Twitter's 'retweet button', there is no bias. Our criterion also includes possible 'retweets of retweets'. Finally, we include the presence of negative sentiment as a covariate, hypothesizing that this may help explain propagation as is hypothesized in news media. User names that are part of the retweet indicator are not counted as a separate mentioning ('@'). To further test the role of newsiness we create a feature which is the logical combination of news and negative.

Table 1. Estimated t-values within the RANDOM and COP15 data for the binary variable encoding the presence of negative sentiment. The t-value is obtained in a general linear model (logistic regression) of the retweet probability versus the independent binary (present/not present) co-variates: negative sentiment, hashtag, mention, and url.

	COP15		STREAM	
	All	**English**	**All**	**English**
N	147,041	136,262	335,236	106,719
Rate News	0.303	0.305	0.226	0.233
t(Negative)	4.889	4.649	2.775	-0.024
t(Negative_newsness)	2.275	1.471	-6.019	3.904
Tweets with Arousal > 0				
N	44,611	42,087	53,473	51,929
t(Negative)	3.276	2.372	-9.725	-9.374
t(Negative_newsness)	1.125	0.180	1.179	1.239

3 Results

The results are discussed in relation to the four research questions Q1-Q4. Details are summarized in table 1.

Q1: The Naive Bayes classifier can detect news. We assign 75% of the sentences in the Brown corpus for training leaving the 11783 for testing. The NLTK naive Bayes classifier reports an accuracy of 84%±1% in the test data.

Q2: 23% of all tweets are news. We apply the classifier to the tweets in the two sets RANDOM and COP15. In the RANDOM sample 23% of the tweets meet the criterion of having a probability of news larger 0.5. If biased, this could be a slight underestimate, as the fraction of news items (i.e., the a priori class probability) in the Brown corpus is about 0.1. In the COP15 sample we find that a larger number of tweets, 31%, are detected as news.

Q3: **Negative sentiment does not promote retweeting in the RANDOM sample.** We trained a number of generalized linear models to test for the significance of negative sentiment in the COP15 and the RANDOM samples. In the RANDOM sample the weight for presence of negative sentiment is slightly negative, however, the t-statistic indicates that we cannot reject the null that there is no effect of negative sentiment. However in the COP15 sample negative sentiment promotes retweeting and the positive weight is highly significant. If the analysis is restricted to the subset of tweets that have non-zero arousal, i.e., positive or negative content, the tendency found in the RANDOM sample of all tweets is strongly amplified, it is a strong promoter of retweet if content is positive.

Q4: Negative sentiment does promote retweeting in news tweets in both the RANDOM and the COP15 samples. To test for the interaction between newsiness and negative sentiment we create a new multiplicative variable (negative sentiment)_

(probability of news). This interaction variable promotes retweeting in both COP15 and RANDOM. However, if we confine ourselves to the aroused tweets the conclusion persists even if it was strongly favored to have positive content in the sample at large.

4 Discussion and Conclusion

There is an ongoing discussion on the graph structure and dynamics in the Twitterverse. A large scale quantitative analysis led Kwak et al. [15] to conclude that Twitter can be viewed both as a social network and a news medium, while the more localized study of Weng et al. [33] found structures more reminiscent of a social network. At the same time the work of Berger and Milkman [2] and general psychological arguments favor the sharing of non-negative content among friends, while classical theories of news diffusion points to increased attention to news if content is negatively framed. This seemingly paradoxical set of results are reconciled in our findings. Using a trained classifier that can reliably detect news, we find that about a quarter of the tweets in a random sample are news, while the more focused sample of tweets relating to the global news event COP15 has a higher fraction nearing one third of news tweets. We note that these figures for news content are higher than estimated in a small informal investigation by San Antonio based analytics firm Pear Analytics. We are in process of developing additional ways of estimating and validating newsness. The differences in news content are reflected in how much we can accelerate retweet by negative framing of content. In a random sample of tweets there is a slight tendency that negative content hinders retweet, albeit it is only a trend, while in the more news driven COP15 sample, negative content is a strong promoter of retweeting. If we look at the interaction term, we find that negative news is more retweeted than positive news. For the non-news segment which may be dominated by social tweets our results support the idea that positive content increase the probability of diffusion [2], whereas in the news segment, our findings confirm the impact of negativity on news awareness and selection as proposed by classic news diffusion theory [6, 8, 26, 29].

In conclusion we find that a simple Naive Bayes classifier can quite reliably detect the presence of news in short communication. We find a relative high amount of news content: 23% in random tweets and 33% in the COP15 data. Investigating the probability of retweeting we find that for generic tweets negative sentiment is detriment to retweeting, while news related content propagates better if negative. *Hence, if you want to be cited: Sweet talk your friends or serve bad news to the public!*

Acknowledgement. We thank the Danish Strategic Research Councils for their generous support to the interdisciplinary project 'Responsible Business in the Blogosphere'. We apologize for the poor typographical quality of the paper induced by the required "MS Word" template. A typeset version of the paper can be obtained at http://arxiv.org/abs/1101.0510

References

[1]. Ahonen, T., Moore, A.: Communities Dominate Brands:Business and Marketing Challenges for the 21st Century. FutureText, London, UK (2005)

[2]. Berger, J., Milkman, K.: Social transmission, emotion, and the virality of online content. Wharton Research Paper (2010)

[3]. Boyd, D., Golder, S., Lotan, G.: Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In: Hawaii Int. Conf. on System Sciences, pp. 1–10 (2010)

[4]. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): Stimuli, instruction manual, and affective ratings. Technical report, Center for Research in Psychophysiology, University of Florida, Gainesville, Florida (1999)

[5]. Dobele, A., Lindgreen, A., Beverland, M., Vanhamme, J., Vanwijk, R.: Why pass on viral messages? because they connect emotionally. Business Horizons 50(4), 291–304 (2007)

[6]. Eilders, C.: News factors and news decisions. theoretical and methodological advances in germany. Communications: The European Journal of Communication Research 31(1), 5–24 (2006)

[7]. Francis, W.N., Kucera, H.: Brown corpus manual. Technical report, Department of Linguistics, Brown University, Providence, Rhode Island, US (1979)

[8]. Galtung, J., Ruge, M.: The structure of foreign news: The presentation of the congo, cuba and cyprus crises in four norwegian newspapers. Journal of Peace Research 2(1), 64–90 (1965)

[9]. Godin, S.: Unleashing the Ideavirus (2001)

[10]. Harcup, T., O'Neill, D.: What is news? galtung and ruge revisited. Journalism Studies 2(2), 261–280 (1981)

[11]. Honeycutt, C., Herring, S.C.: Beyond microblogging: Conversation and collaboration via twitter. In: Proc. 42nd Hawaii Int. Conf. on System Sciences, pp. 1–10. IEEE Computer Society Press, Washington, DC, USA (2009)

[12]. Huberman, B., Romero, D., Wu, F.: Social networks that matter: Twitter under the microscope (December 2008), http://arxiv.org/pdf/0812.1045

[13]. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: WebKDD/SNA-KDD 2007, pp. 56–65. ACM, New York (2007)

[14]. Kozinets, R., de Valck, K., Wojnicki, A., Wilner, S.: Networked narratives. understanding word of mouth marketing in online communities. Journal of Marketing 74, 71–89 (2010)

[15]. Kwak, H., Lee, C., Parkand, H., Moon, S.: What is Twitter, a social network or a news media? In: WWW 2010: Proceedings of the 19th International Conference on World Wide Web, pp. 591–600. ACM, New York (2010)

[16]. Lazarsfeldt, P., Katz, E.: Personal Influence. The Free Press, Glecoe (1955)

[17]. Leskovec, J., Adamic, L., Huberman, B.: The dynamics of viral marketing. ACM Transactions on the Web (2007)

[18]. Lippman, W.: Public Opinion. MacMillan, New York (1922)

[19]. Loper, E., Bird, S.: NLTK: The natural language toolkit. In: Proc. of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics, vol. 1, pp. 63–70. Association for Computational Linguistics, Morristown (2002)

[20]. Miller, V.: New media, networking and phatic culture. Convergence: The Int. Journal of Research into New Media Technologies 14(4), 387–400 (2008)

[21]. Nelder, J.A., Wedderburn, R.W.M.: Generalized linear models. Journal of the Royal Statistical Society, Series A, General 135, 370–384 (1972)
[22]. Peterson, S.: International news selection by the elite press: A case study. Public Opinion Quarterly 45(2), 143–163 (1981)
[23]. Ravikant, N., Rifkin, A.: Why twitter is massively undervalued compared to facebook (2010), http://techcrunch.com
[24]. Rish, I.: An empirical study of the naive Bayes classifier. In: Int. Joint Conf. on Artificial Intelligence, pp. 41–46 (2001)
[25]. Rogers, E.: Diffusion of Innovations. The Free Press, New York (1962)
[26]. Ruhrmann, G., Woelke, J., Maier, M., Diehlmann, N.: Der Wert von Nachrichten im deutschen Fernsehen: Ein Modell zur Validierung von Nachrichtenfaktoren. Leske Budrich, New York (2003)
[27]. Schulz, W.: News structure and people's awareness of political events. International Communication Gazette 30, 139–153 (1982)
[28]. Schwarz, A.: The theory of newsworthiness applied to Mexico's press. Communications: The European Journal of Communication Research 31(1), 45–64 (2006)
[29]. Staab, F.: The role of news factors in news selection: A theoretical reconsideration. European Journal of Communication 5, 423–443 (1990)
[30]. Straughan, M.: An experiment on the relation between news values and reader. International Communication Gazette 43, 93–107 (1989)
[31]. Suh, B., Hong, L., Pirolli, P., Chi, E.H.: Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In: Social Computing / IEEE Int. Conf. on Privacy, Security, Risk and Trust, pp. 177–184 (2010)
[32]. Tarde, G.: The law of Imitation. Holt and Company, New York (1903)
[33]. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twitterrank: finding topicsensitive influential twitterers. In: Proc. of the Third ACM Int. Conf. on Web Search and Data Mining, WSDM 2010, pp. 261–270. ACM, New York (2010)

Semantic Verification in an Open Collaboration Scenario*

Jorge Coelho[1,2] and Mário Florido[1,3]

[1] Artificial Intelligence and Computer Science Laboratory of the University of Porto (LIACC)
[2] School of Engineering of the Polytechnic of Porto (ISEP)
[3] University of Porto, DCC-FC
{jcoelho,amf}@liacc.up.pt

Abstract. In this paper we describe a framework for imposing constraints in the content of webpages and show its adequacy in open collaborative scenarios such as wikis. This framework is the extension of the XML processing language XCentric with syntactic validation, semantic verification and automatic correction of websites. The highly declarative model of XCentric along with the domain-specific verification techniques enable an agile implementation of constraints over web content.

1 Introduction

For sites which rely on collaborative work such as wikis, verifying that its content obeys to some criteria can be a tedious task. Imagine for example verifying the content of a wiki like Wikipedia [11]. There, almost anyone can change any page, that is the essence of the knowledge sharing that Wikipedia encourages but also one of its biggest problems. Not all the content posted in Wikipedia has high quality standards and a constant effort to review all the data is needed. An automatical tool for content reviewing and filtering by applying predefined constraints may be useful in this context. In this paper we present a framework for building such tool relying on syntactic validation and semantic verification of website content with the ability to automatically repair the web pages that don't obey to a given criteria.

The remaining of this paper is organized as follows, in section 2 we explain briefly the main concepts behind the XCentric language. Then, in section 3 we present our extension to verification of content in wikis. In section 4 we present the related work and finally in section 5 we conclude and present future work.

We assume that the reader is familiar with Logic Programming [14].

2 XCentric

XCentric extends Prolog with terms with flexible arity symbols and sequence variables. This approach enables a highly declarative model for querying content in

* Partially funded by LIACC through Programa de Financiamento Plurianual of the Fundação para a Ciência e Tecnologia (FCT).

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 44–53, 2011.

webpages. Constraints of the form $t_1 =^* = t_2$ are solved by a non-standard unification that calculates the corresponding minimal complete set of unifiers. Details about the implementation of this non-standard unification can be found in [6]. In XCentric an XML document is translated to a term with flexible arity function symbol. This term has a main functor (the root tag) and zero or more arguments. Although XCentric translates attributes to a list of pairs, we will omit them for the sake of simplicity. Consider the simple XML file presented bellow:

```
<addressbook>
     <record>
          <name>John</name>
          <address>New York</address>
          <email>john.ny@mailserver.com</email>
     </record>
          ...
</addressbook>
```

The equivalent term is:

```
addressbook(record(name('John'),address('New York'),
                          email('john.ny@mailserver.com')),...)
```

If we want to get the names of the people living in New York and assuming that the document is stored in variable *Doc* we can simply solve the following constraint:

```
Doc=*=addressbook(_,record(name(N),address('New York'),_),_).
```

All the solutions can then be found by backtracking (in variable N).

Note that "_" is an unnamed sequence variable which unifies with any sequence. So, no matter how many records the address book has, we can describe our constraint in a very compact way by focusing on the ones that matter for our purposes. The details of the language and several illustrating examples can be found in [9].

In [7] and [8] XCentric was extended with several features:

- Definition of simple rules for website verification and filtering namely, replacing, deleting and blocking content.
- Use of types for static and dynamic verification of rules.
- Consistency checking between rules (one rule cannot violate another rule).

The type system allows checking the content against schemas and the consistency checking verifies if one rule is not in violation of another rule, for example, when one rule adds some content which is forbidden by another rule.

3 Verification of Content in Wikis

We now proceed showing how XCentric can be used to query and verify content in web pages. We added some simple builtins to XCentric in order to make the verification tasks simpler. A brief description is presented in table 1.

Table 1. Builtins for Verification

Builtin	Descrition
http2pro(U,X)	Get webpage from URL in U and translate it to the internal notation in X.
xml2pro(F,X)	Get XML document from file F and translate it to the internal notation in X.
pro2xml(X,F)	Translate the internal notation for XML document X into the file F.
bind_type(X,T)	Binds the document X to the declared type T.
modified(U,S)	For the URL in U, get the date of last modification and bind it with S.
get_url(T,U)	Parse the text T and extract a (maybe empty) listof URLs to U.
verify_url(U)	True if URL U does not return code 404 (not found).

3.1 Examples

The next examples show the adequacy of XCentric to query, process and verify simple constraints over websites.

Example 1 (Searching data in a document). Given XML-based Wiki page in figure 1 Searching for example the bibliographic references can be done with the simple query (Doc stores the XML file):

```
Doc =*= WikiArticle(_,References(_,bibentry([_],C),_).
```

Resulting in:

```
C = Jorge Coelho and Mario Florido. XCentric: Logic
Programming for XML Processing. 9th ACM International
Workshop on Web Information and Data Management. ACM Press,
2007;

C = SWI-Prolog, http://www.swi-prolog.org/
```

Note that CDATA is seen as an atom and thus its content is not processed.

Example 2 (Content conversion between different formats). Being a relational language, XCentric can easily describe the structures of two documents related with two different wikis and convert data between both. For example, we can write the following simple predicate in XCentric for converting between fragments of two kinds of wikis.

```
translate(<section(title(N),C1)>,<chapter(name(N),C2)>):-
        section_content(C1,C2).
section_content(<>,<>).
section_content(<A1,content(A),A2>,
        <L1,chapter-content(B),L2>):-
                subsection_content(A,B),
         section_content(<A1,A2>,<L1,L2>).
```

```
subsection_content(<>,<>).

subsection_content(<A1,subsection(C),A2>,
                   <B1,subchapter(C),B2>):-
                  subsection(<A1,A2>,<B1,B2>).
```

This relates a *section* element and a *chapter* element, where the pattern of the first argument of *translate* requires the *section* element to contain a *title* element followed by a sequence of *subsection* elements, and the second argument describes the structure of the *chapter* containing a *name* element followed by a sequence of *subchapter* elements.

Variable *N*, which appears in both arguments, specifies that its corresponding subparts, the contents of *title* and *name*, are the same. Predicate *section_content* relates *subsection* in a content element with *subchapter* in the corresponding *chapter* element.

Note the gain in modularity: this predicate can be used in three different ways. 1) to transform an XML document with the format specified by the first argument of *translate* into the format specified by its second argument, 2) to do the opposite transformation, or 3) to guarantee that two different documents in the two different formats are related in the way specified by the predicate (corresponding, respectively, to call it with the first, second or both arguments ground). In a functional-based language these three different behaviors have to be implemented by three different functions.

Example 3 (Verifying at arbitrary depth). XCentric also provides predicates that allow the programmer to find a sequence of elements at arbitrary depth, to search for the *nth* occurrence of a sequence of elements and to count the number of occurrences of a sequence. The predicates are *deep/2*, *deepp/3* and *deepc/3*, respectively.

If we want to know if a reference in the text has a corresponding bibliographic entry one can simply use the following conjunction:

```
deep(ref([],X),XML),deep(bibentry([(number,X)],_),XML)
```

Here we are searching for a sequence of the form <*ref*>*X*</*ref*> in document *XML* and checking if that reference appears in some sequence <*bibentry number="X"*> ...</*bibentry*> in the same document.

If we want to say that it must have only one occurrence in the bibliography then we do:

```
deep(ref([],X),XML),deepc(bibentry([(number,X)],_),XML,1)
```

Note that the 1 in *deepc* means the number of tags *bibentry* must be exactly one.

Verifying at arbitrary depth reveals to be a useful feature, the following examples further illustrate this claim.

Example 4 (Verifying webpage properties). One can verify if the year in a reference containing a URL is valid, for example, by seing if the webpage pointed by the URL wasn't created after the date that appears in the reference (this can be done by processing the HTTP headers):

```
<?xml version="1.0" encoding="utf-8"?>
<WikiArticle>
  <Title >XCentric and Prolog-based XML Processing</Title>
  <Subtitle>One simple example</Subtitle>
  <Abstract >Here we present one example which compares XCentric with traditional
  Prolog-based XML processing.</Abstract>
  <Body>
<Section>
<Title>Introduction</Title>
<Content> XCentric <ref>1</ref> is an extension of Prolog with unification
of terms of flexible arity which enables a simpler and high level querying and
processing of XML data.</Content>
</Section>
<Section>
<Title>Using XCentric</Title>
<Content> In this section we present one simple and ilustrating example of
XCentric compared with classical Prolog.
<Subsection>
<Title>Example</Title>
<Content>Given the following catalog of books:
<![CDATA[ <catalog>
...
<book number="500">
 <author>Simon Thompson</author>
  <name>Haskell: The Craft of Functional Programming (2nd Edition)</name>
  <price>41</price>
  <year>1999</year>
</book>
  ...
</catalog>]]>
 To get all the books with two or more authors using SWI-Prolog <ref>2</ref>
 (which has a quite goodlibrary for processing XML in Prolog) we need the
 following code:
<![CDATA[
pbib([element(_,_,L)]):-  pbib2(L).
pbib2([]).
pbib2([element('book',_,Cont)|Books]):-   authors(Cont),!,pbib2(Books).
pbib2([_|Books]):-pbib2(Books).
authors([element('author',_,_),element('author',_,_)|R]):-

                    write_name(R).
write_name([element('name',_,[N])]):-     write(N),nl.
write_name([_|R]):-write_name(R).]]>
To do the same in XCentric, assuming that the XML file is translated to a term
binding variable $BibDoc$, the following query is enough (note that =*= implements
the flexible arity unification):
<![CDATA[
catalog(_,book(name(N),author(_),author(_),_),_) =*= BibDoc.
]]>
All the solutions can then obtained, one by one, by backtracking. The simplicity
and declarativeness of the second solution speaks by itself when compared to the
first one.
</Content>
</Subsection>
</Content>
</Section>
  </Body><References>
  <bibentry number = "1">Jorge Coelho and Mario Florido. XCentric: Logic
   Programming for XML Processing. 9th ACM International Workshop on Web
   Information and Data Management. ACM Press, 2007.</bibentry>
   <bibentry number = "2">SWI-Prolog, http://www.swi-prolog.org/</bibentry>
  </References>
</WikiArticle>
```

Fig. 1. XML-based wiki document

```
verify_years(Url,XML,N):-
    modified(Url,Stamp),
    deep(bibentry([_],X),XML),
    get_year(X,Y),
    compare(Y,Stamp).
```

This program gets the time of modification of the page in the address given by URL which is related to the bibliographic entry with number *N* and compares both dates. Using a similar strategy one could also automatically include years in the bibliographic references which don't have any, by using the year of creation of the web page (by processing the HTTP headers in a similar way).

Example 5 (URL verification). One could also verify if URLs in bibliography references correspond to valid links by simply using:

```
verify_url(XML):-
    deep(bibentry([_],X),XML),
    get_url(X,U),
    verify_url(U).
```

This works by searching bibliographic entries with URLs and verifying if they are valid by checking if the status is not 404.

3.2 Website Constraints

Definition 1. We define a set of website constraints (SWC) as follows:

- A main rule whose input is a web page (WpageI) and output is a new web page (WpageO) resulting from the input page with the necessary changes in order to obey to the set of constraints imposed:

```
swc(WpageI,WpageO):-
        r1(WpageI,Wpage1),
        …
        rn(Wpage(n-1),WpageO).
```

- Each rule *ri* imposes some action to be taken in case some set of constraints is violated. We call these rules *action rules*. The rule may change the original document in order to make it obey to the constraint set.

Note that *WpageI* and *WpageO* are expressed using our internal representation for XML data and thus *WpageI* resulted from translating a web page from a URL or file by using one of the internal builtins available in the framework.

We now define three builtins associated with action rules:

- *delete*(*S*,*WpageI*,*WpageO*,*L*) - deletes sequence *S* which respect the constraints in *L* from *WpageI* resulting in *WpageO*.
- *replace*(S_1,S_2,*WpageI*,*WpageO*,*L*) - replace sequence S_1 which respect the constraints in *L* by sequence S_2 in *WpageI* by *WpageO*.
- *failure*(*WebpageI*,*L*,*Mesg*) - used when the error is to serious to be automatically solved. Message *Mesg* is shown when the constraints in *L* are violated.

Elements in *L* are, for example, tests in the values in the sequences in order to verify they follow a certain criterion.

Example 6 (Delete rule). Using the same Wiki page of example 1 one can, for example, delete all the references which do not occur in the text by using:

```
delete(ref(R),Wiki1,Wiki2,[not(deep(bibentry([(number,R)],_),Wiki1))]).
```

So, if we have the following XML stored in variable *Wiki1*:

```
<?xml version="1.0" encoding="utf-8"?>
<WikiArticle>
....
<Content> XCentric <ref>3</ref> is an extension of Prolog with unification of terms
of flexible arity which enables a simpler and high level querying and processing of
XML data.
</Content>
...
<References>
  <bibentry number = "1">Jorge Coelho and Mario Florido. XCentric: Logic Programming
  for XML Processing. 9th ACM International Workshop on Web Information and Data
  Management. ACM Press, 2007.</bibentry>
   <bibentry number = "2">SWI-Prolog, http://www.swi-prolog.org/</bibentry>
  </References>
</WikiArticle>
```

By applying the delete and since a reference with number 3 is not available in the references at the bottom of the page it will result in a new XML document in variable *Wiki2*:

```
<?xml version="1.0" encoding="utf-8"?>
<WikiArticle>
....
<Content> XCentric is an extension of Prolog with unification of terms of
flexible arity which enables a simpler and high level querying and processing
of XML data.
</Content>
...
<References>
  <bibentry number = "1">Jorge Coelho and Mario Florido. XCentric: Logic
Programming for XML Processing. 9th ACM International Workshop on Web
Information and Data Management. ACM Press, 2007.</bibentry>
   <bibentry number = "2">SWI-Prolog, http://www.swi-prolog.org/</bibentry>
  </References>
</WikiArticle>
```

where the element *<ref>3</ref>* was deleted.

The *replace* and *failure* rules work in an analogous way. The type system allows checking the content against schemas and the consistency checking verifies if one rule is not in violation of another rule, for example, when one rule adds some content which is forbidden by another rule.

Let's now proceed with a global example containing several rules.

Example 7. Let's use again the Wiki web page presented in example 1. We want to verify the following constraints:

- The document should have non-empty tags referring to sections and subsection titles. In case some title is empty, a warning text should appear in that title content.

- When a link is given in a bibliographic entry it should correspond to an existing web page.
- Subsections without any content should me removed.

We can build the program presented in figure 2 to verify the above constraints over the wiki's webpage.

```
start:-
        xml2pro('./examples/wiki.xml',Wiki),
        bind_type(Wiki,WikiType),
        swc(Wiki1,Wiki2),
        pro2xml(Wiki2,'wikichecked.xml').
swc(In,Out):-
        r1(In,Out1),
        r2(Out1,Out2),
        r3(Out2,Out).
r1(A,B):-
 replace(title(X),title('Title is missing!'),A,
           B,[deep(title(X),A),X=[]]).
r2(A,A):-
failure(A,[deep(bibentry(X),A),get_url(X,Url),not(valid_url(Url))],
                'Valid link not found!').
r3(A,B):-
   delete(subsection(X),A,B,[deep(subsection(X),A),X =*= <_,content([]),_>]).
```

Fig. 2. Wiki's page constraints

It works as follows: the file is retrieved and stored in variable *Wiki* using builtin *xml2pro*. Then, *bind_ type(Wiki,WikiType)* binds the webpage *Wiki* to the associated XML Schema referenced by *WikiType*. Without *bind_type* the compiler ignores the type information and thus would not be possible to find errors at compile time. The rules are described in the remaining lines of the program. Rule *r1* verifies if the titles of subsections have some content, in case they don't, the content of that tag is filled with the message "Title missing!". Rule *r2* verifies if the links in bibliographic entries are not valid by checking if they don't return a 404 error, in case they fail the verification is canceled and the error message "Valid link not found!" is shown. Note that *valid_url/1* accepts a list of URLs for validation and also succeeds with the empty list, which means that is true if no URL is found in the bibliographic content. Rule *r3* searches subsection entries in the wiki webpage which don't have content and deletes them.

3.3 Static Verification of Action Rules

Static verification is made by static analysis of action rules. These rules possibly change the document by means of a delete or replace, thus by having type information describing the document structure (by means of an XML Schema which is not presented due to space limitations), it is possible to verify the validity of these transformations.

Example 8. Using the same scenario as the one presented in example 1. Errors are reported if for example we try to apply some of the following actions:

- *delete(abstract(X),XML1,XML2,[]),* since *abstract* is a tag that cannot be deleted from the document as imposed by the corresponding schema.
- *replace(body(B),references(R),XML1,XML2,[])*, since the *body* element is mandatory as imposed by the corresponding schema.

Warnings are reported if we try the following action:

- *delete(bibentry(B),XML1,XML2,[internal_database(DB),not(DB =*= db(_, bibentry(B),_))]),* here the goal is to delete all the bibliographic entries that don't appear in another XML file in variable *DB*. A warning shows that, if all the bibentry's for a given document are covered, then the XML produced will be invalid since it will have no bibentry elements (at least one must occur as imposed by the related XML schema).

4 Related Work

The work presented here is based on a previous approach that appears in [7] and [8]. There, the theoretical foundations of our declarative approach and how it can be applied to general website verification are described. Our contribution in this paper is the adaptation and application of the previous work to verification of content in wikis. In [2] the authors presented a rewriting-based framework that uses simulation [12] in order to query terms, this was a new rewriting-based language quite different from ours. In [3], the authors present a semi-automatic methodology for repairing faulty websites by applying a set of concepts from Integrity Constraint [17]. In [10] the author proposed the use of a simple pattern-matching-based language and its translation to Prolog as a framework for website verification. In [18] logic was proposed as the rule language for semantic verification. There the authors provide a mean for introducing rules in a graphical format. In [13] the author proposed an algorithm for website verification similar to [5] in expressiveness but based in a different theoretical approach. The idea was to extend sequence and non-sequence variable pattern matching with context variables, allowing a more flexible way to process semistructured data. In [16] the authors present a tool for verification of websites based on a subset of the Xcerpt language [4] and Abductive Logic Programming [15]. The reader can find a detailed comparison between several approaches to verification in [1].

5 Conclusion

We presented a framework for verification of content in wikis using XCentric. Errors can be automatically solved by introducing actions that are executed whenever an error is found. Due to their nature, wikis are vulnerable to vandalism, thus, our framework can be applied to automatically filter content on these websites optimizing the reviewing process and helping to keep high quality standards.

References

1. Alalfi, M.H., Cordy, J.R., Dean, T.R.: Modelling methods for web application verification and testing: state of the art. Softw. Test., Verif. Reliab. 19(4), 265–296 (2009)
2. Alpuente, M., Ballis, D., Falaschi, M.: A Rewriting-based Framework for Web Sites Verification. In: ENTCS, pp. 41–61. Elsevier Science, Amsterdam (2005)
3. Alpuente, M., Ballis, D., Falaschi, M.: Rule-based verification of web sites. STTT 8(6), 565–585 (2006)
4. Bry, F., Schaffert, S.: The XML Query Language Xcerpt: Design Principles, Examples, and Semantics. In: 2nd Annual International Workshop Web and Databases. LNCS, Springer, Heidelberg (2002)
5. Bry, F., Schaffert, S.: Towards a declarative query and transformation language for XML and semistructured data: Simulation unification. In: Stuckey, P.J. (ed.) ICLP 2002. LNCS, vol. 2401, p. 255. Springer, Heidelberg (2002)
6. Coelho, J., Florido, M.: CLP(Flex): Constraint logic programming applied to XML processing. In: Chung, S. (ed.) OTM 2004. LNCS, vol. 3291, pp. 1098–1112. Springer, Heidelberg (2004)
7. Coelho, J., Florido, M.: VeriFLog: A constraint logic programming approach to verification of website content. In: Shen, H.T., Li, J., Li, M., Ni, J., Wang, W. (eds.) APWeb Workshops 2006. LNCS, vol. 3842, pp. 148–156. Springer, Heidelberg (2006)
8. Coelho, J., Florido, M.: Type-based static and dynamic website verification. In: The Second International Conference on Internet and Web Applications and Services, IEEE Computer Society, Los Alamitos (2007)
9. Coelho, J., Florido, M.: Xcentric: A logic-programming language for xml processing. In: Grust, T. (ed.) ACM SIGPLAN Workshop on Programming Language Technologies for XML, PLAN-X (2007)
10. Despeyroux, T.: Practical semantic analysis of web sites and documents. In: Feldman, S.I., Uretsky, M., Najork, M., Wills, C.E. (eds.) WWW, pp. 685–693. ACM, New York (2004)
11. Wikimedia Foundation. Wikipedia: Multilingual, web-based, free content encyclopedia project. World Wide Web (2010), `http://www.wikipedia.org/`
12. Henzinger, M.R., Henzinger, T.A., Kopke, P.W.: Computing simulations on finite and infinite graphs. In: FOCS, pp. 453–462 (1995)
13. Kutsia, T.: Context sequence matching for xml. In: Proceedings of the 1th Int. Workshop on Automated Specification and Verification of Web Sites (2005)
14. Lloyd, J.W.: Foundations of Logic Programming, 2nd edn. Springer, Heidelberg (1987)
15. Mancarella, P., Terreni, G., Sadri, F., Toni, F., Endriss, U.: The ciff proof procedure for abductive logic programming with constraints: Theory, implementation and experiments. TPLP 9(6), 691–750 (2009)
16. Mancarella, P., Terreni, G., Toni, F.: Web sites repairing through abduction. Electr. Notes Theor. Comput. Sci. 235, 137–152 (2009)
17. Mayol, E., Teniente, E.: A survey of current methods for integrity constraint maintenance and view updating. In: Akoka, J., Bouzeghoub, M., Comyn-Wattiau, I., Métais, E. (eds.) ER 1999. LNCS, vol. 1728, pp. 62–73. Springer, Heidelberg (1999)
18. van Harmelen, F., van der Meer, J.: Webmaster: Knowledge-based verification of web-pages. In: Imam, I., Kodratoff, Y., El-Dessouki, A., Ali, M. (eds.) IEA/AIE 1999. LNCS (LNAI), vol. 1611, pp. 256–265. Springer, Heidelberg (1999)

Designing Persuasive Third Party Applications for Social Networking Services Based on the 3D-RAB Model

Isaac Wiafe, Keiichi Nakata, and Stephen R. Gulliver

Informatics Research Centre
Henley Business School
Whiteknights
University of Reading
Reading, UK
i.wiafe@pgr.reading.ac.uk,
{k.nakata,s.r.gulliver}@henley.reading.ac.uk

Abstract. The use of social networks services for promoting business, teaching, learning, persuasion and spread of information continues to attract attention as most social networking services (SNSs) now allow third party applications to operate on their sites. In the field of persuasive technology, the ability of SNSs to build relationships among their users and create momentum and enthusiasm through rapid cycles also give it a greater advantage over other persuasive technology approaches. In this paper we discuss the 3-dimensional relationship between attitude and behavior (3D-RAB) model, and demonstrate how it can be used in designing third-party persuasive applications in SNSs by considering external factors which affects persuasive strategies.

Keywords: Social Networking Services, Persuasive Technology, Behavior modeling.

1 Introduction

The application of social networking services (SNS) continues to spread across various domains as recent SNSs allow third party applications to operate on their sites. Researchers now develop applications to promote teaching and learning, business, marketing, commerce, and news feed, among others. Recently, persuasive technology designers also have implemented third party persuasive applications on SNSs so as to reach a larger audience [1]. Persuasive technology is the use of computing technology to change or shape a person's attitude and/or behavior [2], which has been mainly used in healthcare, environmental and other social purposes [3]. Some persuasive technology designers design their applications to operate on SNSs due to their ability to create momentum and enthusiasm through rapid cycles that promotes behavior and attitude change [4]. Competitive carbon counting [5] is an example of a persuasive technology developed as a third party application designed to persuade users to reduce their energy usage. However, despite the potential of such applications, there is little research on what methods or approach that such third party applications should follow, when designing for persuasion.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 54–61, 2011.

In this research we contribute to the existing study on the use of SNSs as a persuasive technology by demonstrating the need to understand the relationship between a user's attitude towards change, his/her attitude towards change in behavior or maintaining the behavior as compared to his/her current behavior. We demonstrate how the 3D-RAB model [6] can be applied to the design of third party persuasive applications on SNSs.

The remainder of the paper is presented as follows. The next section presents discussions on current research contributions in SNS relating to persuasive technologies. This is followed by a discussion on factors that affect persuasion. Section 3 is on the 3D-RAB model and it is followed by discussions and a conclusion.

2 Persuading via Social Networking Services

Modern computing network technology has made social networks less bounded by providing tools of create and maintain relationships between individuals irrespective of geographical boundaries with social support, information, and a sense of belonging [7]. SNSs are now characterized by mass participation and typical functionalities such as personal profiling of participants, networking of friends, private messaging, discussion forums, events management, blogging, commenting and media uploading [8] enable them to promote persuasive activities. This is because they serve as a platform for discussions and sharing of ideas within a larger community across geographical boundaries. As such both direct and indirect persuasions are possible. What differentiates them from many other persuasive technology platforms is that they are inherently collective and thus make use of group dynamics a powerful factor in the context of persuasion [9].

A research conducted by Fogg explained that the use of Mass Interpersonal Persuasion (MIP) have also become prominent in the area of SNS [4]. MIPs propose six components which promote SNSs as a perfect platform for persuasive activities, namely, persuasive experience, automated structure, social distribution, rapid cycle, huge social graph and measured impact. Persuasive experience is a form of experience that is created to change attitudes, behaviors, or both. As such the creator of the experience aims at making impact on people's lives. Persuasive experience can then be structured as a digital technology allowing the software to present the experience repeatedly, which is known as automated structure. Automated structure enables easy sharing of experiences with other users within a social network, whereas social distribution enhances the ability for the persuasive experience to be shared between peers on the network. In rapid cycling, persuasive experiences are distributed quickly within the network. Fogg argued that the experience is therefore capable of reaching millions of people who are connected through social ties thereby creating a huge social graph. Also the impact of the experience is readily observable by both users and designers as the system provides information on connected peers. He added that though these components existed before, some SNSs have bundled them together and made them more useful for third party persuasive applications [4].

Another research [10] explained that persuasion happens in predictable ways and as such identified patterns of persuasion in online social networks and concluded that there are six patterns of SNSs that serve as persuasive tools. Four of the patterns lead

to a large-scale viral adoption and are native to SNSs. These are: provoke and retaliate, reveal and compare, expression, and group exchange. The other two patterns which are adopted are competition and deception. In addition, in order to design a meaningful persuasive SNSs application, [9] explained that consideration should be given to a system in terms of its affiliation, access and participation.

2.1 Factors Affecting Persuasion

Persuasive technology has introduced a new direction in the study of psychology in relation to information systems design, its usage and acceptance. Nowadays, systems are designed to promote psychological principles and theories in order to change a user's behavior or attitude to a targeted one. For instance, ObiMoPet [11], PowerHouse [12] and UbiGreen [13] are examples of such systems. In this research, the focus is on the relationship between attitude and behavior of users and external factors within the SNS community in relation to the development of third party persuasive applications that promote socially positive behavior. We argue that even though SNSs promote persuasion [4, 9], there is a need for third party persuasive technology designers who aim at using SNSs as the environment to consider the factors discussed below in order to follow a systematic procedure and avoid undesirable effects which can be created by the SNS's environment. Below we discuss the 3D-RAB model [6], which advocates that there are three factors that can be considered for analyzing the cognitive state of a user, namely, user current behavior, user attitude towards target behavior and user attitude towards change or maintaining the change. These three factors are derived from the application of the cognitive dissonance theory which explains the mutual influences between attitude and behavior as explained later in Section 3. There are also two external factors, natural attitude/behavior change and planned attitude/behavior change, which are also discussed in relation to the model.

2.2 User Current Behavior (CB)

One of the factors to consider when designing a third-party persuasive application is users' current behavior (CB). CB is defined as the existing actions of a person in relation to the environment and may be conscious or subconscious, overt or covert, voluntary or involuntary. The designer should consider the target behavior in relation to potential users, i.e. whether users are currently performing the behavior or not. This measure can be positive or negative in relation to the target behavior. Consider an application aiming to promote drivers from not driving under the influence of alcohol. Users who are currently drinking and driving are considered to be exhibiting a negative behavior whereas those not doing so are considered otherwise.

2.3 User Attitude Towards Target Behavior (ATTB)

User attitude towards target behavior (ATTB) is defined as the like or dislike of target behavior; and it can be positive, negative or neutral. If user's attitude supports the target behavior then they are deemed as having a positive attitude towards the target behavior. Again considering the case of drink-driving, users with a positive ATTB are those who believe and accept that one should not drive under the influence of alcohol. The case of the neutral attitude is not considered in the model for simplicity.

2.4 User Attitude Towards Changing/Maintaining Current Behavior (ATCMB)

A user's attitude towards change or maintaining behavior (ATCMB) is a measure of agreement or disagreement of the user in relation to a particular change or maintenance in behavior and can be positive, negative or neutral. In the case of drink-driving, ATCMB is considered to be positive when a user agrees to change to the target behavior (not drink and drive) from the negative behavior or maintain the current positive behavior (not drink and drive), and negative otherwise. Again the neutral attitude was not considered. ATCMB provides information on a user's readiness to change, or to maintain the existing behavior [14]. Positive ATCMB promotes self-efficacy and it has been found to be an important determinant of choice of activities in which users engage, how much energy they expend on such activities, and the degree of persistence they demonstrate in failure and/or adversity [15].

2.5 External Factors in Behavior and Attitude Change

Apart from the 3 internal factors mentioned above, we also consider external factors which could serve as triggers for behavior and attitude changes. As such they can influence users in developing negative or positive behavior or attitude. In other words, since the persuasive application will be operating in an SNS environment there is the need to consider the operations of the SNS itself and identify how it may promote or hinder the persuasive application. In this regard we consider natural attitude or behavior change (NABC). NABC is the behavior or attitude change that may naturally occur as users of SNS interacts with other users in or out of the SNS environment. This may be governed by, e.g., norms, peer pressure, and culture of the SNS community to which the user belongs. Users' experiences result in constant change in their behavior and/or attitude as they interact with others.

Another external factor which could affect persuasive applications in the SNS environment is planned attitude or behavior change (PABC). PABC is found in situations where activities are planned to change people's attitude or behavior; these activities may or may not be automated. Within an SNS community, users form groups in which they plan their activities in a particular way. They are governed by policies and rules of the organization or the associations in order to enable them behave in a particular manner. In designing third party persuasive applications for SNSs, this factor should be considered because it may also promote or hinder the progress of the persuasion. From the above internal and external factors, we present a model which describes the 3-dimensional relationships between the internal factors and how external factors act on them. This is intended to inform designers on how to design persuasive third party applications for SNSs. Section 3 presents the 3D-RAB model and its external factors.

3 The 3D-RAB Model

The 3D-RAB model can enable an application designer to categorize SNS users into groups based on cognitive dissonance states during design so as to present persuasive messages and techniques to support the transition towards a particular target state. In total, 8 different user states can be identified based on the values (positive or negative) of CB, ATTB, and ATCMB (table 1).

According to the theory of cognitive dissonance, two cognitions are considered to be in dissonance if one opposes the other and this creates an unpleasant psychological tension which generates a motivational urge to minimize dissonance by changing an attitude, a belief or behavior [16, 17]. Hence by combining cognition with motivation, attitude and behavior change can be achieved [18]. Based on this premise we argue that cognitive dissonance can be at different levels, which when studied carefully, can be used for selecting appropriate persuasive techniques. Also, these levels of dissonance provide information concerning possible natural transitions that exist as a result of influence of factors which are external to a system. There are four levels of cognitive dissonance, which are strong, moderate, weak and no dissonance.

Table 1. Eight states in the 3D-RAB model

State	CB	ATTB	ATCMB	Cognitive Dissonance	Stability of State	Natural State Transition Tendency	Targeted State
1	+	+	+	No	Stable (+)	1	1
2	+	+	-	Weak	Unstable (+)	1	1
3	+	-	+	Moderate	Unstable (-)	7	1
4	+	-	-	Strong	Unstable (-)	8	2 or 3
5	-	+	+	Strong	Unstable (+)	1	1
6	-	+	-	Moderate	Unstable (-)	8	2 or 5
7	-	-	+	Weak	Unstable (-)	8	3 or 5
8	-	-	-	No	Stable (-)	8	4 or 6 or 7

Strong cognitive dissonance is formed when there is a strong disagreement between one's attitude (either ATTB or ATCMB) and behavior and it results in a strong unpleasant psychological tension with a greater probability that one may change his attitude or behavior in order to eliminate the dissonance. At such a state the user experiences an uncomfortable cognitive state that he or she recognizes the need for a change in attitude or behavior. A moderate dissonance is formed when there is also a disagreement between one's attitude and behavior; however the extent of unpleasant psychological tension experienced in this case is relatively moderate, hence the urge to change attitude or behavior is not intense. In weak dissonance, although there is a form of dissonance it is not significant and the psychological tension created is weak. When one's attitude agrees with his behavior there is no psychological tension. At such a state there is no dissonance. The variation in dissonance creates both stable and unstable states which can be positive or negative in relation to the target behavior. In a positive unstable state a user is more likely to change either his attitude or behavior to favor the target behavior whereas it is vice versa in a negatively unstable state. See table 1.

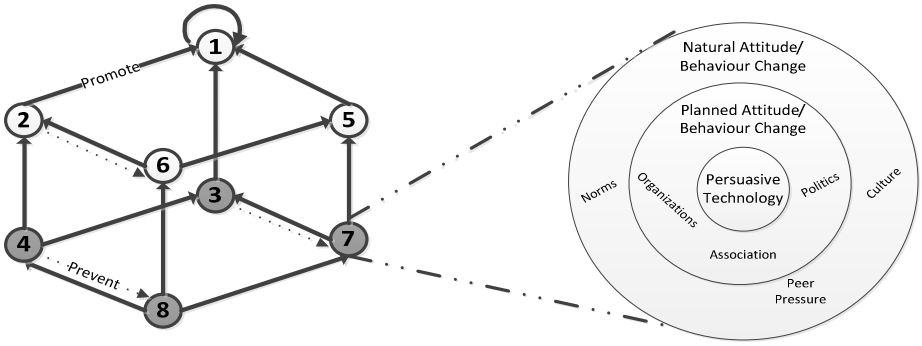


Fig. 1. Graphical representation of external factors and the 3D-RAB adapted from [6]

State 1 is considered to be the ideal state towards which the persuasion is targeted. Conversely, state 8 is considered to be the worst state in terms of persuasion. Therefore, the solid arrows in the diagram in figure 1 shows the routes of persuasion from state 8 to state 1. From table 1, it can be observed that in states 1-4, the target users are already performing the target behavior. However, the model reveals that although they are performing the behavior they have variable levels of cognitive dissonance states which can result in them changing their behavior. External factors can form triggers and make users move naturally to a particular state, which may or may not be in favor of the target. In designing persuasive applications in an SNS environment, designers can employ techniques which will focus on moving the user to the target states in table 1 and in doing so consideration should also be given to the effects of external factors in relation to the target change.

For example, in the case of drink-driving, a consideration can be given to whether or not the SNS community on which the application will be operating promotes drinking. This is because this information will enable the designers to consider how to incorporate techniques which will not condemn drinking completely, but rather focus on drink-driving. Users may ignore the application totally if it makes them believe that it is promoting drinking prohibition.

There are also situations where dissonance creates natural “pulls” or “pushes” towards the target behavior. Hence, without persuasion or persuasive technology, users in the SNS may naturally change as external factors acts on their current state. In such situations persuasive techniques should promoting directions which have natural tendencies of changing users to the target behavior, and possibly inhibit the change towards the negative behavior.

4 Discussion

As mentioned earlier, the use of SNSs as a platform for promoting behavior and/or attitude change presents an opportunity to persuasive technology to exploit their features. We argued for the need for a systematic strategy to achieve the target behavior, and proposed the use of the 3D-RAB model. One of the major challenges in using the 3D-RAB model for designing persuasive technology is to identify at which

state the user is in. In this regard, SNSs have the ability to collect user feedback which makes it possible for designers to collect information on, and possibly changes in, users and thus tailor persuasive approaches according to users' states. For instance in the case of drink-driving persuasive design, designers can implement strategies which can collect user states through daily status updates or tweets. Such information can then be used to identify the cognitive state the user is in, which in turn informs the persuasive application on the needed techniques which can be applied to effectively motivate or persuade a user. Moreover, feedback information will provide data on the effectiveness of the persuasive application. This is because designers will be in the position to identify and track progress of users as they transit from one state to another as they use the application.

Amidst these advantages of SNSs platforms for persuasion, there are some limitations which can result in undesirable effects on persuasion. Often, persuasive technology designers ignore the possibility of boomerang effects (the ability of the system to create a counter-persuasion); but this can happen in persuasive design [19] and in the case of SNS persuasion such an occurrence will have a negative effect on the target behavior relative to the persuasive system which do not operate on SNS. This is because the ability of SNS to spread information rapidly (which is useful in persuasion) can make users spread negative information about the system rapidly. In addition, existing applications are object-centered [20] and do not allow explicit connections between users from other different SNS application [8]. This limits the use of SNS for persuasive activities, since it creates a bounded community of users and their shared objects, which hampers interoperability.

5 Conclusion

In this paper we have discussed the need for third party persuasive applications designed to operate on SNSs to follow a systematic approach and demonstrated how the 3D-RAB model can be applied to such designs. We presented the relationship between the internal factors that determines the state of a user and the external factors which may promote or hinder persuasive activities naturally. However, since this model is still at the conceptual stage, subsequent research will focus on substantiating its theoretical claims with empirical evidence. Also, investigations will be conducted to identify appropriate techniques to facilitate or prevent transition of states. This will serve as a tool for designing effective persuasion which takes into account the external factors of the SNS community.

References

1. Fogg, B.: A behavior model for persuasive design. In: 4th International Conference on Persuasive Technology. ACM, California (2009)
2. Fogg, B.J.: Captology: the study of computers as persuasive technologies. ACM New York, New York (1997)
3. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems (2009)

4. Fogg, B.J.: Mass interpersonal persuasion: An early view of a new phenomenon. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 23–34. Springer, Heidelberg (2008)
5. Foster, D., Blythe, M., Cairns, P., Lawson, S.: Competitive carbon counting: can social networking sites make saving energy more enjoyable? ACM, New York (2010)
6. Wiafe, I., Nakata, K., Moran, S., Gulliver, S.: Considering user attitude and behaviour in persuasive systems design. In: European Conference on Information Systems, Helsinki, Finland (2011)
7. Wellman, B.: Computer networks as social networks. Science 293(5537), 2031 (2001)
8. Ahmadi, N., Jazayeri, M., Lelli, F., Nesic, S.: A survey of social software engineering. IEEE Computer Society Press, Los Alamitos (2008)
9. Khaled, R., Barr, P., Aumann, Y., Biddle, R.: Investigating social software as persuasive technology. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 104–107. Springer, Heidelberg (2006)
10. Weiksner, G.M., Fogg, B.J., Liu, X.: Six patterns for persuasion in online social networks. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 151–163. Springer, Heidelberg (2008)
11. Wiafe, I., Nakata, K.: A semiotic analysis of persuasive technology: An Application to Obesity Management. In: 12th International Conference on Informatics and Semiotics in Organisations, SciTePress, Reading (2010)
12. Bang, M., Torstensson, C., Katzeff, C.: The powerHhouse: A persuasive computer game designed to raise awareness of domestic energy consumption. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 123–132. Springer, Heidelberg (2006)
13. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.: UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. ACM, New York (2009)
14. Prochaska, J., DiClemente, C.: Toward a comprehensive model of change. Treating addictive behaviors: Processes of change, 3–27 (1986)
15. Oldenburg, B., Glanz, K., Ffrench, M.: The application of staging models to the understanding of health behaviour change and the promotion of health. Psychology & Health 14(3), 503–516 (1999)
16. Festiger, L.: A theory of cognitive dissonance. Row, Peterson, New York (1957)
17. Aronson, E.: Back to the Future: Retrospective Review of Leon Festinger's A Theory of Cognitive Dissonance. The American Journal of Psychology 110(1), 127–137 (1997)
18. Griffin, E., McClish, G.: A first look at communication theory. McGraw-Hill, New York (1991)
19. Segerståhl, K., Oinas-Kukkonen, H.: Distributed user experience in persuasive technology environments. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 80–91. Springer, Heidelberg (2007)
20. Cetina, K.: Sociality with objects: Social relations in postsocial knowledge societies. Theory, Culture & Society 14(4), 1–30 (1997)

A Self-adaptive Clustering Scheme with a Time-Decay Function for Microblogging Text Mining

Chung-Hong Lee and Chih-Hung Wu

Dept of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan
leechung@mail.ee.kuas.edu.tw, williamwu.tw@gmail.com

Abstract. Online microblogging services such as Twitter allow users to post very short messages related to everything ranging from mundane daily life routines to breaking news events. This phenomenon has changed the way for information acquisition. In this paper, we present an instinctive method with a time-decay function which corresponds to the natural propagation of social networks for clustering real-time text streams collected from Twitter. Compared to most previous studies, we follow natural cascading behaviors of event lifecycle to develop a self-adaptive clustering model for online event detection. Also, we construct an expandable similarity matrix which is capable of evaluating microblogging posts with incomplete semantic features. Experimental results show that the proposed method is a sensible solution to monitoring momentous real-time events and utilizing the text streams to facilitate the management of social networking data.

Keywords: real-time mining, social networks, microblogs, self-adaptive term weighting.

1 Introduction

Social networks have become one of the primary information sources and act as important tools which collect and propagate real-life information. Microblogging, which consists of user generated contents, fulfills the task for driving a faster way of communication to eliminate the gap between desktop and mobile users. This phenomenon has changed the way for information acquisition. As such a powerful information tool, a key challenge confronted in real-time event mining is the validation of consecutive message streams based on their temporal factors.

There is an increasing need for automatic real-time text mining techniques to analyze and present correlative and emergent messages to a general reader in a timely and efficient manner. For example, applying Twitter posts to learn about earthquake occurrence[1], and analyzing the properties of Twitter posts for crisis communication [2]. Meanwhile, these services also come with new challenges due to these short messages are propagated explosively and generate information cascades which reflect our social behaviors. However, there is no salient solution to evaluate the validity of an instant event from the messages because no test corpus is available or well prepared for analyzing the text streams. These difficulties might make the text mining tasks become more inefficient and unpredictable.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 62–71, 2011.

The goal of our research is to help users eliminate information overload and automatically detect real-time hot event topics from microblogging messages. In our work such continuous messages will be clustered according to their concepts, and represented as waves of real-time event groups for knowledge discovery and social networking management. These converged clusters will help users discover potential incidents what people may concern. However, our objective is not to define new events on our system, instead, we apply machine learning methods to generalize real-time streams which are propagated from microblogs. This study not only considers development of the text-stream clustering method but also detection of evolvement of real-life events. In order to reach the aims mentioned above, we applied three major methods to carry out our research work. First, we analyzed the properties of information propagation and formulated a time-decay function, which obeys the natural of the power law, to find out the event lifecycle in a social network. Second, a weighting approach is applied to deal with messages with complete and incomplete features and, fit them into the dynamic time slot. Third, an unsupervised competitive-based clustering method is designed to aggregate online messages for discovering and tracing evolvement of real-time hot events according to their concept similarity.

2 Related Work

Since lots of users have used microblogs such as Twitter, as their primary interactive and collaborative platform to exchange information, recent studies started to evaluate the social influence through diffusing linkages and contents of individuals. For instance, Java analyzed the user intensions associated at a community level and showed users connecting with each other with similar intentions[3]. Banerjee categorized content-indicative and usage-indicative keywords together to capture and disseminate users' thoughts and actions for contextual analysis[4]. Moreover, Cheong classified thousands of tweets into three categories as long-term, medium-term and short-term topics, and then traced unique message identifier (UID) to analyze trend patterns [5]. Honeycutt analyzed conversational exchanges on the public timeline and commented that microblogging evolves to become more conversational and collaborative[6]. In the study of user-to-user content transfer, [7] examined the interplay of online social networking content, and the result shows that the sharing in social graph follows the cascading behavior. Furthermore, Weng extended the idea of PageRank to their TwitterRank algorithm and reported the phenomenon of reciprocity[8]. The insight that messages published in social networking websites would descend their influence steeply into more tiers and time of propagation. It is impracticable for a user to sort through all messages without any preprocessing because the messages a person can read in a certain time is much less than the amount generated within the same period and via several sources.

In recent studies, more researchers concentrated their attentions on web stream mining but mainly extended from traditional algorithm; for instance, Gong proposed a text stream clustering algorithm TSC-AFS(text stream clustering based on adaptive feature selection) which applies incremental processing and variance-based feature selection to cluster newswire articles [9]. A news stream can be separate by evaluating their temporal fact or statistical distribution. Nallapati investigated a

clustering process that combines a time-decay function and a dependency model with an event-time-ordering function[10]. Incremental TF-IDF approaches and time windows to limit prior context for previous stories are applied to the problem domain[11, 12]. They tested a decaying-weight function where documents are removed in time have progressively less influence on the current decision. Brants used similar incremental TF-IDF approach to find new events, but they applied sliding windows to replace fixed windows[13]. However, a fixed time window might hardly satisfy real distribution of events. Up to this point, some researchers used time-based summarization to smooth the organization of news topics. For instance, Swan used the $\chi 2$ measurement to extract and generate clusters with named entities and noun phrases which capture the information corresponding to major news topics covered in the corpus[14].

3 Self-adaptive Term Weighting Scheme

Our approach is based on a competitive learning-like algorithm to formulate a time-decay function. The self-adaptive model developed in this work mainly contains clustering and term weighting techniques. Let $\boldsymbol{C} = \{\boldsymbol{c_1}, \ldots, \boldsymbol{c_{\mathcal{E} \le \mathcal{E}_{max}}}\}$ be the feature vectors of event clusters, and $\boldsymbol{\mathcal{E}_{max}}$ is the maximum permissible number of online threads. Meanwhile, these functions employ the total set of concurrent event clusters as $\forall \boldsymbol{i}\ \boldsymbol{C_i} \neq \emptyset$ and $\forall \boldsymbol{i} \neq \boldsymbol{j}\ \ \boldsymbol{C_i} \cap \boldsymbol{C_j} = \emptyset$. Let $\boldsymbol{r}$ be the latest state of processing a new message with no restriction since real-time messages are posted continuously. We list the parameters as **Table 1** to describe the related equations.

Table 1. List of parameters

Parameter	Description
$\boldsymbol{M}$	collected real-time messages
$\boldsymbol{m_r}$	latest message
$\boldsymbol{\eta}$	self-adaptive slowly varying parameter
$\boldsymbol{\alpha}$	power-law exponent
$\boldsymbol{n}$	total number of messages after $\boldsymbol{c_i}$ being created
$\boldsymbol{t_{c_i}}$	initial timestamp of $\boldsymbol{c_i}$
$\boldsymbol{t'_{c_i}}$	latest updated timestamp of $\boldsymbol{c_i}$,
$\boldsymbol{t_{m_r}}$	timestamp of $\boldsymbol{m_r}$
$\boldsymbol{N_{M,t}}$	total number of messages at time $\boldsymbol{t}$ in $\boldsymbol{M}$

Our purpose is to compare the set of $\boldsymbol{M} = \{\boldsymbol{m_1}, \cdots, \boldsymbol{m_r}\}$ to each cluster. The clusters are compared and updated each time while a new message $\boldsymbol{m_r}$ is presented to the algorithm, and the winner is the cluster that lies closest to $\boldsymbol{m_r}$**.** The winner will absorb $\boldsymbol{m_r}$ into its vector and be updated so as to evolve the real-time event cluster, and then all clusters are updated and toward to fade-away. Since the instant messages are posted continuously, a better practice is to follows the cascading behavior of social information propagation to segment each event discussing thread as mentioned above. In this study, we use a time-decay function, $d(\cdot)$, to determinate the lifecycle of an event cluster and to erase slack messages which is denoted as:

$$d\left(t_{c_i}, t'_{c_i}, t_{m_r}, n\right) = \eta n^{-\alpha} e^{\frac{-(t'_{c_i} - t_{m_r})}{1+(t_{c_i} - t_{m_r})}} \tag{1}$$

Where $\boldsymbol{\alpha > 1}$ and $\boldsymbol{\eta} = \begin{cases} 1, & n < 2 \\ ln\, n, & n \geq 2 \end{cases}$ controls the slope and optimizes our scheme according to the number of real-time posts. The similarity between $\boldsymbol{m_r}$ and $\boldsymbol{c_i}$ would be determined by a threshold $\boldsymbol{\theta}$. $\boldsymbol{m_r}$ is considered dissimilar to $\boldsymbol{c_i}$ and cannot be join to c_i accurately if:

$$sim(c_i, m_r) d(\cdot) < \theta \tag{2}$$

In the self-adaptive term weighting method, the number of messages and words are constrained in a finite scale, and each event cluster will be diminished by the time-decay function. As shown in Fig. 1, the self-adaptive model uses the features and adjusts weights in accordance with their dynamic status in the online thread.

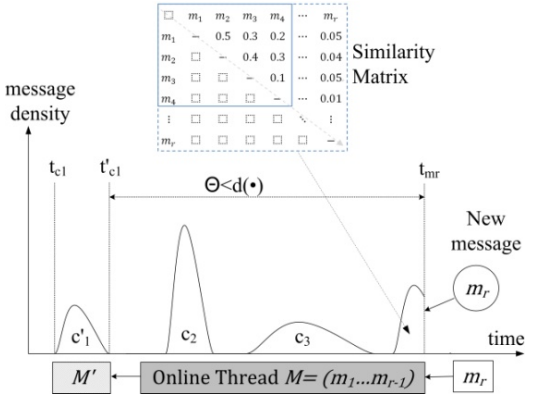


Fig. 1. Self-adaptive term weighting model

According to observations of information cascading behavior and event lifecycle of microblogs, a better practice to cluster real-time messages is to follow a human nature in information spreading (e.g. daily routines). Messages are posted as waves, and meanwhile potential events could be clustered according to their semantic relation in a limited window. Differing from previous studies used an *incremental TF-IDF*[13] with a fixed window size, each cluster and thread might vary in the amount of features. That is to say that, the system should not only input the new coming messages but also remove the faded clusters. The factor *Term Frequency-Inverse Corpus Frequency* (i.e. *TF-ICF*) [15] which references a historical document distribution to approximate the IDF value, which may not fit in real-time message streams because the global value counteracts the burst factor of real-time posts. In our self-adaptive term weighting method, document frequencies $\boldsymbol{df_{M,t}(w)}$ is not only incremental but also being adjusted by the size of online thread in time step $\boldsymbol{t}$.

$$df_{M,t}(w) = df_{M,t-1}(w) + df_{m_r,t}(w) - df_{M',t-1}(w) \tag{3}$$

Where $\boldsymbol{df_{M,t}}(\boldsymbol{w})$ denotes the number of messages in the online thread that contains word $\mathbf{w}$, $\boldsymbol{df_{m_r,t}}(\boldsymbol{w})$=1 when a new message $\boldsymbol{m_r}$ is added, and $\boldsymbol{M'}$ is the message set of faded cluster. Then, the term weighting function used for $\boldsymbol{m_r}$ could be noted as equation (4), and the normalization value $\boldsymbol{Z_t(m_r)}$ is as equation (5). Due to the collected words follow the general trend of Zipf's law[16](or Pareto distribution), an easy way to determine the uninformative terms in this heavy tail distribution is to follow their natures. As noted above, instant message might be lack of useful features due to the insufficient words in a posted text. Messages are propagated in a cascading behavior on social networking website[17], and thus we compare a new message to each existed cluster by estimating their semantic relationship and then place it to the most similarity cluster. After iterating, each identified cluster will accumulate messages until the end of its lifecycle, and this process yields an incremental vector for each cluster. It is worth mentioning that the kernel similarity function can be manipulated on demand. To compare the term vectors of $\boldsymbol{c_i}$ and $\boldsymbol{m_r}$, we use the traditional cosine measure in this preliminary scheme as shown in equation (6).

$$weight_{M,t}(m_r, w) = \frac{1}{Z_t(m_r)} tf(m_r, w) \cdot log \frac{N_{M,t}}{df_{M,t}(w)} \tag{4}$$

$$Z_t(m_r) = \sum_w tf(m_r, w) \cdot log \frac{N_{M,t}}{df_{M,t}(w)} \tag{5}$$

$$sim(c_i, m_r) = max_{i=1,\cdots,\mathcal{E}} cos\left(v_{c_i}, v_{m_r}\right) c_{in} \tag{6}$$

Where $\boldsymbol{v_{c_i}}$ represents the vector terms of $\boldsymbol{c_i}$, $\boldsymbol{v_{m_r}}$ represents the vector terms of $\boldsymbol{m_r}$, and $\boldsymbol{c_{in}}$ represents the recent number of messages in $\boldsymbol{c_i}$, and then the $\boldsymbol{m_r}$ would be absorbed into winner event cluster as:

$$\boldsymbol{c_i} = \boldsymbol{v_{c_i}} \cup \boldsymbol{v_{m_r}} \tag{7}$$

The clusters with more messages will outlive others, and we manipulate $\boldsymbol{c_{in}}$ as the factor to stretch the lifecycle out over an extended period while such events are discussed accumulatively. In this study, only the winner cluster has to be updated. In order to mine real-time messages, we apply a set of expandable message-to-message similarity matrixes Φc_i (see Fig. 1) for evaluation and considering the issue with lack of useful features among messages by calculating their distance to the cluster center as:

$$\forall j < r \;\; sim(m_j, m_r) = cos(m_j, m_r) + \left(\sqrt{(c_i - m_j)^2 - (c_i - m_r)^2}\right) \tag{8}$$

4 Experiment and Result

The goal of this preliminary experiment is to corroborate our scheme and evaluate the incidence of temporal factors. Since the number of posts might be varied during a period of time, a feasible approach is to adjust the size of time window according to the frequency of incoming messages and the event lifecycle. That is, the time window should not only be incremental but also be curtailed.

4.1 Evaluation of Performance

A total number of 2,385,350 online tweets (from 23 Aug, 2010 to 29 Aug 2010) are collected randomly through Twitter Streaming API. The specific symbols (such as "RT") and stop words (including English, Spanish, German, French and Italian) are removed in preprocessing. Furthermore, in order to reduce inflected words in short messages, we also applied the Porter stemming algorithm to remove the commoner morphological and inflexional endings. It is worth mentioning that so far there is no ideal corpus for evaluating the cluster performance of real-time messages from social networks.

We also followed the experience of [17] to set our initial power-law exponent α from -1.0 to -2.0 to rate the stability of the time decay-function. Moreover, different threshold values θ are also rated to evaluate the clustering performance. In order to evaluate the quality and effectiveness of our scheme, we first fine-tuned the power-law exponent $\boldsymbol{\alpha}$ from 1.0 to 2.0 ($\theta = 0.000001$) as shown in Fig. 2. The result corresponds with the temporal aspects of blogsphere as noted in [17, 18], and the number of clusters is convergent in a finite amount of time when $\alpha = 1.5$. Table 2 reports the length of event life-cycle which corresponding to different power and threshold factors.

Compare with the result shown in Fig. 2, the average cluster lifetime is cut down due to the time-decay function is mainly dominated by the power-law exponent $\boldsymbol{\alpha}$, and it also increases the number of clusters. Meanwhile, we might modulate both $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ according to the frequency of incoming messages and their bursty conditions. Table 2 also suggests that the window size of a cluster was adjusted automatically according to the frequency of messages $\boldsymbol{n}$ and the slowly varying parameter $\boldsymbol{\eta}$ (see Max Lifetime and Min Lifetime). This is more in tune with the information cascading behavior such as social networks.

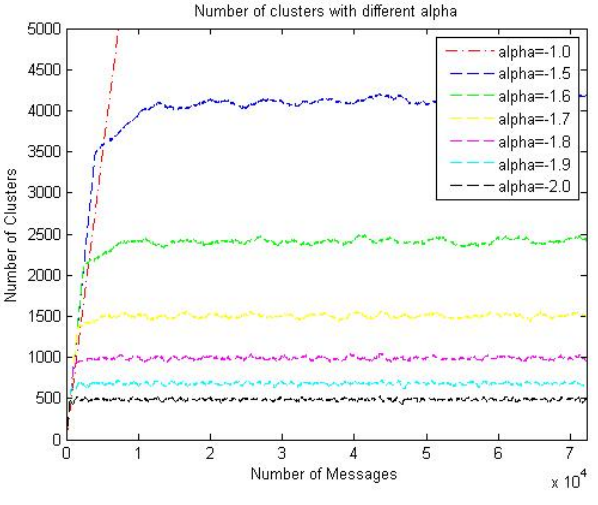


Fig. 2. Fine-turning of power-law exponent

Table 2. Evaluation of power-law exponent α and Threshold θ (time unit: Sec.)

Theta	Power	Average Lifetime	Max Lifetime	Min Lifetime	Max size of cluster
0.00005	-1.4	1380	2338	904	30
0.00005	-1.5	746	1260	423	30
0.00005	-1.6	428	897	240	30
0.00005	-1.8	174	419	61	20
0.00005	-2	85	358	1	13
0.000005	1.5	3589	3779	3363	28
0.00001	1.5	2397	3956	1682	36
0.00005	1.5	746	1260	423	30
0.0001	1.5	449	900	241	30

4.2 Evaluation of Hot Event Detection

To identify the proposed approach, we utilized the *Manila hostage crisis* event as a case study to testify our model. 400 posts, which include hostage, manila and philippiens, were collected from Twitter API. As mentioned above, the window size of online thread depends on the number of messages $\boldsymbol{n}$. Fig. 3($\boldsymbol{\alpha} = -1.5,\ \boldsymbol{\theta} = 0.01,$ cluster size ≥ 2) shows an interesting result respecting to the time-decay function which is discussed how to determinate the cluster lifecycle and erase slack messages. Posts including *hostage, manila* and *philippiens* words were bursting from Aug 23th to Aug 24th. The number of clusters also increased rapidly, and the cluster lifetime was transient. On the other hand, we saw the cluster lifetime has a lasting period due to low update rate of messages in ordinary days.

For more detail, we manipulated different power-law exponent $\boldsymbol{\alpha}$ and threshold values $\boldsymbol{\theta}$ to rate the window size for event detection. The parameters are arranged to fine-tune the scheme function as shown in Fig. 4, and the matrix illustrates the distribution of different cluster sizes. Both $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ coordinate the cluster lifecycle, and this analysis provides clues to determine the value of $\boldsymbol{\eta}$ in different message frequency. In the case of this two dimensional analysis, clusters will be faded away as quickly as the strict of threshold and power-law exponent, and each isolated message is categorized as a spot. For instance, clusters are almost tightened within three messages when $\boldsymbol{\alpha} = -2$ and $\boldsymbol{\theta} = 0.01$, but spread to fifteen clusters when relaxing the constraints. A better practice to leverage the cluster lifetime and online post frequency is to use a slowly varying parameter $\boldsymbol{\eta}$, which controls the distribution of clusters according to the number of messages in it, to adapt time-decay function.

When the frequency of online message is dispersed, the window size will be expanded for acquiring more messages. On the other hand, messages will be clustered immediately, and the cluster lifetime will be curtailed since it is appraisable. In order to facilitate real-time event detection in social networks, the function of a self-adaptive window size is critical because people send regular posts in accordance with their daily routines, however, bursts of messages were often happened at the time when they were shocked by those accidents. This phenomenon could be calculated on the cascading behavior of word distributions. Furthermore, clustered results might be mistaken if the cluster lifecycle is simply separated by a fixed or an incremental window size.

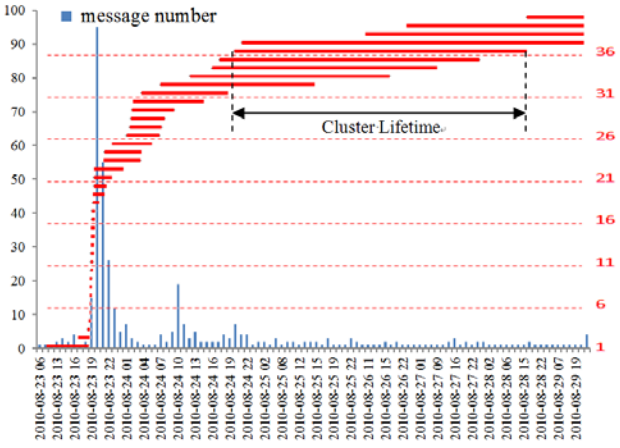


Fig. 3. Message and cluster distribution of Malila hostage crisis event

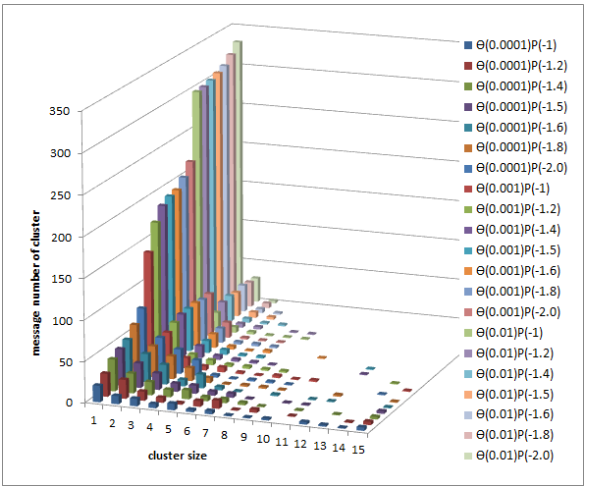


Fig. 4. Distribution matrix of different cluster size

5 Conclusion and Future Works

In this paper, we present a competition-based clustering method with a time-decay function corresponding to the natural propagation of microblogs for clustering real-time streams collected from Twitter. The proposed clustering method is used to facilitating hot event detection tasks. By following the natural cascading behavior of event lifecycle, we develop a self-adaptive TF-IDF model to manipulate real-time events. Also, we construct an expandable similarity matrix which is capable of evaluating microblogging posts with few useful semantic features. The proposed scheme provides an effective way to extract hot events and trace their evolvements.

Experimental results show that our proposed clustering method provides a feasible solution to monitor momentous real-time events and utilizing the text streams to facilitate the management of social networking data.

Some system development will be involved in our future work. First, the distributions of concurrent words will be compared between long-term and short-term variables for extracting bursty features in messages. Second, we will integrate our other developed weighting methods utilizing the bursty features of events to improve the efficiency of real-time message clustering in social networks. Third, both cluster dispersion and bursty variability will be integrated as ranking factors to increase feasibility of the developed event mining method.

References

1. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proceedings of the 19th International Conference on World Wide Web, pp. 851–860. ACM, Raleigh (2010)
2. Heverin, T., Zach, L.: Microblogging for Crisis Communication: Examination of Twitter Use in Response to a, Violent Crisis in Seattle-Tacoma, Washington Area. In: Conference Microblogging for Crisis Communication: Examination of Twitter Use in Response to a 2009 Violent Crisis in Seattle-Tacoma, Washington Area (2009)
3. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, pp. 56–65. ACM, San Jose (2007)
4. Banerjee, N., Chakraborty, D., Dasgupta, K., Mittal, S., Joshi, A., Nagar, S., Rai, A., Madan, S.: User interests in social media sites: an exploration with micro-blogs. In: Proceeding of the 18th ACM Conference on Information and Knowledge Management, pp. 1823–1826. ACM, Hong Kong (2009)
5. Cheong, M., Lee, V.: Integrating web-based intelligence retrieval and decision-making from the twitter trends knowledge base. In: Proceeding of the 2nd ACM Workshop on Social Web Search and Mining, pp. 1–8. ACM, Hong Kong (2009)
6. Honeycutt, C., Herring, S.C.: Beyond Microblogging: Conversation and Collaboration via Twitter. In: Proceedings of the 42nd Hawaii International Conference on System Sciences, pp. 1–10. IEEE Computer Society, Los Alamitos (2009)
7. Bakshy, E., Karrer, B., Adamic, L.A.: Social influence and the diffusion of user-created content. In: Proceedings of the Tenth ACM Conference on Electronic Commerce, pp. 325–334. ACM, Stanford (2009)
8. Weng, J., Lim, E.-P., Jiang, J., He, Q.: TwitterRank: finding topic-sensitive influential twitterers. In: Proceedings of the Third ACM International Conference on Web Search and Data Mining, pp. 261–270. ACM, New York (2010)
9. Gong, L., Zeng, J., Zhang, S.: Text stream clustering algorithm based on adaptive feature selection. Expert Systems with Applications 38, 1393–1399 (2011)
10. Nallapati, R., Feng, A., Peng, F., Allan, J.: Event threading within news topics. In: Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management, pp. 446–453. ACM, Washington, D.C (2004)
11. Yang, Y., Pierce, T., Carbonell, J.: A study of retrospective and on-line event detection. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 28–36. ACM, Melbourne (1998)

12. Yang, Y., Carbonell, J.G., Brown, R.D., Pierce, T., Archibald, B.T., Liu, X.: Learning approaches for detecting and tracking news events. IEEE Intelligent Systems and their Applications 14, 32–43 (1999)
13. Brants, T., Chen, F., Farahat, A.: A System for new event detection. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, pp. 330–337. ACM, Toronto (2003)
14. Swan, R., Allan, J.: Automatic generation of overview timelines. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 49–56. ACM, Athens (2000)
15. Joel, W.R., Yu, J., Thomas, E.P., Brian, A.K., Mark, T.E., Ali, R.H.: TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams. In: Conference TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams, pp. 258–263
16. Li, W.: Random texts exhibit Zipf's-law-like word frequency distribution. IEEE Transactions on Information Theory 38, 1842–1845 (1992)
17. Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N., Hurst, M.: Cascading Behavior in Large Blog Graphs. In: SIAM International Conference on Data Mining, SDM (2007)
18. Barabasi, A.-L.: The origin of bursts and heavy tails in human dynamics. Nature 435, 207–211 (2005)

Weighted Co-authorship Network Based on Forgetting

Milos Kudelka, Zdenek Horak, Vaclav Snasel, and Ajith Abraham

VSB-TU Ostrava, 17. listopadu 15,
77900 Ostrava-Poruba, Czech Republic
{milos.kudelka,zdenek.horak,vaclav.snasel}@vsb.cz,
ajith.abraham@ieee.org

Abstract. In this paper we focus on the analysis of weighted networks and their properties. We describe a new way to weigh network vertices and edges based on the Forgetting curve. We denote the weight as a stability changing gradually over time. Based on the stability, we propose new measures. For our experiments we have selected the DBLP database, therefore we can evaluate our approach on a real network with more than 830,000 vertices.

Keywords: weighted network, forgetting curve, measure, experiment, DBLP.

1 Introduction

Publication activities make up a complex and steadily growing social network. In such a network, the evolution and dynamics contains much hidden information. In this paper we focus on the analysis of ties in a co-authorship network. The computer science bibliography database (DBLP) is a good example of a large dynamic network that contains a lot of implicit information. Various authors work together through time; their cooperation can be suspended for some time and restored later. Therefore these relations have different weights over time as well as particular authors possess a different weight (from the viewpoint of their surroundings).

In our paper we describe a novel approach for weighing the vertices and edges of a social network (see also [10]). This approach is based on the so called Forgetting curve. We propose several measures which can describe the authors' behavior over time. For an experiment proving our hypothesis we have selected a large network of authors. The important aspect of our approach is its low computational complexity. Thanks to this, we can present our own solution serving as an experimental sandbox. This online solution delivers all the results presented in this paper.

The next section of this paper discusses approaches for weighted network analysis related to our work. The third section formalizes the notion of the forgetting curve, while the fourth section adapts this notion for weighted networks. In the fifth section we discuss specific measures based on stability.

2 Related Work

The analysis of general complex networks is well-described in [3] and [4]. Liu et al. in [11] provides a good overview of Social Network Analysis, co-authorship networks

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 72–79, 2011.

and a combination of both. They also compared the results of the analysis using classical SNA and PageRank and its modification AuthorRank, respectively. Further coefficients can be found; e.g. in Newman [12]. Hart [8] provided an interesting survey on co-authorship – its models, grounds, etc. Han et al. [7] introduced the concept of supportiveness, which captures co-authorship ties in a non-symmetric way.

Huang and Huang [9] addressed two main problems of most visualization techniques - the problematic application in large-scale networks and the difficulty to incorporate historical data in one artifact. Elmacioglu and Lee [6] presented statistics calculated from the DBLP dataset along with a comparison of weighted and unweighted variants of SNA coefficients used to identify important authors.

Opsahl et al. [13] discusses two aspects of weighted networks – the number of ties versus their strength and presents an approach combining these two aspects. Some of our proposed measures benefit from this combination too. The architecture of weighted networks is investigated in [2] along with several measures constructed as an extension of classic network analysis coefficients for weighted networks.

Barabasi et al. [1] is focused on the evolution of the social network of co-authorship. Co-authorship can be also considered as a suitable area for link prediction (see for example [14]). A more general approach to the time aspect of social networks is discussed in [15]. The presented Noise Tolerant Time-varying Factor Graph Model is based on three intuitions: (i) users' actions at some particular time are influenced by their friends historic actions, (ii) users' actions are usually based on their previous actions; (iii) users' actions at the same time have a (strong) correlation. This method simultaneously models the social network structure and predicts social actions of users. This paper also contains a brief survey of dynamic social networks analysis.

We address similar issues as mentioned works, but our approach is strongly based on historical data and in a certain way complies with the presented intuitions, but differs in using the forgetting curve. The motivation is that we believe that this approach complies with user expectations and human behaviour. There are some basic principles arising from the classic network analysis which play an important role in the weighted network analysis and also in our work (such as the analysis of connected triplets). Note that many of the mentioned papers use networks of researchers for their evaluation.

3 Forgetting Curve

The human brain stores information which is fixed in memory by its frequent usage, but which can – when not used – also fade from the memory. This process is very complex. However many experiments have already been done (see e.g. [5]), which lead to a fairly exact description of functions involved in the memorization and forgetting of information. We review the social network as a human brain which learns and forgets. The reason is that the vertices of the network are people having these functions in their brains.

In the following we understand the term social network as an undirected weighted graph. During the calculations of edge and vertex weights, we use time-dependent information and values related to forgetting.

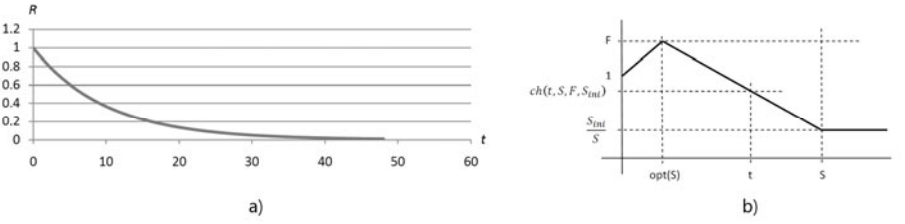


Fig. 1. a) Forgetting curve, b) calculation of $ch(t, S, F, S_{ini})$ in time t

The forgetting curve (see Figure 1a) defines the probability that a person can recall information at time t since previous recall. It can describe long-term memory and it is usually expressed using the following equation:

$$R = e^{-\frac{t}{S}}, \tag{1}$$

where R (memory retention) is the probability of recalling information at time t since the last recall, e is the Euler number, t is the time since the last recall and S (relative strength of memory; stability) is the approximated time since the last recall for which the information is stored in memory.

The computation depends on the type of memory, especially on the estimated time S (this value is not constant in the long term). For simplicity, assume that if we work with information for the first time, then the time of storing information in memory is $S_{ini} > 0$ and this default value is constant.

An important feature of long-term memory is that after information recall at the time $t > 0$, the time of storing information in memory S changes. The change is dependent on the previous time S and on the time t of recall. Ideally, the reproduced recall multiplies this time (in comparison with the previous value) by factor $F > 1$.

The other important feature of the long-term memory is that immediate (too early) reproduced recall of information has no big effect on the learning. On the other hand, the late reproduced recall (in time near S) causes substantial forgetting. There is an optimal time between these two extreme situations in which the reproduced information recall causes a high level of remembering (and consequently the maximum increase of time S by factor F). In the ideal case (reproducing the information in optimal time), the remembering of information is gradual and very effective – after each recall, the time of storing information in memory S (remembering) is multiplied by factor F. Therefore for the calculation of S_{new} we have to consider several things which we discuss in detail in [10]. For the purpose of this paper we can simplify the process to the equation $S_{new} = ch(t, S, F, S_{ini}) \cdot S$, where one choice of the ch function (based on linear functions) is illustrated on Figure 1b.

4 Weighted Network Based on Stability

Interactions between particular pairs of vertices continuously take place in the social network. If we consider these interactions as an experience stored in memory, then the

ties between the vertices of the network are more stable if this network "learns" these interactions. As a result we assume that the more interactions occur between the two vertices, the more stable the vertices and the tie between them are. Therefore we can understand the social network as a set of variously stable vertices and ties.

The properties of vertices and ties change over time, depending on how often and in what time two vertices interact. For the calculation of the properties of ties we use the forgetting curve. It is an analogy to the learning and forgetting of reproduced information. For each vertex and tie we define two time-changing characteristics:

- *Edge stability ES* is the estimated time for which the tie remains active (since given time *t*), while *Active edge* is a tie, for which holds that $ES > 0$ in time *t*.
- *Vertex stability VS* is the estimated time for which the vertex remains active (since time *t*), while *Active vertex* is a vertex, having $VS > 0$ in time *t*.

4.1 Experimental Dataset

On December 12, 2010, we downloaded the DBLP dataset in XML[1] and preprocessed it for further usage. First of all, we selected journal volumes and conferences held by IEEE, ACM and Springer (55,367 journal records and 13,052 conference records). For every record we identified the month and year of the publication. In the next step we extracted all authors having at least one published paper (839,180 authors). Using this information about authors and their papers, we created a set of cooperation between these authors consisting of 4,603,474 items. The cooperation is understood to be the co-authorship of one paper. Using the information about the conference date, we accompanied this cooperation by time information. We also ignored the ordering of author names as it is impossible to investigate the particular ordering protocol (by alphabet, by contribution, etc.) and hence all co-authors are given equal importance.

We have truncated the selected time period to June 2010 to get the most complete dataset. Then we divided the entire recorded publication period of conferences (the first record from 1963) into one-month time periods. If during one month an author has published a paper with another co-author in at least one conference or journal, then we set one *interaction* for both authors (vertices) and the tie between author and his/her co-author (edge) for this month (as an interaction we understand a publication activity in one month). For each vertex and edge we obtained a list of months in which the interactions occurred (if a particular edge has many interactions, then its adjacent vertices have at least the same number of interactions).

Remark 1. If an author published in one month without any co-authors, the interaction is counted separately (each vertex has one self-edge for this kind of publications).

Then we applied the forgetting curve to compute the stability of every author and tie in a specified month. To set the parameters we have used the following values: $S_{ini} = 12$, $F = 1.2$. Interpretation: if an author publishes for the first time, then he will be active in the network for the next 12 months. If he publishes again in the optimal time, his stability is multiplied by 1.2. These values have been chosen on the basis of our experiments as highly corresponding to the expert expectations. Small change of the values will not cause major changes to the network.

[1] Available from `http://dblp.uni-trier.de/xml/`

As an example, we selected an author who has the most records in DBLP - Philip S. Yu. Figure 3 (left) shows the evolution of stability of this author from 1985. As evident, thanks to his high and regular publication activity, the stability grew almost continuously. In Figure 3 (right) can be seen the evolution of stability of the edge with co-author Kung-Lung Wu (the most stable edge of Philip S. Yu to June 2010).

At the end of the time interval only 200,626 authors and 409,797 ties were active (the others had stability equal to zero – from more than 830,000 authors and 2,500,000 ties). The first ten authors and co-authorship ties according to the stability in June 2010 are shown in Table 1. The stability does not depend only on the number of interactions (and the number of publications consequently) but also depends on how often and how regularly these interactions occur.

Remark 2. The stability calculation for each vertex and edge has linear time complexity with the number of interactions.

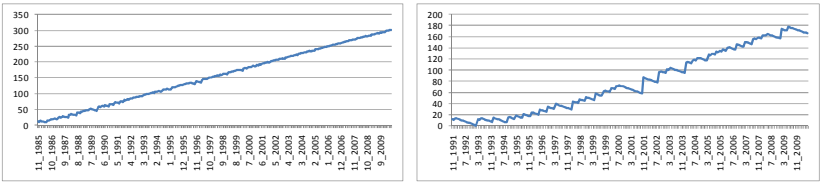


Fig. 2. Stability evolution of Philip S. Yu and between Philip S. Yu and Kung-Lung Wu

Table 1. Top 10 vertices and edges by stability

#	AUTHOR(vertex)	CO-AUTHORS (edge)
1	Jeffrey D. Ullman	Andrzej Ehrenfeucht - Grzegorz Rozenberg
2	Edward G. Coffman Jr.	Didier Dubois - Henri Prade
3	John E. Hopcroft	Divyakant Agrawal - Amr El Abbadi
4	Grzegorz Rozenberg	Irith Pomeranz - Sudhakar M. Reddy
5	Robert Endre Tarjan	Pankaj K. Agarwal - Micha Sharir
6	Oscar H. Ibarra	Evelina Lamma - Paola Mello
7	Arnold L. Rosenberg	Evangelos Kranakis - Danny Krizanc
8	C. A. R. Hoare	P. M. Melliar-Smith - Louise E. Moser
9	Erol Gelenbe	Richard C. Wilson - Edwin R. Hancock
10	Kurt Mehlhorn	Enrico Macii - Massimo Poncino

5 Measures Based on Stability

Stability in the context of the entire network shows the importance of a particular author or relationship between authors at a particular time. We proposed three new measures related to the view of an author in the context of his/her surroundings.

5.1 Cooperativeness

The basic motivation is to describe the vertex in relation to important (stable) vertices in its surroundings. The current stability of the vertex, which may be rather high, is

not solely dependent on the number of ties or the strength of these ties (e.g. authors with high stability may have no co-authors). As an important tie for its surroundings we consider a tie having an adjacent vertex with high stability. The cooperativeness also takes into account the stability of the relation between these vertices:

$$Cooperativeness(v) = \sum_i \sqrt{ES(e_i) \cdot VS(v_i)}\,, \tag{2}$$

where e_i and v_i are edges and vertices adjacent to the vertex v. The first ten authors according to the cooperativeness in June 2010 are shown in Table 2.

5.2 Solidity

The aim of the solidity is to give priority to strong ties over the weak ones. As a consequence, the solidity describes vertices whose ties to their surroundings are strong. Solidity takes into account only ties having at least some minimal stability:

$$Solidity(v, stab) = \sum_i (ES(e_i) - stab)\,, \tag{3}$$

where e_i are edges adjacent to the vertex v, for which holds $ES(e_i) > stab$. In our experiment we have used $stab = 12$ (as a default value of stability S_{ini}). The first ten authors according to the solidity in June 2010 are shown in Table 3.

5.3 Dependency

An interesting feature of co-authors may be their relationship in the stronger/weaker terminology. This is a measurement of a relationship with another author in the context of the group of his/her co-authors. This value indicates how stable (w.r.t. to the stability) the ties between the vertex and the neighbors of the second vertex are (incl. the tie between the inspected vertices). This is a non-symmetric measure; at least one common neighbor is the precondition for nonzero value.

$$Dependency(v_1, v_2) = \frac{\sum_{e_i \in CE(v_1, v_2)} ES(e_i) \cdot DepRatio(e_i)}{\sum_{e_i \in E(v_1)} ES(e_i)}, \tag{4}$$

$$DepRatio(e_i) = \frac{ES(v_1, v_2)}{ES(e_i) + ES(v_1, v_2)}, \tag{5}$$

where $CE(v_1, v_2)$ is a set of edges between vertices v_1, v_2 and their common neighbors, $DepRatio(e_i)$ is a relative dependency of the edge e_i on the vertex v_2 (on the edge v_1, v_2) and $E(v_1)$ is a set of all edges adjacent to the vertex v_1. As an example, we present the evolution of the tie of aforementioned authors (see Figure 4).

Table 2. Top 10 vertices by cooperativeness

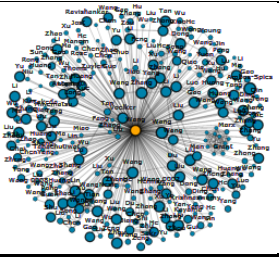
#	AUTHOR(vertex)	BEST AUTHOR NETWORK
1	Wei Wang	
2	Lei Wang	
3	Wei Chen	
4	Philip S. Yu	
5	Li Zhang	
6	Jiawei Han	
7	Wei Zhang	
8	Hai Jin	
9	Yan Zhang	
10	Luca Benini	

Table 3. Top 10 vertices by solidity

#	AUTHOR(vertex)	BEST AUTHOR NETWORK
1	Mahmut T. Kandemir	
2	Luca Benini	
3	Philip S. Yu	
4	Floriana Esposito	
5	Francky Catthoor	
6	Mario Piattini	
7	Hai Jin	
8	Jiawei Han	
9	Leonard Barolli	
10	Enrico Macii	

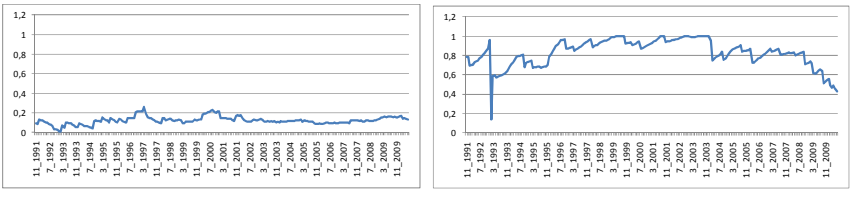


Fig. 3. Dependency Philip S. Yu on Kung-Lung Wu and vice versa

6 Conclusion

In our paper we have described a new way to weigh vertices and edges in the social network. The application of our approach to the co-authorship network shows results consistent with expert expectations. In this case, the stability based on the forgetting curve indicates the strength of social connection between vertices and the social strength of particular vertices. Using stability, we proposed measures capturing the relation between vertex and its surroundings in detail.

Our approach is proving to be easily applicable even for large networks. Results presented in this paper are verifiable in our online system http://www.forcoa.net, which provides an interface to the complete visualized results over the dataset described in this paper. Due to the limited extent of this paper, we have not addressed the direct comparison with existing approaches and analysis of properties of the

whole stability-based network. We will address these issues in our future work as well as prepare a larger author-oriented study.

Acknowledgment. This work is partially supported by SGS, VSB-Technical University of Ostrava, Czech Republic, under the grants No. SP2011/172 and No. GA 102/09/1494.

References

1. Barabasi, A.L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. Physica A: Statistical Mechanics and its Applications 311, 590–614 (2002)
2. Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted network. Proceedings of the National Academy of Sciences of the United States of America 101, 3747–3752 (2004)
3. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: Structure and dynamics. Physics Reports 424, 175–308 (2006)
4. Costa, L.F., Rodrigues, F.A., Travieso, G., Boas, P.R.V.: Characterization of complex networks: A survey of measurements. Advances in Physics 56, 167–242 (2007)
5. Ebbinghaus, H., Ruger, H.A., Bussenius, C.E.: Memory: A contribution to experimental psychology (1885/1913)
6. Elmacioglu, E., Lee, D.: On six degrees of separation in DBLP-DB and more. ACM SIGMOD Record 34, 33–40 (2005)
7. Han, Y., Zhou, B., Pei, J., Jia, Y.: Understanding Importance of Collaborations in Coauthorship Networks. In: SIAM Int. Conference on Data Mining, pp. 1112–1123 (2009)
8. Hart, R.L.: Co-authorship in the academic library literature: a survey of attitudes and behaviors. The Journal of Academic Librarianship 26, 339–345 (2000)
9. Huang, T.H., Huang, M.L.: Analysis and Visualization of Co-authorship Networks for Understanding Academic Collaboration and Knowledge Domain of Individual Researchers. In: Int. Conference on Computer Graphics, Imaging and Visualization, pp. 18–23 (2006)
10. Kudelka, M., Horak, Z., Snasel, V., Abraham, A.: Social Network Reduction Based on Stability. In: International Conference on Computational Aspects of Social Networks (2010)
11. Liu, X., Bollen, J., Nelson, M.L., Van de Sompel, H.: Co-authorship networks in the digital library research community. Information Processing & Management 41 (2005)
12. Newman, M.: Who is the best connected scientist? A study of scientific coauthorship networks. Complex networks, 337–370 (2004)
13. Opsahl, T., Agneessens, F., Skvoretz, J.: Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32, 245–251 (2010)
14. Pavlov, M., Ichise, R.: Finding Experts by Link Prediction in Co-authorship Networks. In: 2nd Internation ExpertFinder Workshop, pp. 42–55 (2007)
15. Tan, C., Tang, J., Sun, J., Lin, Q., Wang, F.: Social action tracking via noise tolerant time-varying factor graphs. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1049–1058 (2010)

Towards Usage-Centered Design Patterns for Social Networking Systems

Dina Hussein, Ghada Alaa, and Ahmed Hamad

Faculty of Informatics & Computer Science, The British University in Egypt (BUE)
Cairo - Suez Desert Road, El Sherouk City, 11837 - P.O. Box 43, Egypt
{dina.hussein,ghada.alaa,ahmed.hamad}@bue.edu.eg

Abstract. The role of social networks is becoming increasingly relevant in the recent years. While social networks are becoming more ubiquitous, a profound approach guiding the analysis, design and development of such social systems is becoming increasingly important. This research focuses on deriving functionalities that classify behavior of social networking applications. This led to identifying three basic/elementary usage patterns which include; *inter-connectivity*, *file sharing* and *tagging*, while the intersection between these patterns led to other secondary/supportive usage patterns which include *content authoring*, *content remixing*, *content aggregation & syndication* and *content streaming*. By interlinking the proposed patterns of usage with underlying techniques and Web 2.0 tools we aim to provide insight into re-usable elementary building blocks common to any type of social networks. As a result a usage-centered roadmap is presented that identifies the intersection of several Web 2.0 tools and techniques used to realize the three identified social networking patterns of usage.

1 Introduction

The rapid evolution of Internet technologies have witnessed the emergence of new Web elements, such as blogs, wikis, social networking, social bookmarking, and other related applications referred to as Web 2.0. While Web 1.0 paradigm has been related to passive, just receptive users, Web 2.0 relies mainly on rich user participation through individual authorization or co-creation of content, and content publishing/ sharing that would shift traditional passive content toward a user-generated one [8]. Social Networks in this sense could be considered as a cornerstone Web 2.0 application [1]. Social Networks refers to online community websites that inter connect users and enhance sharing of common interests and activities [2]. Since its proliferation, Social Network (SN) features have evolved in an interesting way. It started from communication through instant messaging, chatting and group mailing. Then expanded further to include defining circle of trusted friends that cover friend of a friend (FOAF) concept, as well as sharing resources of textual and other multimedia formats. Recently, SN covered collaborative commenting, tagging and rating features [3], these content recommendation services are based on user-generated tags and user interests extracted from their profiles. Thus, an essential element for realizing SNs

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 80–89, 2011.

covers "user profiles" [4, 5], where data about members are collected and stored for further sharing among acquaintances and to be processed to gather specific interests and preferences of users. User profiles enable dissemination of users' information either publicly to all members of the SN or keep it private to the circle of trusted friends. SN applications support both textual and pictorial information about users, whereas an increased number of SNs is supporting video and/or audio profiles (6).

Hence SNs bring together major concepts of Web 2.0; they facilitate rich inter-connectivity among participants, enhance file sharing and active participation of members in a usable way to include even those with poor technical expertise. A long list of SN applications have been introduced to cope with the continuous drops in personal computer prices, increases in computers' capacity and speed, the greater and cheaper broadband networks access and the improvements in techniques that were introduced by Web 2.0.

Social networking is an increasing hot topic in software analysis and design. Applications of social networking are built around communities of common interest, professional communities of expertise and geographical proximity networks [20]. Examples of social networking communities are LinkdIn, Friendester and Facebook where such sites are growing in population, taking advantage of rich user profiles to create links among members.

Current research in the area of social computing is aiming to establish theories and set principles related to social network behavior and SN evolution dynamics. This should help understand the phenomenon of SN building and their implications on analysis, design, implementation and maintenance of SNs. This would further guide SN future trends and evolution of SN applications that prove the longevity of such systems.

Patterns on the other hand represent a major concept in software design and modeling, they aim to discover re-usable building blocks to be taken under consideration [21]. They act as guide for software architects, to help them identify recurrence and reduce redundancy in software design. Two common types of patterns are widely known; architectural patterns, that offer a structural scheme for software systems, and design patterns that offer behavioral/functional scheme for software systems [22].

In this work we focus on tailoring design patterns that is based on SNs common usage/functionality derived from literature to reach taxonomy of SNs usage patterns.

The paper is organized as follows. In section 2 we start with presenting taxonomy that define Web 2.0 design patterns; we explain how the suggested design patterns build on and are affected by WEB 2.0 architecture, technology and techniques in use. In section 3 we surveyed the literature for techniques used to build social networking platforms. In section 4 a usage-centered classification of SNs is presented along with interlinking each functionality/usage with underlying SN techniques and architecture. Section 5 concludes our work.

2 Web 2.0 Patterns of Usage

In a previous research, the authors have derived seven usage patterns for Web 2.0 applications which are; Inter-connectivity, Content authoring, Content tagging &

rating, Content aggregation & syndication, Content remixing, Content streaming and File sharing. This classification has arisen from surveying multiple attributes related to functionalities of Web 2.0 tools and applications [8]. We can summarize the different Web 2.0 patterns of usage with some examples in Table 1.

Table 1. Web 2.0 patterns of usage, adobted from [8]

Design Pattern	Description	References
Interconnectivity	Connecting people and enhancing interactions and participation among them are major goals of Web 2.0. Blogs let Internet users to participate through posting their comments. Similarly, social networks allow rich inter-connectivity between participants, such as chat, posting notes and status updates, sharing data, and utilizing user profiles to capture members' interests.	[1, 20, 27, 28, 29]
Content authoring	This Web 2.0 design pattern refers to authorization to co-create and publish content. Blogs for example give the blogger individual authority to generate his/her own post, whereas wikis and Office 2.0 enable group authorizations; i.e. several participants can edit each others' work.	[20, 27, 28, 29]
Content tagging	A challenge appeared with the rise of Web 2.0 applications that cover mainly user generated content, is how to classify content? The traditional taxonomy is becoming more difficult to apply with the huge diverse background of participating users. This gave rise to a classification technique called folksonomy; in which the responsibility of classifying content is given to the community along with the responsibility to create and maintain that content. The community decides where this content belongs to by the use of "tags". Tagging enhances social search as applied in social networking websites. Furthermore, tagging can rely on so-called content cloud which is a technique that visually represents frequently used tags.	[1, 20, 26, 27, 28, 29]
File sharing	With the advancements in P2P (peer-peer) networks and the relatively low cost of digital media devices, such as digital cameras several Internet users can now upload and share media files over the Internet. P2P networks fasten the exchange of files by opening direct sessions between the users without the need to stream content or files through the server, like in social networking chat rooms, file sharing services websites, Skype, online games, etc.	[2, 28, 29]
Content aggregation & syndication	While having continuously added/updated content over the Internet it is important for users who are interested in multiple, scattered sources of information to keep up with such updates. An aggregator or feed reader can be used to centralize all the recent changes in the sources of interest, and a user can easily visit the reader/aggregator to view recent additions and updates. The feed aggregator can syndicate several Web content, such as news headlines, blogs, file sharing resources, etc. This can be also coupled with mashups; re-mixing aggregated content like in Yahoo Pipes for example.	[5, 20, 27, 28, 29, 30]
Content remixing	This involves mashup-based services that refer to Web applications combining data from more than one source into a single integrated tool. This is implemented by accessing open APIs and available data sources to produce another meaningful application. Paul Rademacher's housingmaps.com for instance combines Google Maps with Craigslist apartment rental and home purchase data sources to create a richer housing search tool. Similarly, members of social networks could customize their own pages by adding Applications to further extend the functionalities typically offered by the social networks, Facebook Apps is an example of content remixing in social networks.	[28, 29]
Content streaming	Audio or video files can be transmitted as live or archived files. Streaming techniques imply content to be streamed directly to users, not downloaded from or a single server computer, thus it uses P2P networks. This makes access feasible for slower connections. Applications that use streaming include webcasts and virtual reality games. Several social networks offer streaming of audio files and creating personal channels and building libraries of favorite tracks, e.g. YouTube.	[5, 28]

The significance of each suggested pattern may vary according to the type of Web 2.0 application under consideration. For example for an e-learning system we found out that they depend more on inter-connectivity, content authoring, file sharing and content streaming. On the other hand, in e-commerce applications, we found out that inter-connectivity, content aggregation and content tagging & rating, in form of comments and reviews are the key patterns [8].

By applying the proposed Web 2.0 design patterns on a real life example, eBay (http://www.ebay.com/) we found out that the pattern of *tagging* is missing, we argued that applying *tagging* on such example, would enhance searching for a certain product, as a more reflective and informative tags might be added by users who had already used a product. Proposed tags may also be further extended to include pattern matching [20]. For example, recommending some links that match or extend the user's purchased items. Mashups (re-mixing of heterogeneous applications and widgets by building on open APIs and service-oriented architecture concept) may also add value to the website [8].

In the following sections, we'll be applying the Web 2.0 usage patterns on SNs to depict the key patterns that would act as a backbone for SNs applications and other supportive patterns.

3 Survey of Techniques Underlying SNs

Several contributions tried to investigate the set of functionalities characterizing various social network applications that build on concepts of Web 2.0. The term of "Social Web" is presented in several literatures to denote "Connectivity" with the aim to enhance social interactions through the use of Wikis, Blogs and Podcasts [5, 4]. This indicates the value of "Collaboration" that would further encourage the activity of sharing knowledge, experience and resources. "Flexibility" is another feature that is realized in Social Web through reusability of open APIs and remixing of Web content with flexible content classifications known as "Folksonomies" [19]. With the use of "Folksonomy" a quicker and more relevant access to specific topics that rely on "Groups of Interests" enabled and created through SNs becomes possible.

Thus, "Architecture of Participation" is a very important feature of social networks in general as they enable information to be liberated from the traditional control enforced by content creators. This encourages "Collective Intelligence", where analyzing reviews and tags written, edited, and published by Internet users to obtain a collective pool of knowledge[19]. In this sense folksonomies and tagging could be considered as user generated ontology.

Authors in [5] indicated that social networking provide social tools that collect information about members known as *"User Profiles. Also,* SNs connect members together through a practice of *"Sharing"* information of interest. Members can *"Inter-link"* their personal pages together, where *"Searching"*, *"Locating"* and *"Recommending"* are important features. The creation of *"Communities of Practice"* is offered by such SNs through *"User-Created Groups"* that discuss and share information about a certain topic of interest. *"Announcements"* could be also delivered to members to highlight a certain update or event. Authors in [5] argue that SNs could be created by bundling several features of Web 2.0 and social system

applications, e.g. blogs, instant messaging, collaborative filtering, social book marking, folksonomies, social search, file sharing and podcasts, where the later are media files shared over the internet and made available to syndicated download through RSS feeds.

Similarly, [4] signifies *"User Profiles"* as a major feature in SNs where data about users could be stored and used for further recommendation of content and matching with similar users. *"Personalization"* is another feature, where users can create their own Web pages and create user-generated content; examples are blogs, wikis, photos and multimedia sharing. Additionally *"Links"* exist to interlink Web pages of SN members with other friends who share common interest. "*Search*" is also paramount to benefit from such Web page inter-linking to find vast amount of information that exist within the network of friends. Search facility helps expanding the network of acquaintances by finding other friends with common interest. The authors also present SNs as applications that enable the creation of *"Communities of Practice"* that are based on discussion forums and chat sessions between members and to share knowledge by means of blogs and wikis to create and leverage "social interactions*"*, *"content recommendation"*, *"user-centered classifications"* and "*monitoring"* of content.

Major features of SNs are provided in [9]; users represent members of SNs who register firstly by providing personal information known as "*User Profiles*". *"Links"* exist to connect members of SNs and enabling such "*Links*" to public would facilitate browsing and locating friends with common interests. *"Groups"* is the last and third feature presented by the authors in [9] to notify "*Groups of Interest*". These could be created among members to enable them to discuss, send messages, notifications or simply share information and various multimedia files among them. Furthermore, the authors in [6] elaborate that one of the major features of SNs is *"User Profiles"* where members can provide information about themselves that could be used to define a "Circle of Friends".

Major features of SNs are provided in [9]; users represent members of SNs who register firstly by providing personal information known as "*User Profiles*". *"Links"* exist to connect members of SNs and enabling such "*Links*" to public would facilitate browsing and locating friends with common interests. *"Groups"* is the last and third feature presented by the authors in [9] to notify "*Groups of Interest*". These could be created among members to enable them to discuss, send messages, notifications or simply share information and various multimedia files among them. Furthermore, the authors in [6] elaborate that one of the major features of SNs is *"User Profiles"* where members can provide information about themselves that could be used to define a "*Circle of Friends*".

In [11] authors were concerned to provide a classification scheme for social networking portals to include *"User Profiles"* that enable personalized interactions between network subscribers based on user preferences. These can help join together people from certain geographic areas, or simply those who share common interests. *"Share"* resources is also an important feature of SNs, along with *"Participation"* allowing members to jointly create content rather than stick to pre-defined content offered by traditional Web. *"Semantic Search "*is also mentioned as a feature that would rather enhance conventional search approaches. This makes use of simple links or matching techniques for a set of keywords enhanced by the use of "Ontologies*"* where domain-specific vocabulary and definitions are utilized to enhance query results.

Thus, we can conclude that SNs utilize several tools to cover instant messaging, blogs, wikis, chat forums, podcasts, etc., as well as techniques to cover notifications, content tagging content, reviews, comments and finally content recommendation based on users' contributions and reviews. In Table 2 we provide a summary of major SNs techniques.

Table 2. Social networks features

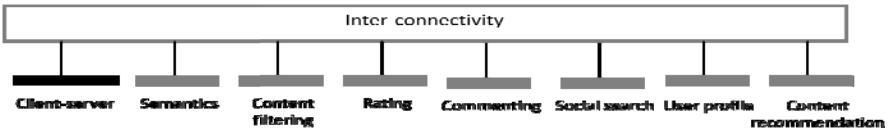
Technique	Description
User profiles	SNs in its simplest definition is a platform that enable members to easily create "profiles" where information about themselves is provided as well as their interests, a member can then identify a circle of his friends [6]. Through the information provided in each member's profile "Links" could be created to establish "Connections" among several personal pages and to expand the circle of friends based on the information provided in the user's profile [5].
Content recommendation	Web sites that provide information or selling products or offering services use recommendation systems to suggest relevant content to its users. Amazon (http://www.amazon.com/), for example, uses collaborative filtering to recommend new books or movies to its users. Recommendation systems would basically find users with common interests, and suggest new products that were liked by other users with similar opinions [13].
Social search	While active participation is highly encouraged in all SNs, through sharing various resources and adding one's own tags, this enhances searching for a particular subject among a collective list of resources shared by others in a SN who shares the same search interest. Sharing content in such a way and making it searchable would leverage communities of interest and expertise [4]. Searching for news and updates via live feeds posted by members of SNs who are witnessing certain events is another powerful tool offered by SNs, Twitter (http://www.twitter.com/) and Youtube (http://www.youtube.com/) have been widely searched for live updates and videos for the demonstrations that took place all around Egypt in Jan 2011.
Aggregation & syndication	Growth of SNs has given rise to tools that notify users on updates with a chosen social network platform [5]. An aggregator or feed reader can be used to centralize all the recent changes in the sources of interest, and a user can easily visit the reader/aggregator to view recent additions and changes [15]. The feed aggregator can syndicate several Web content, such as news headlines, blogs, multimedia resources, etc. Examples are iGoogle (http://www.google.com/ig) where a user can add news, photos, weather, and stuff from across the web to his/her own personalized page.
Commenting	In order to leverage social interaction, multiple participants can add and share their comments. Collaborative commenting on related blogs for example and the support for feedback allows a group to rate the contributions of others and encourage active participation [10].
Rating	Content rating is a feature that enables users to rate and provide reviews about different products and sellers. For example in eBay (http://www.ebay.com/) a constant need to leverage online trust and help sellers establish a profound reputation that will lead to customer loyalty.
Semantics	Web semantics technologies is a major feature in SNs, as web intelligence is an important part of SNs applications, these are included in search engines, recommendation of a content based on a user profile and personalization of members pages based on his favorite content online.
Content authoring	This feature of SNs encourages co-creation, self classification, tagging, of a content and then publish it over the Internet. Blogs for example give the blog creator individual authority to authorize a post, tag it and publish it online [1].
Content filtering	While SNs relies mainly on reciprocal trust built and encouraged within participants and interest groups, some deceitful conduct might though exist, giving the fact that a online communities are becoming more ubiquitous with information of every kind, there is an increased demand to assure the quality of content which have spurred the demand to filtering system. Digg (http://www.digg.com/) for example, accepts submissions of stories from users who consider worthy of attention. Users can then vote for, or 'dig and bury' that reflects a good example of community based filtering in action [13].

4 Towards Usage-Centered Design Patterns of Social Networks

After surveying SNs tools and techniques derived from literature in previous section, we are now moving toward a higher level of abstraction to form SN design patterns.

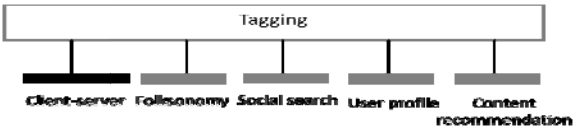
By applying Web 2.0 usage patterns on SNs, basic usage patterns has been derived which include; *inter-connectivity*, *tagging* and *file sharing*. In Fig 1.1, 1.2, 1.3 basic SNs design patterns are represented along with their related underlying techniques.

a- Inter-connectivity



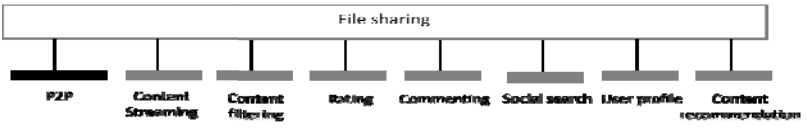
While interconnectivity we can this pattern supports creating network effects through inter-connecting users, like in blogs, interests based social networking sites. It is based on a *client-server* architecture where the server interconnects the different users/clients. Web *Semantics* is and technique that exists to facilitate the *social search* and *content recommendation* which is mainly offered to an existing rich *user profile*. Active participation is also encouraged through *commenting* and *rating*.

b- Tagging



This pattern enables users to classify and rate Internet content. Users add their own tags to certain content, as well as provide their own reviews and rating. It is based on a client-server architecture where the server interconnects the different users/clients. *Folksonomy* is a known technique where search keywords are extracted from user tags rather than using ontologies. *Social search* is a technique use to take advantage of tagging, while *content recommendation* based on a rich *user profiles* is other techniques that rely mainly on tagging.

c- File sharing



This pattern represents sharing multimedia files over the Internet, e.g. images or video files. It is based on peer-peer networks regulated by a server that contains the files (referred to as hybrid peer-peer networks). This is required in order to speed up exchange of files as the main design criteria are speed of transmission. Often this pattern is coupled with a supportive function; *content filtering* to ensure the quality of a content, *rating* and *commenting* to encourage active participation, as well as features making use of the file sharing pattern like *social search*, *user profiles* and *content recommendation*.

Fig. 1.1, 1.2, 1.3 *Interconnectivity*, *File sharing* and *Tagging* usage patterns

Remark 1. *In figures 1.1, 1.2, 1.3 the black rectangles refer to architectures whereas the grey ones refer to techniques.*

Some intersections among SNs usage patterns exists, which means that come some patterns rely mainly on the same techniques in order to function. This intersection could lead to another supportive usage patterns.

The intersection between *inter-connectivity* and *file sharing* for instance will lead to *content remixing* where members of SNs could customize their pages by adding and sharing apps and gadgets to offer more functionalities not basically offered by the social network.

Whereas the intersection between *inter-connectivity, file sharing* and *tagging* will lead to *content authoring* pattern where the user is enabled to share a post or note, that could be any multimedia file, tag this contribution and publish it, *content aggregation & syndication* is yet another existing pattern in that intersection that facilitate web feeds from multiple sources to be subscribed to and displayed in a single web page.

Finally *content streaming* pattern exists within the *file sharing* usage pattern, where shared multimedia files may rely on streaming techniques. Fig 2 shows three basic design patterns, with others supportive ones led by their intersection.

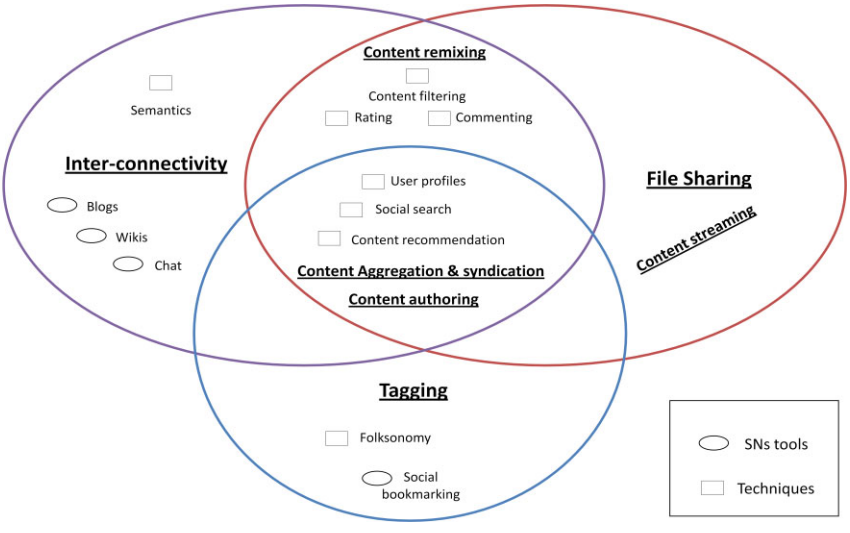


Fig. 2. Roadmap for SNs usage patterns and related tools and techniques

5 Conclusion

Basic social network patterns of usage have been identified, to include; *"Interconnectivity"*, *"File sharing"* and *"Tagging"*. Other supportive patterns include *"Content remixing", "content aggregation & syndication", "content authoring" and "content streaming"*. Web 2.0 tools, related techniques and architecture has been derived from literature to combine proposed SN patterns of usage with their related

features. For example, *"User profiles"*, *"Content recommendation", "Social search"* and *"Aggregation & syndication"* were found common among the three usage patterns, these features could act as core building blocks for SNs.

We conclude that Web 2.0 applications with its broader scope focuses more on Inter-connectivity, file sharing and content remixing, whereas the later pattern has a little existence within social networking application where tagging instead came as a core pattern along with inter-connectivity and file sharing. The intersection between our SNs design patterns will lead to other supportive ones, e.g. the intersection between inter-connectivity and file sharing will lead to content remixing, while the intersection between inter-connectivity, file sharing and tagging will lead to content authoring and content aggregation & syndication, and content streaming exists within the file sharing design pattern.

While SNs are still evolving as concepts and applications, we argue that the introduced SN patterns of usage would represent a useful contribution and a roadmap guiding the design and development of SNs. We argue that the proposed classification would be useful for both researchers and practitioners by giving an insight into functional and technical requirements of SNs. This would improve conceptualization on SN specific behavior and help better analyze and design SNs.

The research is still in its first stages and future work will cover applying identified patterns of usage on several real case studies in order to verify the matching of the introduced classification that have been derived from several literatures and published case studies.

References

1. Oreilly, T.: What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Communications & Strategies (1), 17 (first quarter 2007)
2. Kroski, E.: The social tools of Web 2.0: opportunities for academic libraries. Choice 44(12), 11–21 (2007)
3. Franklin, T., Harmelen, M.: Web 2.0 for Content for Learning and Teaching in Higher Education. Franklin Consulting and Mark van Harmelen (2007)
4. Barsky, E., Purdon, M.: Introducing Web 2.0: social networking and social bookmarking for health librarians. J. Can Health Libr. Assoc. 27, 65–67 (2006)
5. Boulos, M., Wheeler, S.: The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. Health Information and Libraries Journal 24, 2–23 (2007)
6. Messinger, P.R., Stroulia, E., Lyons, K., Bone, M., Niu, R.H., Smirnov, K., Perelgut, S.: Virtual worlds - past, present, and future: New directions in social computing. Decis. Support Syst. 47(3), 204–228 (2009)
7. Boyd, D.M., Ellison, N.B.: Social network sites: definition, history, and scholarship. Journal of Computer-Mediated Communication 13 (1), article 11 (2007)
8. Hussein, D., Alaa, G., Hamad, A.: Web 2.0 Based service-Oriented E-learning systems: Recurrent Design and Architectural Patterns. In: Proceedings of 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2010), June 9-11, pp. 227–232. The University of Greenwich,Article 11, London, United Kingdom (2010)
9. Measurement and Analysis of Online Social Networks IMC 2007, October 24-26, San Diego, California, USA. Copyright, ACM 978-1-59593-908-1/07/0010 (2007)

10. Pata, K., Valjataga, T.: Collaborating across national and institutional boundaries in higher education – the decentralized iCamp approach. In: Association for the Advancement of Computing in Education (AACE), pp. 353–362 (2007)
11. Neumann, M., O'Murchu, I., Breslin, J., Decker, S., Hogan, D., Macdonaill, C.: Semantic social network portal for collaborative online communities. Journal of European Industrial Training 29, 472–487 (2005)
12. Li, X., Guo, L., Zhao, Y.E.: Tag-based social interest discovery. In: WWW (2008)
13. Lerman, K.: Social networks and social information filtering on digg. Submitted to International Conference on Weblogs and Social Media (2006)
14. Lerman, K., Jones, L.: Social browsing on Flickr. In: Proceedings of International Conference on Weblogs and Social Media, Boulder, Co, USA (2007)
15. Chowdhury, S., Landoni., M.: News aggregator services: user expectations and experience, vol. 30 (2), pp. 100–115. Emerald Group Publishing (2006)
16. Raunik, A.: An investigation of streaming - webcasting and webconferencing technologies in the US and UK. VALA Travel Scholar (2004), http://www.valaconf.org.au/vala2006/papers2006/100_Raunik_Final.pdf
17. Sharp, D.: Smart Internet 2010—Social Networks. Melbourne, Australia: Swinburne University of Technology/Smart Internet Technology CRC Pty Ltd (March 2006)
18. Kear, K.: Peer learning using asynchronous discussion systems in distance education. Open Learning 19, 151–164 (2004)
19. O'Reilly, T.: What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software (September 30, 2005),http://www.oreillynet.com/lpt/a/6228 (accessed January 29, 2011)
20. McAfee, A.: Enterprise 2.0: The Dawn of Emergent Collaboration. MIT Sloan Management Review 47 (3) (2006)
21. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M.: Pattern-Oriented Software Architecture: A System of Patterns. John Wiley and Sons, Chichester (1996)
22. Koch, N.: Modelling Service-Oriented Systems with UML4SOA. In: 3rd SENSORIA Summer School (SENSUS 2009), Keszthely, Hungary (2009)
23. Mathes, A.: Folksonomies: cooperative classification and communication through metadata (2004), http://www.adammathes.com/academic/computer-mediatedcommunication/folksonomies.pdf (accessed January 29, 2011)
24. Gruber, T.: Ontology of folksonomy: a mashup of apples and oranges. International Journal on Semantic Web and Information Systems 3(1), 1–11 (2007)
25. Godwin-Jones, R.: Tag clouds in the blogosphere: electronic literacy and social networking. Language Learning and Technology 10(2), 8–15 (2006)
26. Cosh, K., Burns, R., et al.: Content Clouds: Classifying Content in Web 2.0 Library, vol. 57(9), pp. 722–729. Emerald Group Publishing (2008)
27. Chawner, B.: Spectators, no players: information managers" use of Web 2.0 in New Zealand, vol. 26(5), pp. 630–649. Emerald Group Publishing (2008)
28. Ganesh, J., Padmanabhuni, S.: Web 2.0: Conceptual Framework and Research Directions. In: AMCIS 2007 Proceedings (2007)
29. Anderson, P.: What is Web 2.0? Ideas, technologies and implications for education. JISC Technology and Standards Watch (February 2007), http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf
30. Chowdhury, S., Landoni., M.: News aggregator services: user expectations and experience, vol. 30(2), pp. 100–115. Emerald Group Publishing (2006)

Ranked Clusterability Model of Dyadic Data in Social Network

R.B. Fajriya Hakim[1], Subanar[2], and Edi Winarko[2]

[1] Statistics Department, Faculty of Mathematics and Natural Sciences
Universitas Islam Indonesia,
Jalan Kaliurang KM 14.5 Sleman, Jogjakarta, Indonesia, 55584
[2] Mathematics Department, Faculty of Mathematics and Natural Sciences,
Universitas Gadjah Mada
Sekip Utara, Jogjakarta, Indonesia 55528
hakimf@fmipa.uii.ac.id

Abstract. The dyads relationship as a substantial portion of triads or larger structure formed a ranked clusterability model in social network. Ranked clusterability model of dyads postulates that the hierarchical clustering process starts from the mutual dyads which occur only within clusters then stop until all of the mutual dyads grouped. The hierarchy process continues to cluster the asymmetric dyads which occur between clusters but at different levels. Then the last process is clustering the null dyads, which is clustered at the end of the hierarchy after all of asymmetric dyads grouped and occur only between clusters at the same level of the hierarchy. This paper explores a ranked clusterability model of dyads from a simple example of social network and represents it to the new sociomatrix that facilitate to view a whole network and presents the result in a dendrogram network data. This model adds a new insight to the development of science in a clustering study of emerging social network.

Keywords: dyad data, ranked clusterability model, social network, network actor.

1 Introduction

Social networking sites have bombarded the Internet and become one of the most past time trends of all time which have a wide range of people who used them on a daily basis over more than past ten years since the Internet itself. Social Networking sites are a growing phenomenon that have become something more interesting and useful on users' communication. The analysis of social networks, on the other side, become a truly emerging field also. This field has emerged as a key technique in modern social psychology, business, marketing, information science and organisational studies. Social network theory representing a social relationships in terms of nodes and their ties. Nodes are the individual users or actors within the networks, and ties are the relationships or linkages between the actors. The focus of the analysis lies on the relationships and the ties between actors within the network and there are many

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 90–99, 2011.

opportunities to use a form of analysis that would enable a view of how actors interact with each other. The study about how actors interact with each other will provide new knowledge to the assesment of overall network. The segmentation or clustering about the interaction of actors in the social network has become a focus on marketing science or political sciences. Different in the unit analysis of conventional statistics that use individual as an object of observation, the social network analysis use an entity that consisting of a collection of individuals and the linkages among them. Network analysis focus on two actors and their ties (dyads), three actors and their ties (triads) or larger systems (subgroups) or entire networks, but in the actual networks, most of the information is at the dyad level and in the real social networks in which the dyad is necessarily embedded, all those other actors intertwine with one or more of its members. A substantial portion of triadic or subgroups or entire networks structure is explained by nodal and dyadic features [4]. Any dyad needs to exist within such groups and is therefore accountable to them to some extent of grouping analysis. In its present form it dates back to Davis and Leinhardt [2] and briefly summarized by Wasserman and Faust [1] that introduced ranked clusterability model for network data and described it as a building that involves a three type of dyad connection. This paper gives a detailed account of the dyadic approach to ranked clusterability analysis, including how this method considering a level of connection between dyad and differs from the more common individualistic approach when the dyad and not the individual are used as the unit of analysis and presents the result in the dendrogram network data.

2 Literature Review

2.1 Studies about Dyad

Kenny et al. [7] have explored about dyadic data analysis. They considered that any relation in the social network occur in dyad relationship because the dyad is fundamental unit of interpersonal interaction and interpersonal relations. Most of their methodological and data-analytic approaches used conventional statistical method and do not give more emphasize about three types of dyadic relationship which involves mutual, null and asymmetric relationship that could make dyad grouping. Yablonsky [8], insisted that two actor interaction or dyad is a foundation of both individual and group strength in the social network. His research explored a theory of the dyad and its dimensions and also implications to each member's individual social atom and emphasize the dyad as the nucleus of its own social atom in such sociometric situations. DeCoster [6] gave the flowchart to examine data from groups and dyads using Anova, He considered about the independene and nonindependence as a relation of an actor in the analysis of variance techniques. Aurifeille and Medlin [9] has applied a study about dyad in business science, especially about segmentation of business partnership. They describe the business partnership as a dyadic structure that influences the firm's performance and using the relationship modes of merging, teaming and sharing. Their dyad segmentation strategies comply the capacity of dyadic relationship to reflect the modes of partnership. Menon and Elkan [10] studied the problem of predicting labels associated with dyad members using new method

which learns latent features within a log-linear model in a unsupervised way. They also compare their new approach which maximizing predictive accuracy for both dyad observations and item labels to existing methods for within-network classification. Mizruchi and Marquis [11] contrasted the dyadic analysis with both individual and system level analysis in their study of corporate political behaviour and suggested that eventhough analysis of individual and system-level are useful but the dyadic analyses are a flexible means to examine the effects of multiple networks at multiple levels. All of the researchers mentioned above emphasize on the analysis of dyad relationship. This paper will use the idea of Davis and Leinhardt [2] that introduced a ranked clusterability model in the study of grouping the two actors in the network and presenting the model in the sociomatrix and dendrogram network data.

2.2 The Simplest Form of Sociometric Choices Data

It is commonly to present the information about social network using graphs, but they become complicated and difficult to see the patterns since there are many actors and kinds of relation [5]. Figure 1 shows a simple directed graph of friendship choices among A, B, C and D.

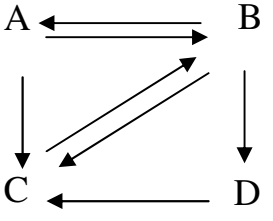


Fig. 1. A digraph of friendship choices

By converting those directed graph to the matrix, the rows will represent the source of directed ties, and the columns the targets, then the dyadic relationships of each actor shown in Table 1. Each cell of the array describes a relationship between the actors which is coded zero if the relations being absent and coded one if the ties being present. There are four possible dyadic relationships for directed data: A and B are not connected, A sends to B, B sends to A, or A and B send to each other, meanwhile there are only two possible relationship for undirected data, A and B connected or not connected.

Table 1. Friendship choices

Chooser:	Choice:			
	A	B	C	D
A	0	1	1	0
B	1	0	1	1
C	0	1	0	0
D	0	0	1	0

Using a sociomatrix representation (Figure 2a), each cell on the main diagonal (e.g. A likes A, B likes B) corresponds to self choice is assumed to be zero [3]. In this matrix, we could see that A chooses C, but C does not choose A. This is an example of an asymmetric matrix that represents directed ties (ties that go from a source to a target). That is, the element *i*, *j*; where *i* and *j* represent the numbers of rows and columns of matrix, does not necessarily equal the element *j*, *i*. If the ties in our matrix representation were undirected, the matrix would necessarily be symmetric; that is, the element *i*, *j* would be equal to element *j*, *i*.

By simple similarity analysis we could treat this data structure using conventional attribute data. We can see similarities between the choices of actors with each other by comparing each row. Meanwhile, by comparing each column we can see the similarities between the selected actors. This method is a simple and useful to see which actors who have a similar position in the network. The study of the position of an actor or an actor embedded among others in a network is a first major emphasis in network analysis. The second major emphasis in network analysis is comparing the cells above and below the main diagonal to check whether there is reciprocity in choices (e.g. A choose C, did C choose A?) and this leads the analyst to see whether the patterns of all individual choices give rises to more overall network pattern [5]. Many studies used by network analysts treat the network data as a conventional data and describing the network data as just a special form of conventional data. This paper trying to clustering actors by looking the network data in a different ways.

3 A Ranked Clusterability Model of Dyadic Data

3.1 A New Sociomatrix

In this paper, we present a different ways to display a network data matrix. We combine cells in the bottom and the top half of the main diagonal to be the cells in the top half of the main diagonal. For this, we move all cells in the bottom half of the main diagonal matrix to the top half of the main diagonal matrix as shown in Figure 2b, from this matrix we could see the reciprocity of ties and form the hierarchy in a network.

	A	B	C	D
A	0	1	1	0
B	1	0	1	1
C	0	1	0	0
D	0	0	1	0

	A	B	C	D
A	0|0	1|1	1|0	0|0
B		0|0	1|1	1|0
C			0|0	0|1
D				0|0

Fig. 2. (a) Sociomatrix of table 1. (b) A sociomatrix choices network.

In this new sociomatrix that contains directed data, the conventions are the first value in the entries (1|0) presents the direction of a relation as running from a row element to a column element and the second value (0|1) presents the direction of a relation as running from a column element to a row element. This matrix helps us to see which actors that connect to each other or reciprocal and which are not. If the

entry in the cell is (1|1) means that the two actors are connected. If the entry in the cell is (1|0), the actor who is on the row chooses the actor who is in the column, but not vice versa. If the entry in the cell is (0|1) means that the actor who is in the column chooses the actors in a row, but not vice versa. If the entry in cell (0|0) means that the two actors are not connected. The entries contained in any main diagonal is (0|0) for the reason mentioned in [3]. The table also allows us to quickly see whether the network data is symmetric (undirected ties) or asymmetric matrix (directed ties).

3.2 A Ranked Clusterability Model

Wasserman and Faust [1] and Davis and Leinhardt [2] stated that one needs to examine three kinds of dyads when studying ranked clusterability model for complete signed graphs, there are: ++ or (1|1) dyads, in which both actors are connected; - - or (0|0) dyads, in which both actors are not connected and +- or (1|0) or (0|1) dyads, in which one actor chooses his pair but not vice versa. The ranked clusterability model is started by treating a group's ranking structure as a series of ordered levels [2]. The ordered levels was thinked as stories in a building, in the sense that people on a given floor do not differ in level, any two persons on different floors are unambiguously ordered by level, then floors and ceilings mark the levels in a building, and the stories form a complete order. Wasserman and Faust [1] have explained briefly in which ranked clusterability postulates that (1|1) dyads occur only within clusters and (0|0) dyads occur only between clusters at the same level of the hierarchy or order of clusters. The interesting (0|1) or (1|0) dyads also occur between clusters, but at different levels.

3.3 A Level of Connection

We see from the Fig. 2b, the entries in cell (A, B) and (B, C) are (1|1), that is, they are reciprocal. For the clustering process, we need to get an additional measure to differentiate between (A, B) and (B, C). Borgatti [12] and Hanneman and Riddle [5] have described about measurement of centrality and flows of networks including geodesic distances, Katz [13], Hubbell, Hoede and Taylor's measurement. In this paper, we will use the ideas of the geodesic distance and the Taylor's measurement. The geodesic distance examines a shortest possible connection between a pair of actors. Taylor's measurement consider the direction of connections, the column marginal of network matrix subtracted from the row marginals that means the balance between each actor's sending connections and their receiving connections [5]. This paper try to combines those two ideas for a level of connection of pair of actors. In any network, if n is the number of actors, there are (n * (n-1)) unique ordered pairs of actors (AB different from BA) and the exclusion of self-ties relationship. We have four actors in our network example, with directed ties, there are 12 possible relationships. We will see how much the dyadic relationship that occurs with A and B, whether it be as a source or as a target.

First, the directed ties from A to B. In this case we want to see how much of an actor A sends to the other actors in addition to actor B. This means that the actor A is not just liking an actor B but he did share it with other actors. We assume that there are no possibility that choices have different strengths attached to them and also they have the same kind of choice types. Matrix in Fig. 3 is already added each row and

	A	B	C	D	Total
A	-	1	1	0	**2**
B	1	-	1	1	3
C	0	1	-	0	1
D	0	0	1	-	1
Total	1	**2**	3	1	

(a)

	A	B	C	D	Total
A	-	1	1	0	2
B	1	-	1	1	**3**
C	0	1	-	0	1
D	0	0	1	-	1
Total	**1**	2	3	1	

(b)

Fig. 3. Matrix to count the level of connection from actor A to B (a) and from B to A (b)

column marginals to help the calculations. The value of connections from A to B could be seen from Fig. 3a, which row and column are shaded to show the total choice of A and the total how many actors who choose B.

Hence, the level of connections of a directed ties from A to B is (2+2)/12 = 4/12. *Second*, because A and B are reciprocal, then we count the level of connections of a directed ties from B to A. B not only choose A, but he also choose C and D, while, A is the only choosed by B. The calculation showed by the shaded row and column in the Figure 3b. The value of the connections of a directed ties from B to A is (3+1)/12 = 4/12. Hence, the level of connections a dyadic relationship of A and B is (4/12) + (4/12) = 8/12. We could add this level of connection as an additional measurement of matrix in the Fig. 2b and present it on the Fig. 5 for all calculations of dyadic relationships.

3.4 A Formal Definition

If the nodes are numbered in some arbitrary way from 1 to *n*, then we may create a new useful matrix representation of the digraph as follows. Note that we use a common matrix **X** whose the entry X_{ij}, *i* and *j* represent the numbers of row and column of matrix, respectively. We will also use the entry X_{ji} that alter the position of actor *i* in the rows and actor *j* in the column to be the actor *j* in the rows and actor *i* in the columns from the matrix X_{ij}, where *j* and *i* represent row and column, respectively. Now, let **X** be a *n* by *n* matrix whose (*i*, *j*) entry is $X_{ij} \mid X_{ji}$ where,

$$X_{ij} = \begin{cases} 1 & if\ i \to j \\ 0 & otherwise \end{cases}$$

Note that $i \to j$ means that there is a directed line from node *i* to node *j*; and a directed line from node *j* to node *i*,

$$X_{ji} = \begin{cases} 1 & if\ j \to i \\ 0 & otherwise \end{cases}$$

The main diagonal of **X** corresponds to self choice in the sociometric context, and for many reasons it is convenient for us to assume that $X_{ij} \mid X_{ji}$ equals to zero if $i = j$.

The row and column sums of a sociomatrix X will be denoted by $\{\sum_{i=1}^{n} X_{ij}\}$ or $\{\sum_{j=1}^{n} X_{ji}\}$ and $\{\sum_{j=1}^{n} X_{ij}\}$ or $\{\sum_{i=1}^{n} X_{ji}\}$, respectively. In the sociometric context, $\{\sum_{i=1}^{n} X_{ij}\}$ (or $\{\sum_{j=1}^{n} X_{ji}\}$) is the number of choices made by individual *i* (or *j*, since *i* alter the position of *j*), while $\{\sum_{j=1}^{n} X_{ij}\}$ (or $\{\sum_{i=1}^{n} X_{ji}\}$) is the number of choices received by individual *j* (or *i*, since *i* alter the position of *j*). The level of connection $\gamma(X_{ij} \mid X_{ji})$ of node *i* to node *j* and node *j* to node *i* is

$$\gamma(X_{ij} \mid X_{ji}) = \begin{cases} \dfrac{\sum_{i=1}^{n} X_{ij} + \sum_{j=1}^{n} X_{ij}}{n(n-1)} + \dfrac{\sum_{j=1}^{n} X_{ji} + \sum_{i=1}^{n} X_{ji}}{n(n-1)} & \text{if } i \text{ and } j \text{ mutual relations} \\ \dfrac{\sum_{i=1}^{n} X_{ij} + \sum_{j=1}^{n} X_{ij}}{n(n-1)} & \text{if } i \rightarrow j \\ \dfrac{\sum_{j=1}^{n} X_{ji} + \sum_{i=1}^{n} X_{ji}}{n(n-1)} & \text{if } j \rightarrow i \\ 0 & \text{otherwise} \end{cases}$$

The sociomatrix **X** can be described as in figure below

	1	2	...	*J*	...	*n*
1	$X_{11} \mid X_{11}$; $\gamma(X_{11} \mid X_{11})$	$X_{12} \mid X_{21}$; $\gamma(X_{12} \mid X_{21})$	...	$X_{1j} \mid X_{j1}$; $\gamma(X_{1j} \mid X_{j1})$	...	$X_{1n} \mid X_{n1}$; $\gamma(X_{1n} \mid X_{n1})$
2		$X_{22} \mid X_{22}$; $\gamma(X_{22} \mid X_{22})$	...	$X_{2j} \mid X_{j2}$; $\gamma(X_{2j} \mid X_{j2})$	...	$X_{2n} \mid X_{n2}$; $\gamma(X_{2n} \mid X_{n2})$
...			...	...	...	...
i				$X_{ij} \mid X_{ji}$; $\gamma(X_{ij} \mid X_{ji})$	...	$X_{in} \mid X_{ni}$; $\gamma(X_{in} \mid X_{ni})$
...					...	...
n						$X_{nn} \mid X_{nn}$; $\gamma(X_{nn} \mid X_{nn})$

Fig. 4. A new sociomatrix

3.5 Hierarchical Clustering Process

Ranked clusterability model of dyads postulates that the clustering process starts from the mutual dyads that occur only within clusters until all of the mutual dyads grouped then the hierarchy process continue to cluster the asymmetric dyads that occur between clusters but at different levels and the last process is clustering the null dyads, which is clustered at the end of the hierarchy after all of asymmetric dyads grouped and occur only between clusters at the same level of the hierarchy. The hierarchical process uses the following steps.

Step 1. Form a new sociomatrix of sociomatrix choices network.

	A	B	C	D
A	0|0	1|1 (8/12)	1|0 (5/12)	0|0 (0)
B		0|0	1|1 (9/12)	1|0 (4/12)
C			0|0	0|1 (4/12)
D				0|0

Fig. 5. A Sociomatrix with the level of connections

Step 2. Find the maximum value in the matrix using ranked clusterability model.
The clustering starts from the mutual dyads and then the asymmetric dyads and finally the null dyads. The matrix (Fig.5) shows that there are mutual dyadic relationships i.e., AB and BC or both of them are a reciprocal relationships. Then the value that we will use to start the clustering is the maximum value of the level of connections between AB or BC. We choose BC because their degree (9/12) is higher than AB (8/12). The hierarchical clustering start from BC.

	BC	A	D
BC	1|1(9/12)	1|1(8/12)	1|0(4/12)
A		0|0	0|0 (0)
D			0|0

Fig. 6. First step of clustering

Step 3. Modify the sociomatrix according to step 2.
From step 2, BC occupies the first column, next column is occupied by A and D. The value in columns BC and A obtained from the search for maximum value between the BA with CA. The value of BA is (1\|1\|(8/12)) while the value of CA is (1\|0\|(5/12)) therefore, we select the value of BA (Fig. 6). This was also done for the entries of other cells (Fig. 7).

	BCA	D
BCA	1|1(8/12)	1|0(4/12)
D		0|0

Fig. 7. Second step of clustering

this clustering process continues until found one large cluster (Fig. 8) and the last clustering process is

	BCAD
BCAD	1|0(4/12)

Fig. 8. The final step of clustering

Step 4. Draw the final dendrogram.
Then we can get the dendrogram of actors in the network (Fig. 9).

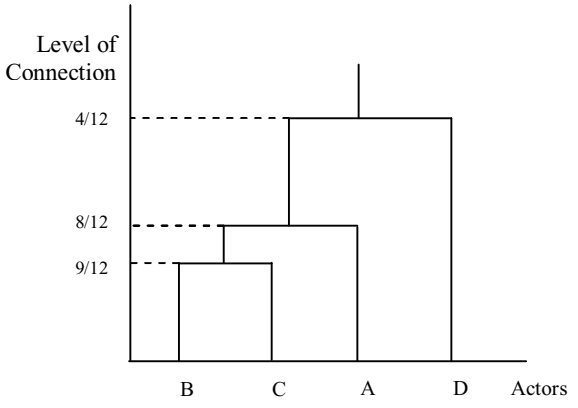


Fig. 9. Dendrogram network data

From the dendrogram we could see that actor B and C as a basis for the friendship choices in the network. Eventhough both of AB and BC are mutual relationship but most of the information is at the both two actor B and C, B chooses all of friends in the network and C has been chosen by all of friends in the network. If we would like to make two cluster based on ranked clusterability of dyadic data then the group is (A, B, C) and (D). D has the lowest level of connection of dyad relationship with other friends. This simple example of ranked clusterability gives a new insight in how to view the relationship among actors in the network using sociomatrix, level of connections and dendrogram network data.

4 Conclusion

This paper shows the ranked clusterability model idea from Davis and Leinhardt [2] using new sociomatrix and its level of connection. The new sociomatrix has already help us to understand about ranked clusterability model that can be presented in a dendrogram network data. In this simple example of choice network data we assume that there are no possibility that choices have different strengths attached to them and also they have the same kind of choice types. For the next research, a real network data will be used and will be compared to other methods that usually used to cluster

the dyads data. Eventhough this research starts from simple example of friendship choices data, the result gives us a new insight in how to view the clustering and relationships among actors in the network.

References

[1] Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1994)

[2] Davis, J.A., Leinhardt, S.: The Structure of Positive Interpersonal Relations in Small Groups. In: Berger, J., Zelditch, M., Anderson, B. (eds.) Sociological Theories in Progress, vol. 2, Houghton-Mifflin, Boston (1971)

[3] Holland, P.W., Leinhardt, S.: Local Structure in Social Networks. Sociological Methodology 7, 1–45 (1976)

[4] Faust, K.: Very local structure in social networks. Sociological Methodology 37(1), 209–256 (2007)

[5] Hanneman, R.A., Riddle, M.: Introduction to Social Network Methods. University of California, Riverside (2005)

[6] DeCoster, J.: Using ANOVA to Examine Data from Groups and Dyads (2002), `http://www.stat-help.com/notes.html` (retrieved from the web January 14, 2011)

[7] Kenny, D.A., Kashy, D.A., Cook, W.L.: Dyadic Data Analysis. Guilford Publications, New York (2006)

[8] Yablonsky, L.: The Sociometry of the Dyad. Sociometry 18(4), 357–360 (1955)

[9] Aurifeille, J.-M., Christopher, J.M.: A Dyadic Segmentation Approach toBusiness Partnerships. European Journal of Economic and Social Systems 15(2), 3–16 (2001)

[10] Menon, A.K., Charles, E.: Predicting labels for dyadic data. Data Mining and Knowledge Discovery 21, 327–343 (2010)

[11] Mizruchi, M.S., And Marquis, C.: Egocentric, sociocentric, or dyadic? Identifying the appropriate level of analysis in the study of organizational networks. Social Networks 28(3), 187–208 (2006)

[12] Borgatti, S.P.: Centrality and Network Flow. Social Networks 27, 55–71 (2005)

[13] Katz, L.: A New Status Index Derived From Sociometric Analysis. Psychometrika 18(1) (1953)

Forensic Analysis of Residual Information in Adobe PDF Files

Hyunji Chung, Jungheum Park, and Sangjin Lee

Center for Information Security Technologies,
Korea University, Seoul, Korea
{foryou7187,junghmi,sangjin}@korea.ac.kr

Abstract. In recent years, as electronic files include personal records and business activities, these files can be used as important evidences in a digital forensic investigation process. In general, the data that can be verified using its own application programs is largely used in the investigation of document files. However, in the case of the PDF file that has been largely used at the present time, certain data, which include the data before some modifications, exist in electronic document files unintentionally. Because such residual information may present the writing process of a file, it can be usefully used in a forensic viewpoint. This paper introduces why the residual information is stored inside the PDF file and explains a way to extract the information. In addition, we demonstrate the attributes of PDF files can be used to hide data.

Keywords: PDF, Residual Information, Data hiding, Information leakage, Digital evidence.

1 Introduction

People create their electronic documents using various application programs, such as Microsoft Office, Adobe Acrobat, and so on. Although the use of document files is widespread, few people recognize that 'hidden' data exist in files. The reason for using the word 'hidden' is that identification of these data is not possible using its own application program.

J. Park et al. introduced 'hidden' data in MS PowerPoint files. Even after the contents of an MS PowerPoint file have been deleted or edited, they can still exist inside the file as residual information. This is because file-saving algorithm used in this application[2]. Similar to MS PowerPoint files, Adobe PDF (Portable Document Format) files can also contain 'hidden' data. Therefore, it is necessary to investigate such 'hidden data' in a digital forensic viewpoint.

Adobe PDF files have been largely used for various purposes such as writing personal documents and distributing official documents at enterprise. In particular, for enterprise purposes, some confidential documents were created using applications like Microsoft Word or PowerPoint. And they were distributed after transforming into Adobe PDF files. In the past, PDF files were just application for distribution. However, the content of PDF files can be edited using various PDF editors in these days.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 100–109, 2011.

The following is a hypothetical case related to Adobe PDF files. Company “Y” recruited new business partner for new technologies. As a result, both company “A” and company “B” were on the shortlist. The company “B” bought up a core member of the company “A” in order to modify final proposal and be business partner with company “Y”. Finally, the modified file was submitted, and company “B” was selected as a partner. Some three years later, certain suspicious elements were detected in the competitive selecting process, following which investigation of the business was undertaken. The only evidence was the Adobe PDF file created for the final proposal after long time.

In the case, forensic examiners were able to investigate some evidences in Adobe PDF file. This is because the contents have been deleted or edited can exist as residual information inside the file due to the file-update mechanism.

This study consists of seven sections. Section 2 introduces the existing studies on Adobe PDF files. Section 3 describes the internal structure of Adobe PDF files. Section 4 represents the reason that PDF files can include residual information. Section 5 proposes a way to extract the residual information. Section 6 describes a data hiding method using a PDF update mechanism. Finally, Section 7 represents the conclusion of this study.

2 Related Studies

In 2008, Didier and Matthew introduced that updated PDF files can include revision history by updating mechanism of Adobe PDF. That is, they explained it is possible to know revision history in a PDF file when the file is modified and saved using ‘Save’ instead of ‘Save As’ function[1][5].

Didier was developed tools, such as pdf-parser.py, pdfid.py for analysis of PDF files. The tools help examiners to identify an internal structure of PDF files. However, it is difficult to identify specific ASCII, Unicode text for digital forensic investigation[5].

Also, Matthew was developed a tool, pdfresurrect. The pdfresurrect shows the revision number and time of a PDF file. This tool makes investigators to confirm revision history. However, since it cannot verify what data is modified, there is a limit to utilizing in digital forensic purposes[1].

Furthermore, this study researches about unused area in PDF documents from the viewpoint of digital forensics. This paper introduces a tool for analyzing PDF files. It is developed with the special function of extracting text data(both ASCII and Unicode) at each revision. It is easy to compare modified version with original version. Therefore, the result of this paper is useful for investigating Adobe PDF files. Also, this paper explains that unused area in PDF files can be used for data hiding.

3 Internal Structure of PDF Files

As shown in Figure 1, a PDF file consists of Header, Body, Cross-reference table (hereinafter called xref), and Trailer[3].

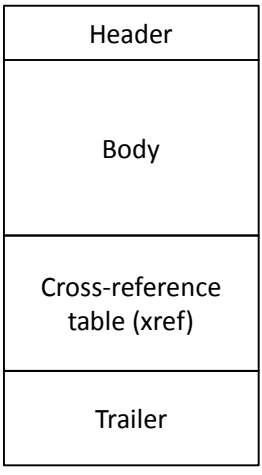


Fig. 1. Basic Structure of PDF files

The first line of a PDF file is header that has a version number. The trailer is at the end of a PDF file. It has byte offset of the last xref[3].

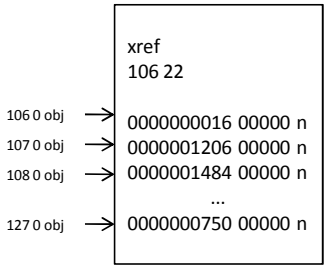


Fig. 2. Example of cross-reference table(xref)

Xref contains information of all objects in a PDF file. The example of the xref is shown in Figure 2. There can be several xrefs, but Figure 2 shows only one of them. The term of "106 22" means that there exists twenty-two objects from the 106th object. Then, the first ten places in the next line represent the position of the 106th object as a byte offset. The second five places mean the generation number, and the value of '00000' is allocated as it is first generated. For the third place, 'n' or 'f' can be positioned where 'n' is the object, which is being used, and 'f' is the object, which is not used (i.e., free)[3].

Many objects exist in a PDF file. Objects are logically made of a tree. As shown in Figure 3, /Root, /Pages, /Kids, /Page and /Contents represents the type of objects.

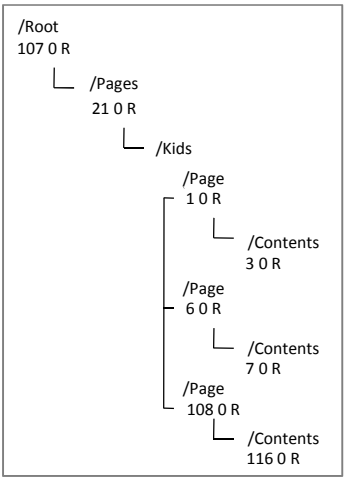


Fig. 3. Internal structure of a PDF file

Although there are various different types in addition to the type presented in Figure 3, this study refers the main types only. In Figure 3, the root of the internal structure tree is the /Root type. The /Root contains an indirect reference, "107 0 R". In detail, xref in Figure 2 includes byte offset of "107 0 obj". /Pages type exists at byte offset 1206. Continuously, /Pages has indirect reference, "21 0 R". The "21 0 obj" is /Kids type, and it is possible to access the position of "21 0 obj" through proper xref. /Kids consists of one or more indirect /Page references. It is possible to know how many pages PDF file has. This is because a /Page type means each page inside a file. Figure 3 is an example of the PDF file that has three pages. In figure 3, the one of /Page types has indirect reference, "1 0 R". The "1 0 obj" is /Contents type, and the position of it can be accessed using xref. Also, /Contents has indirect reference, "3 0 R". The "3 0 obj" is the body of a page, and the data will be acquired through xref that saves the position as a byte offset.

In addition, the body data can be a compressed(or encoded) format. Thus, it is necessary to decompress(or decode) the data in order to confirm plain text.

4 Residual Information in PDF Files

4.1 Update Mechanism of PDF Files

Lots of users recently build their electronic documents using specific application programs like Microsoft Office 2007 and store them as PDF files through PDF transformation process. The Adobe Acrobat 8.0 was used in this experiment. Table 1 shows the experiment procedure.

Table 1. Experiment procedure

Step1	Create three-page contents (texts and images) using Microsoft Office Word 2007 and store it as a PDF file using the function of "Save as Adobe PDF".
Step2	Open the PDF file using Adobe Acrobat.
Step3	Modify the content in the first and second pages of the original PDF file into different texts and images after deleting the existing them.
Step4	Resave it using the function of "Save" after completing the modification.

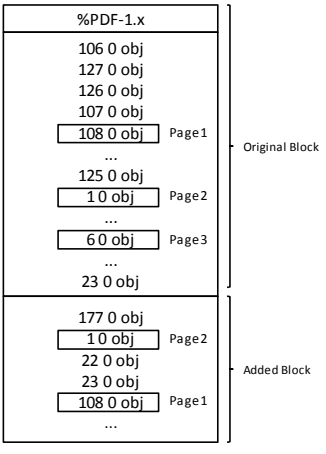


Fig. 4. Structure after completing the experiment

Figure 4 shows the internal format after finishing the experiment. Repeating Step 3 and Step 4 attaches 'added block' to 'original block'. "108 0 obj" contains contents of the first page. "1 0 obj" has them of the second page and "6 0 obj includes them of the third page.

Both "108 0 obj" and "1 0 obj" exist in 'added block', and "6 0 obj" doesn't exist in 'added block'. This is because the third page is not modified. As shown in Figure 5, it can be seen that the size of the modified file is bigger than the size of original one.

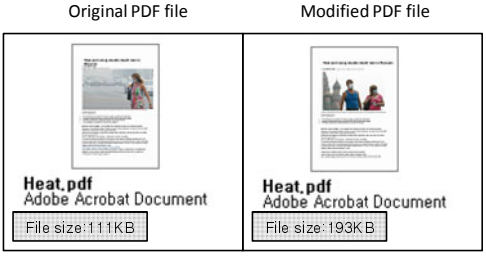


Fig. 5. Comparison of original file size and modified file size

Residual information is generated when the file is updated. Figure 4 and 5 show this feature. This feature improves the efficiency of saving a PDF file. In other words, it takes less time than a full save of the file. The reason for using 'residual' is that it cannot be identified by PDF application. However, if a user saves the file using the "Save As" function, the application does reconstruct the entire structure.

4.2 Residual Information In PDF Files

Figure 6 explains the concept of residual information. In Figure 6, the left side is xref of the modified file, and the right side is the internal structure. "108 0 obj" and "1 0 obj" exist in both original and added blocks. "108 0 obj" in 'added block' is appended after modifying the file. Therefore, "108 0 obj" in 'original block' exist in the file as the residual information.

As the first and second pages are modified, the objects related to the modification are included in both 'original block' and 'added block'. The first page after the modification does not use the 108th object in the 'original block' but use the 108th object in the 'added block'. The 108th and the 1st objects in the 'original block' are not used by the PDF file viewer.

5 Extraction of Residual Information

5.1 Recovering the File to Original Version

In figure 6, the first page of the PDF file after the modification uses the 108th object (byte offset 119465) in the 'added block'. Also, the second page uses the 1th object (byte offset 114777) in the 'added block'.

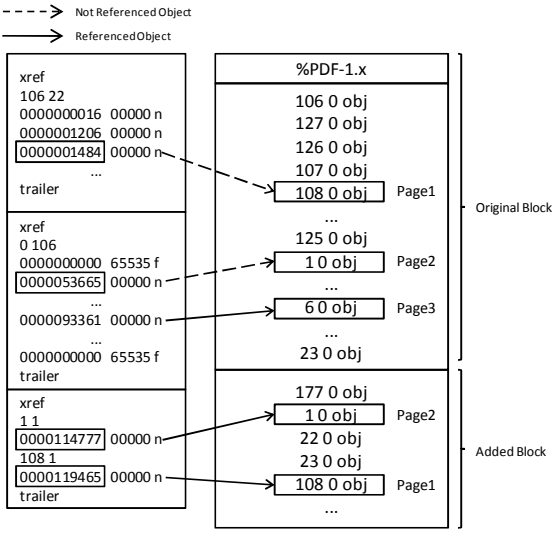


Fig. 6. State of reference in modified file

In Figure 8, to access the contents of original file, there is one methodology that it changes byte offset 114777 into byte offset 53665. In addition, it replaces byte offset 119465 with byte offset 1484. Then, the file is recovered to original version. Also, the contents of the original file can be viewed using a PDF file viewer.

Figure 7 shows the contents identified by a PDF file viewer. It is verified that the size of the recovered file is the same as the modified file and shows the same contents as the original file.

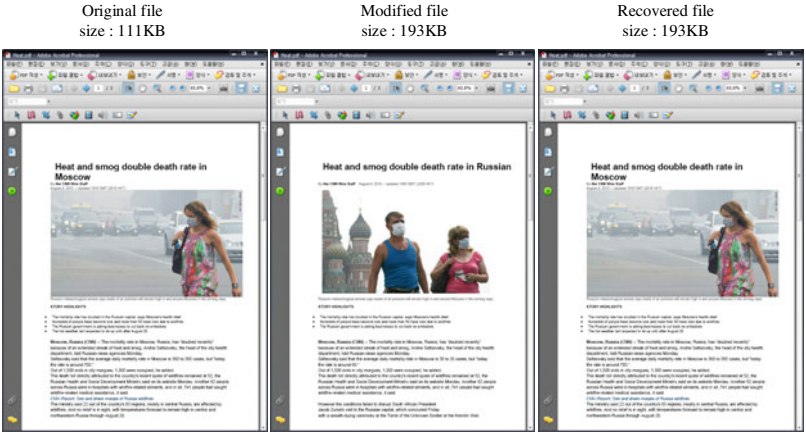


Fig. 7. Size and contents of original file, modified file, and recovered file

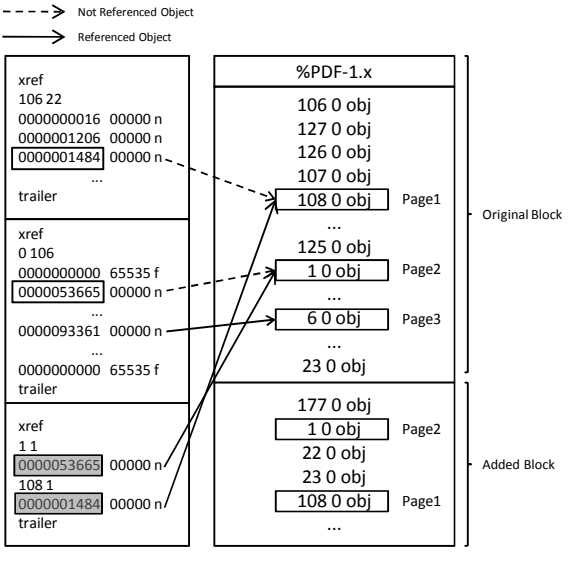


Fig. 8. Method of recovering the file

5.2 Extraction of Texts and Images

Another method of extracting residual information is to directly extract text and images. Since the rules of storing text and images in PDF files are presented in [3], the contents can be extracted directly.

In case of the text data, there are two types of ASCII and Unicode. The text data can be extracted by decompressing (or decoding) the data of a target object. Although decompressed (or decoded) data is acquired, it is difficult to extract exact text data. Since the PDF file saves the text with some information such as position, shape and so on, it is necessary to do additional parsing process.

Also, the images in the content can be directly extracted because images are stored as their own image formats.

5.3 ExPRI(Extractor for PDF Residual Information)

To investigate the forensic attributes of Adobe PDF file, ExPRI(Extractor for PDF Residual Information) has been developed. It is difficult for investigator to change manually byte offset using methodology in Figure 8.

The ExPRI is able to recovery previous files and to extract the text at each storing point. Left side in Figure 9 shows result of extracting modified PDF file's contents. Right side in Figure 9 shows the result of extracting recovered PDF file's contents. Modified and recovered file in Figure 7 demonstrate that ExPRI extracts text correctly.

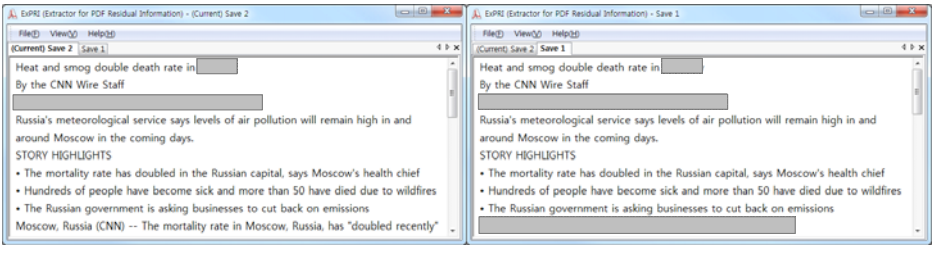


Fig. 9. ExPRI(Extractor for PDF Residual Information)

6 Data Hiding Method Using File-Update Mechanism

6.1 Data Hiding Technique 1

The actually viewed section is the part of 'valid data' of the modified file presented in Figure 10. It means that although the area of 'hidden data' exists inside the file, it is not used. This area can be used to hide data. Even though 'hidden data' is stored in a PDF file, the data cannot be identified using a PDF file viewer. Since the data is compressed (or encoded), it is not possible to verify such data using a simple method like Strings. Therefore, it can be used as a method that simply hides huge amounts of data.

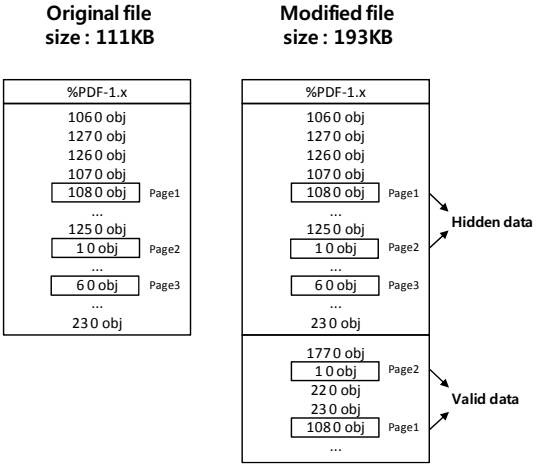


Fig. 10. Hidden and valid data in the modified file

6.2 Data Hiding Technique 2

Because the area of 'hidden data' presented in Figure 10 is not used, the area can be used for hiding the data. Figure 11 illustrates the example of such a way. As the data is hidden using this method, it is not possible to verify the data by extracting the residual information. Also, it is very strong way to hide data with encryption algorithms. Thus, it is very difficult to investigate such hidden data.

These two methods presented above represent an advantage that is able to hide huge amounts of data compared with the method proposed by Shangping Zhong et al. [4].

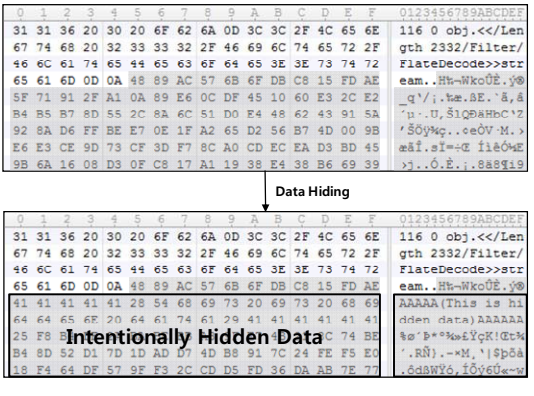


Fig. 11. Data hiding technique 2

7 Conclusion

In digital forensic investigation, there are some cases which documents files are analyzed as evidences. Because researches on electronic document forensics are not enough until now, the analyses have been focused on the contents that can be easily verified by using specific applications. However, this method is insufficient for analyzing PDF files because there are 'hidden data' inside them. In this vein, it is expected that the method introduced in this study is useful to investigate Adobe PDF files. Also, this study will help to improve the admissibility of electronic documentary evidence.

In this study, we analyzed the structure and attributes of Adobe PDF files in a digital forensic viewpoint. It is possible to trace the previous work in a file using its residual information remained in the file based on the update mechanism of an Adobe PDF editing program. In addition, because the area of residual information is not identified by a PDF file viewer, it can be used as a way of data hiding.

For future study, the tool that has been developed at the present time will be completed. It can be used as a forensic analysis module for PDF files because it includes previous file recovery, text and image extraction, and metadata extraction functions for each storing point. In addition, a study on the method that intentionally hides data to the unused area where residual information are stored will be deeply conducted.

Acknowledgments. This work was supported by the IT R&D program of MKE/KEIT [10035157, Development of Digital Forensic Technologies for Real-Time Analysis]

References

1. Davis, M.R.: Faith in the Format: Unintentional Data Hiding in PDFs. 757 Labs.com (2008)
2. Park, J., Lee, S.: Forensic investigation of Microsoft PowerPoint files. Digital Investigation 6, 16–24 (2009)
3. Adobe Systems.: Document management - Portable document format Part 1: PDF 1.7, 1st edn. (2008)
4. Zhong, S., Cheng, X., Chen, T.: Data Hiding in a Kind of PDF Texts for Secret Communication. International Journal of Network Security 4(1), 17–26 (2007)
5. Stevens, D.: Solving a Little PDF Puzzle (2008), http://blog.didierstevens.com/2008/05/07/solving-a-little-pdf-puzzle/

Digital Quantum Forensics: Challenges and Responses

Richard E. Overill

Department of Informatics, King's College London, Strand, London WC2R 2LS, UK
richard.overill@kcl.ac.uk

Abstract. Recent developments in technology suggest that within the next one to two decades some form(s) of quantum computing device will become viable. Once such devices become generally available they will doubtless be employed by cyber-criminals to perform brute-force decryption tasks that were previously infeasible. This paper attempts to address the question of how digital forensic investigation techniques will need to develop in order to respond to the challenges posed by such future generation computing devices.

Keywords: quantum computation; digital forensics; quantum forensics.

1 Introduction and Background

The current state-of-the-art in quantum computing suggests that, despite much theoretical and experimental work over the past 25 years, the topic is still in its infancy, both as regards the technologies in which it can be implemented and also the range of algorithms available.

Current candidate technologies include laser-excited atomic ion traps using beryllium or calcium atoms, bulk liquid-phase and solid-phase nuclear magnetic resonance (NMR), and superconducting solid-state circuits operating at liquid helium temperatures (*circa* 4 degrees Kelvin). To date the maximum number of participating quantum bits (qubits) that has been reached is 12 [1], which in principle enables problems of size $2^{12} = 4096$ to be addressed.

The currently available algorithms for quantum computers include Shor's $O(n^3)$ integer factorization and discrete logarithm [2, 3], which could be employed cryptanalytically to attack both RSA-like and EC-based public-key cryptosystems, and Grover's $O(\sqrt{n})$ search algorithm [4, 5], which could be used in an exhaustive search of the key-space of a conventional symmetric cryptosystem such as AES. Due to the intrinsic nature of quantum computation, as outlined in the following section, these quantum algorithms will be capable of compromising many state-of-the art cryptosystems in a matter of minutes or even seconds of real-time. For example, Shor's algorithm could be used to factor numbers100 decimal digits long in a fraction of a second. Quantum speed-up inevitably means that cryptographic cover times which were previously unbreachable cannot be expected to remain intact in the future. In the context of anticipating future large-scale cyber-crimes executed with the aid of quantum computers, this paper aims to address the key issues of the quality and the quantity of digital forensic evidence that may be recoverable from these devices, and also to indicate how such evidence may be analyzed and interpreted for subsequent forensic use in judicial proceedings.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 110–114, 2011.

2 Quantum Computation

Quantum computation [6] exploits the evolution in Hilbert space of a coherent linear superposition of quantum states such that each component of the superposition follows a distinct computational path. A final decohering operational step, which is equivalent to performing a measurement or an observation, then extracts the required output with high probability. Whereas in classical computing a bit is a binary digit which has a value which is either zero or one, in quantum computing a single qubit can in general be an unequal linear superposition (or 'mixture') of the basis states zero and one:

$$|\Psi> = \alpha|1> + \beta|0>, \text{ where } \alpha^2 + \beta^2 = 1$$

From an n-particle quantum system an n-qubit register (qureg) may be constructed:

$$|\Psi_n> = |1> \otimes |1> \otimes \ldots \otimes |1> \equiv |11 \ldots 1>$$

Applying to this qureg a Walsh-Hadamard transformation, represented by a unitary operator U_H which transforms a single basis eigenstate into an equal superposition of the two basis eigenstates, generates a coherent equal superposition of all 2^n possible inputs:

$$|\Psi_n>' = U_H \otimes U_H \otimes \ldots \otimes U_H |11 \ldots 1>$$

$$= \left(1/\sqrt{2^n}\right) \times (|1> + |0>) \otimes (|1> + |0>) \otimes \ldots \otimes (|1> + |0>)$$

$$= \left(1/\sqrt{2^n}\right) \times (|11 \ldots 1> + |11 \ldots 0> + \ldots + |00 \ldots 0>)$$

Note that applying a linear (n-1) number of operations to the qureg yields a register state which is a superposition of an exponential (2^n) number of terms. This exponential performance is one crucial characteristic of the potential power of quantum computation which distinguishes it from other unconventional computational models, such as DNA computation, swarm computation or evolutionary computation, that scale linearly with N, the number of molecules, particles or entities in the system [7].

The prepared register state is input into a quantum logic network constructed from quantum logic gates. Two important kinds of gate are the one-qubit unitary gates, such as the Hadamard, Pauli, and phase-shift gates, and the two-qubit controlled-unitary gates, such as the controlled-NOT (C-NOT) gate. The quantum logic network is designed to carry out a specific quantum computation by unitarily evolving the input state into the output state, which is a coherent linear superposition of all the possible outputs from that computation. It is then necessary to extract the single required component from the output state, and this is generally achieved by means of a process, referred to as controlled decoherence, that repeatedly enhances the amplitude of this component at the expense of all of the remaining components' amplitudes. It is then highly probable (although not absolutely certain) that an observation or measurement of this state will yield the required component.

Note that it is not fruitful to attempt to directly observe the superposed output state prior to controlled decoherence since the act of measurement will cause the state to decohere spontaneously (or 'collapse') to a single randomly selected component with

a consequential loss of information. Note also that as a consequence of the no-cloning theorem [8, 9] it is not possible to clone an exact copy of an arbitrary unknown quantum state for preservation and subsequent analysis, although approximate copies of quantum states can be made.

3 Quantum Forensics

The topic of quantum forensics has so far received virtually no consideration from researchers [10]. In order to understand the possibilities for digital forensic examination and analysis of quantum computers, it is convenient to employ the conventional division into *in vivo* (live) forensics and *post mortem* forensics. From what has already been said above it will be clear that it is not possible to perform live system forensics on a quantum computer, since any observation or measurement made on an evolving superposed state will cause it to collapse to a single randomly selected component. At that instant the states of the system and the observer or measuring apparatus in their environment become correlated (or 'entangled') and useful state information leaks from the system into the composite super-system from which it cannot be recovered.

At first sight, *post mortem* forensics offers a somewhat more encouraging prospect. After the termination of a quantum algorithm and the recovery of an output state generated by controlled decoherence, a single classical output may remain for conventional digital forensic recovery and analysis. Of course, conventional anti-forensics techniques may be used to remove this trace. However, since there is no possibility of observing the intermediate (non-classical) states of the quantum computer *post mortem*, it will not be possible to reconstruct a timeline for the evolution of the quantum computation. This restriction places quite severe limitations on both the quantity and the quality of the evidential traces that may ultimately be recoverable. Since the probative value of digital evidence depends to a considerable extent upon the contextual support it receives from ancillary evidence, it would appear to be very difficult to build a judicially convincing case from the available digital forensic evidence alone.

At first sight, therefore, it would appear that with the advent of practical quantum computers the task of cyber-law enforcement will become significantly more challenging. Not only will quantum cyber-criminals be able to gain access to devices capable of breaking a variety of both symmetric and asymmetric state-of-the-art cryptosystems well within any realistic cover time, but their activities will also leave behind a minimal set of recoverable forensic traces for use in any subsequent prosecution. It is this doubly challenging prospect that has given rise to the present study.

A road-map for future digital forensic investigations of quantum cyber-crimes should therefore focus attention on the maximum amount of information that can be elicited from the recoverable evidence, which could potentially amount to just a single trace. If it could be demonstrated with high probability that the *only* feasible route leading to the creation of a single recovered evidential trace necessarily involved the unauthorized perpetration of the cryptanalytic process as alleged by the prosecution, then it might indeed be possible to mount a plausible judicial case.

In fact, recent research in this area, using the actual data from real-world criminal cases involving peer-to-peer file-sharing [11] and online auction fraud [12] respectively, has demonstrated that the application of computational complexity theory [13] and the Keystroke-Level Model (KLM) [14] within the framework of an Operational Complexity Model (OCM) [15], together with the use of Bayesian inference networks for reasoning about digital evidence [11, 12], does indeed offer a realistic prospect of quantifying the posterior odds and the likelihood ratios of competing hypotheses concerning the feasible routes leading to the formation of the recovered digital evidential traces [16]. In particular, by enumerating and analyzing all feasible routes j by which the recovered evidential traces could have been produced, the likelihood ratio of feasible route i is given by:

$$\Lambda_i = \Pr(E|H_i) \,/\, \Sigma_{j \neq i} \Pr(E|H_j)$$

and the posterior odds of feasible route i is given by:

$$O_i = \Pr(H_i|E) \,/\, \Sigma_{j \neq i} \Pr(H_j|E)$$

where $\Pr(E|H_i)$ is the probability that evidential trace E is found given that the hypothesis that feasible route i was taken is correct, and $\Pr(H_i|E)$ is the probability that the hypothesis that feasible route i was taken is correct given that evidential trace E was found.

4 Summary and Conclusions

We have examined how the operation of a quantum computer in the hands of cyber-criminals attempting to break a state-of-the-art public-key or private-key cipher may result in the creation and recovery of one or more digital forensic evidential traces. However, very the nature of the quantum computational process precludes the creation and recovery of any useful information concerning the intermediate states in the evolution of the quantum computation. Furthermore, it also precludes the possibility of live forensics *per se*. It is probable that the lack of this ancillary evidence will prove to be a significant barrier to mounting successful judicial prosecutions in cases of serious quantum cyber-crime. However, recent and ongoing developments in quantifying the probative value of digital forensic evidence may be able to offer at least a partial resolution of this problem. Further detailed research in this area would therefore be well motivated and of great potential value.

References

1. Reich, E.S.: Quantum Computers Move a Step Closer. Nature 467, 513 (2010)
2. Shor, P.W.: Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In: Proc. 35th Annual Symposium on Foundations of Computer Science, pp. 124–134. IEEE Press, Los Alamitos (1994); quant-ph/9508027
3. Shor, P.W.: Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM J. Comput. 26, 1484–1509 (1997)

4. Grover, L.K.: A Fast Quantum Mechanical Algorithm for Database Search. In: Proc. 28th Annual ACM Symposium on the Theory of Computing, pp. 212–219 (May 1996); quant-ph/9605043
5. Grover, L.K.: Quantum Mechanics helps in Searching for a Needle in a Haystack. Phys. Rev. Letters 79, 325–328 (1997)
6. Mermin, N.D.: Quantum Computer Science: an Introduction. Cambridge University Press, Cambridge (2007)
7. Overill, R.E.: Parallelism of Paradigms in Non-Classical Computation. In: Stepney, S. (ed.) Proc. Intl. Workshop: The Grand Challenge in Non-Classical Computation, York, UK (April 2005)
8. Wootters, W.K., Zurek, W.H.: A Single Quantum Cannot be Cloned. Nature 299, 802–803 (1982)
9. Dieks, D.: Communication by EPR devices. Physics Letters A 92(6), 271–272 (1982)
10. Konstadopoulou, A.: From Quantum Security to Quantum Forensics. In: Proc. 1st Conference on Advances in Computer Security and Forensics (ACSF), Liverpool, UK, pp. 105–111 (July 2006)
11. Kwan, M., Chow, K.-P., Law, F., Lai, P.: Reasoning about Evidence using Bayesian Networks. In: Proc. 4th Annual IFIP WG 11.9 International Conference on Digital Forensics, Kyoto, Japan (January 2008); Advances in Digital Forensics IV, ch. 12, pp. 141–155. Springer, Heidelberg (2008)
12. Kwan, Y.K., Overill, R.E., Chow, K.-P., Silomon, J.A.M., Tse, H., Law, Y.W., Lai, K.Y.: Evaluation of Evidence in Internet Auction Fraud Investigations. In: Proc. 6th Annual IFIP WG 11.9 International Conference on Digital Forensics, Hong Kong (January 2010); Advances in Digital Forensics VI, ch. 7, pp. 95–106. Springer, Heidelberg (2010)
13. Papadimitriou, C.H.: Computational Complexity. Addison-Wesley, Reading (1994)
14. Kieras, D.: Using the Keystroke-Level Model to Estimate Execution Times, University of Michigan (2001), `http://www.cs.loyola.edu/~lawrie/CS774/S06/homework/klm.pdf`
15. Overill, R.E., Silomon, J.A.M., Chow, K.-P.: A Complexity Based Model for Quantifying Forensic Evidential Probabilities. In: Proc. 3rd International Workshop on Digital Forensics (WSDF 2010), Krakow, Poland, pp. 671–676 (February 2010)
16. Overill, R.E., Silomon, J.A.M.: Digital Meta-Forensics: Quantifying the Investigation. In: Proc. 4th International Conference on Cybercrime Forensics Education & Training (CFET 2010), Canterbury, UK (September 2010)

Recovery Techniques for Deleted Email Items in Digital Forensic Context

Chorong Jeong, Keun-gi Lee, and Sangjin Lee

Center for Information Security Technologies,
Digital Forensic Research Center,
Korea University, Seoul, Korea
{cutycom_,lifetop,sangjin}@korea.ac.kr

Abstract. Corporations use email as their primary method for internal communication and business processes. By their nature, the email messages are in general used for major business processes that contain large amounts of business information. Technological improvements also enable an email system to handle large amount of data. When there is a criminal activity like Technology leakage, an email message can be highly likely to become critical evidence; however, there is a high chance that a suspect will intentionally delete the message. Thus, the ability to recover deleted email will be very critical to prove the wrongdoing in this case. This paper analyzes how various email client applications adopt their deletion systems and any possible recovery methods are introduced accordingly.

Keywords: email; email forensics; recovery; recovery email.

1 Introduction

Email is a communication method for individuals, but is also a core of major business processes. Email messages are used to enable smooth running of the business with employees. So, a lot of business information is preserved as Emails.

Email can be used in corporate crimes, such as Technology leakage. In the investigation of Technology leakage, emails are subjected to forensic investigation. Generally, an email is stored the common information, such as sender and receiver, attached files, date and time, references, etc. Email forensics refers to the collection, extraction and analysis, restoration of the emails related to cyber crime. Based on the common information extracted from the seized high capacity email, the correlation with the event and the correlation analysis between the sender and the receiver can be investigated. There are two methods of deleting an email message: deleting it from the archive's recycling bin, and deleting it using the Shift-Delete key combination (deletion method without using recycle bin in each email client). Though these methods of deletion are different, with the exception of several specific files the information in the email deleted at an unallocated area remain as it is. However, there are difficulties in analyzing the unallocated area, because in most cases, the data cannot be kept in perfect condition [1].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 115–122, 2011.

When investigating corporate crimes, such as Technology leakage, there is a high possibility that the suspects have already deleted the emails related to the event. Therefore, restoring deleted emails is helpful to understanding the overall flow of the event, and because the major clues for the event can be provided through them, it is very critical.

To recover deleted emails, email recovery tools are used. Commonly, specific tools exist for each format such as DBX recovery tool, PST recovery tool. therefore these tools are difficult to analyze different file formats. Also, recovery tools are not performed perfectly. so a new recovery method needs to be sought.

This paper explains the structure and the recovery method of deleted email messages within the email files used at Mozilla Thunderbird (MBOX), Outlook Express (DBX), and Microsoft Office Outlook (PST, OST).

2 Structure and Recovery Method of Deleted Files

In June of 2009, Campaign Monitor gave a questionnaire survey to 30 thousand customers on their current email clients for 6 months. Fig. 1 shows the results [2].

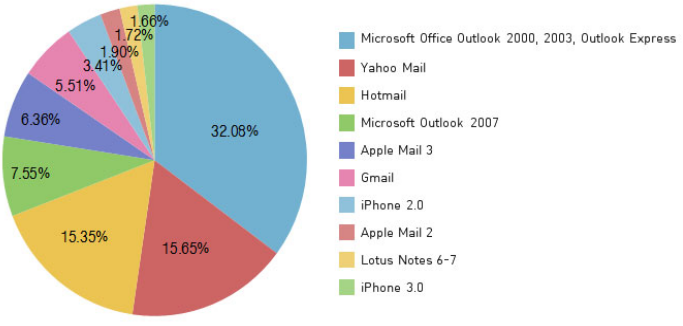


Fig. 1. Usage ratio for personal email client

The clients that most people use are Outlook 2000, 2003 and Express, and if web mail usage is excluded, Outlook 2007 is the next most used email client. In addition to Outlook, Lotus Note 6-7 is also being used by corporations. In this section, we discuss the methods that can be used for the recovery of deleted emails in Outlook, Outlook Express, IBM Lotus Notes, Mozilla Thunderbird and Windows Live Mail.

Of these, the files used in Windows Live Mail (EML) and IBM Lotus Notes (NSF) are not recoverable. The reasons are as follows.

The file format used by Windows Live Mail is EML. EML is a MIME format. This file saves a single email at a single file. Deleting an email means deleting the EML file, so the email deleted in the file no longer exists.

IBM Lotus Notes is groupware that used by the company. its email is stored in NSF file format. This file stores all emails in folders, such as Inbox, Sent, Trash, etc. NSF file stands for Notes Storage File and saves information such as electronic

business documents, schedules, contacts, and email information based on Lotus Notes groupware and also saves files created based on application developed by the user [3].

While the server is running, each Domino DB file in the data folder of Domino server is saved in NSF format. However, when the email is deleted, it gets overwritten with 0xAA, etc. Therefore, deleted email messages in NSF are not recoverable.

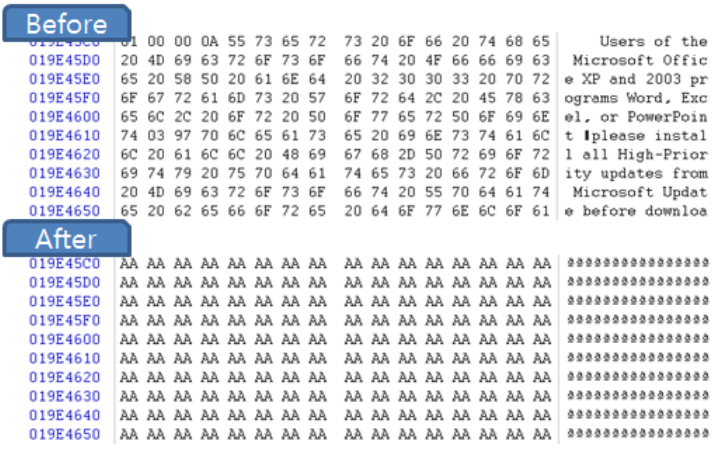


Fig. 2. Email messages and Deleted email messages in NSF File

2.1 Mozilla Thunderbird (MBOX)

Thunderbird is an email client provided by Mozilla. An email message in Thunderbird consists of a file pair comprised of a file without a file extension and an .msf file. Files without file extensions compose the email in MBOX format, and several emails are continuously stored.

The .msf files are indexing files to manage emails within a file not having a file extension.

Table 1. MBOX File format and .msf File format example

Example(MBOX File format)	Example(.msf File format)
From - Mon Feb 21 20:09:06 2011 X-Account-Key: account3 X-UIDL: 1297918244525.42499.koreaLINK,S =1297918244525.4249900000 X-Mozilla-Status: 0009 X-Mozilla-Status2: 00000000 X-Mozilla-Keys: ...	// <!-- <mdb:mork:z v="1.4"/> --> < <(a=c)> // (f=iso-8859-1) (B8=numHdrsToKeep)(B9=daysToKeepBodie s)(BA=keepUnreadOnly) (BB=useServerDefaults)(BC=cleanupBodies)(BD=applyToFlaggedMessages) (BE=LastPurgeTime)(BF=account)(C0=dateRe ceived)(C1=ProtoThreadFlags) ...

email files are created and are seperated by folder, such as Inbox, Sent, etc.

To recover deleted emails in Thunderbird, the MBOX files must be extracted. Because the MBOX files consist of text, if they are continuously extracted in the unit of MIME format, the deleted emails are easily recoverable. However, it cannot be verified whether the email messages are deleted or undeleted with this kind of file alone. To verifiy this, an additional analaysis of the .msf file is necessary.

2.2 Microsoft Outlook Express (DBX)

Microsoft Outlook Express is an email client included in Windows OS. It can be used in Windows 95, 98, and XP without installing Microsoft Office. The file format in which emails are saved in Outlook Express is DBX. DBX exists as a single file per folder, with file names such as Inbox, Sent, Trash, Drafts, and Unsent.

Because the deleted email information at DBX file remains in the deleted header, to recover the deleted emails, the deleted message list root pointer and the deleted tree list root pointer existing at the file header must be referred[4].

Table 2. A Part of DBX file Header

Index	Type	Description
...	...	...
0x48	int4	Root pointer to the deleted message list
0x4C	int4	Root pointer to the deleted tree list
...	...	...

If pointer values described at the Table 2 is exist, this means that there is a deleted email. If the index moves to the pointer with a value of 0x48, the deleted object and the related offset and structure appear as follows[5].

Table 3. Deleted object header and body format (Deleted email message)

The Header		
Position	Type	Description
(1)	int4	Object marker
(2)	int4	Length of field (6)
(3)	int4	Length of this object, (2) + 0x14
(4)	int4	Not used, not cleared
(5)	int4	Pointer to the next deleted object
The Body		
Position	Type	Description
(6)	data	Free space, not cleared, length (2)

The data structure above consists of the deleted object, the current offset and object size as relevant information, and the offset and data values of the deleted object. When this structure is viewed as a binary structure of the actual file, it shows as Fig. 3.

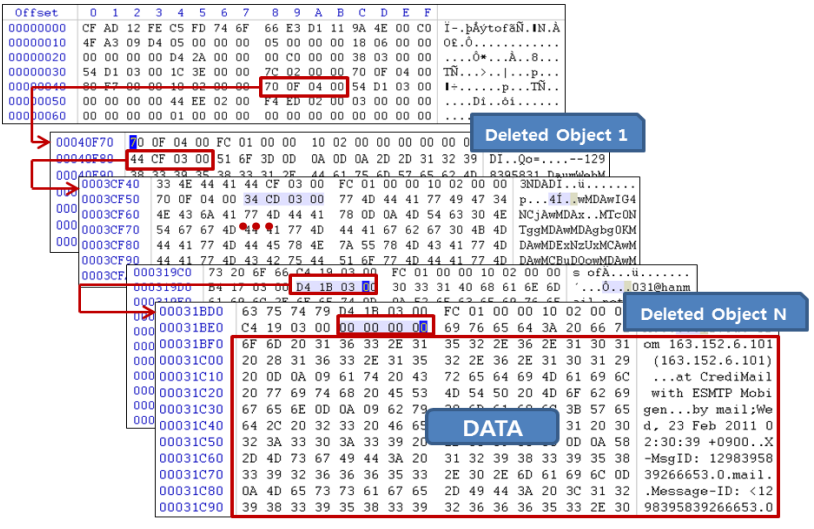


Fig. 3. The process of searching for deleted email messages in DBX

The 0x48 offset value of the file header is 0x00040F70. At this location, the first deleted object of the root list can be seen. After verifying the header 0x14 bytes of the deleted object structure, the 0x1FC bytes data below are read. These data are the information of the deleted email, and normally are the context of the deleted email. After reading the data, the pointer of the next deleted object is read at the structure head, and the location moves to the next corresponding value. This procedure proceeds as much as the deleted list size.

At the deleted list within the DBX, the recently deleted emails are recoverable, and because the complete email information is not obtainable, a partial email recovery such as email context is possible.

2.3 Microsoft Office Outlook (PST, OST)

Microsoft Outlook is an email client that is part of the Microsoft Office suite, and the file formats for saving email messages are PST and OST. PST (Personal Storage Table) is a file to store the message and other items of the computer, and OST (Offline Storage Table) is a file that enables the user to work offline when not connected to Exchange Sever by storing a copy at a local location.

A PST file consists of NBT (Node B-Tree) and BBT (Block B-Tree), and NBT is indexed to be managed for efficient Node search. BBT manages the Nodes to store the actual data of email[6].

In order to restore deleted email at PST, the following analysis is necessary.

2.3.1 AMAP, PMAP Analysis

In order to approach the deleted area, the allocated information within the PST files must first understanding. The AMAP (Allocation MAP Page) and the PMAP (Page MAP) are analyzed.

AMAP stands for Allocation MAP Page. The first offset of AMAP is 0x4400, and its unit is 512 bytes, which is the unit for a page. 1 bit of AMAP is mapped with a block of 512 bytes, and if it is 1, it means allocation, and if it is 0, it means unallocated area. One page shows a data section of 253,952 bytes, so AMAPs are placed at 253,952 byte intervals[6,7].

PMAP stands for Page Map. The first offset of PMAP is 0x4600, and consists of page units of 512 bytes. 1 bit of PMAP is mapped with a page of 512 bytes, and if it is 1, it means allocation, and if it is 0, it means unallocated area. Because the PMAP of a page shows a data section of 2,031,616 bytes, PMAPs are placed at 253,952 byte intervals[6,7].

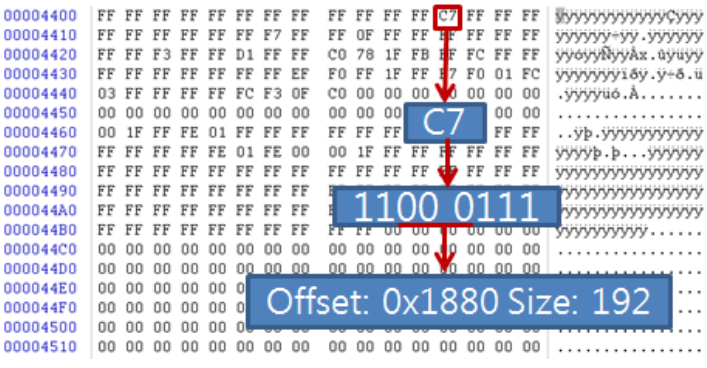


Fig. 4. The starting offset and the size of the unallocated area at AMAP

In order to understand the unallocated area, information of 1-bits saved as 0 at AMAP, and PMAP is necessary. Based on all AMAP and PMAP information, by only extracting bits with 0 value and saving the start offset and size, the respective unallocated lists of AMAP and PMAP are configured.

2.3.2 Creation B-Tree Consisted of Email to be Restored

In order to recover the deleted emails, BTPAGE, which are the elements of B-Tree, must be searched. They can be searched with the unallocated list of AMAP and PMAP that has been configured in advance. Because BTPAGE is basically configured with 512 bytes, the page of B-Tree is mostly searched using the offset and the size of the unallocated list of PMAP. In this process, it must be checked whether the offset and the size of the unallocated list exist at the effective range of the data section of actual file, and when they are included in this range, they are inserted into the B-Tree. If BTPAGE is Leaf page, a Leaf page refers to another new block or TC, PC. TC is saved the attribute data. And PC is saved the actual email data.

Also, if the block size is bigger than 512 bytes and is a multiple of 512 bytes, the possibility that it is a table where the email message's attribute data are saved or a

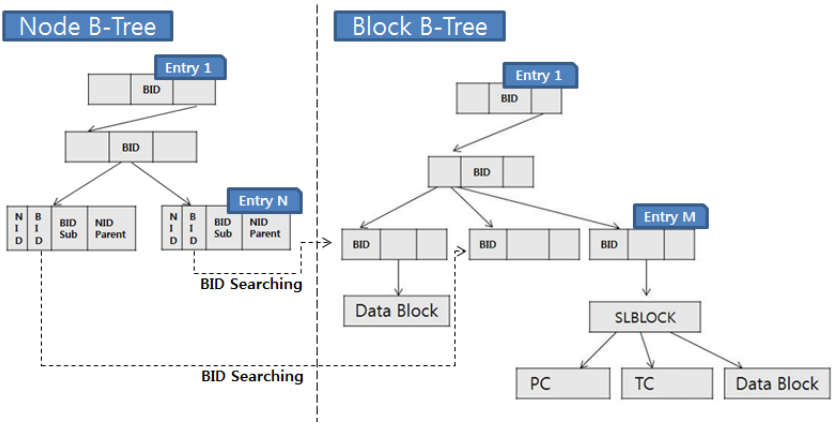


Fig. 5. NBT and BBT Composition

table where the actual data are saved is higher than the possibility that it is a page to compose a B-Tree. Therefore, these data offsets must be also recognized.

2.3.3 Deleted Email Item Extraction Based on B-Tree Information

If the B-Tree configuration is complete, the deleted emails must be extracted. The entry nodes saved at the Leaf page of BBT normally store the offset of TC(or PC) of a SubBlock. In other words, this offset refers to the TC or the PC storing information, such as email title, sender, receiver, receiving time, and mail context of the deleted

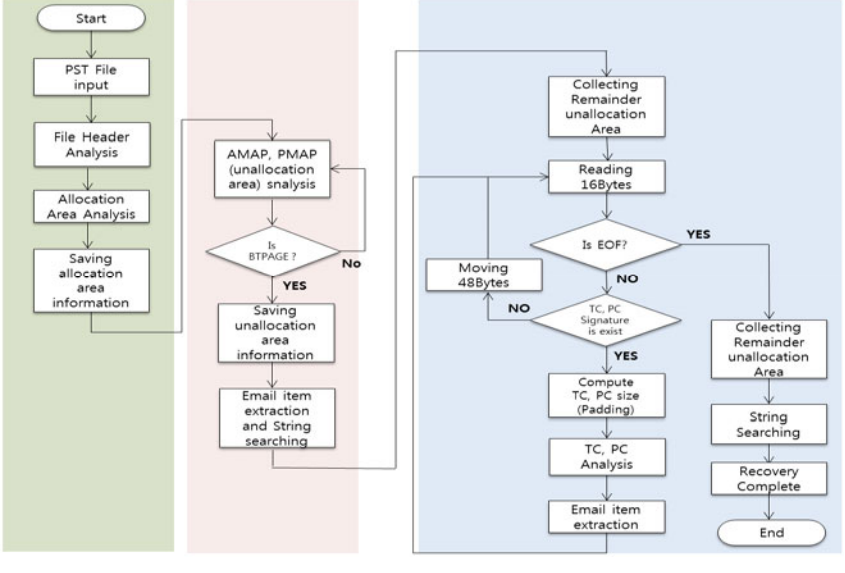


Fig. 6. Deleted email recovery algorithm in PST file

email. If it is an entry node saved at a Leaf page of NBT, the entry node of the leaf page of the corresponding BBT is found by searching an entry node having a unique ID, the NID value that each entry node has. If an entry node is searched, the subsequent procedure is the same.

Therefore, to extract as much information as possible from a deleted email existing within a file, the carving method utilizing the signature of tables saving the email data must be adopted to the recovery algorithm. First, after creating NBT and BBT managing the undeleted emails existing within a PST file, TC(or PC) corresponding to the offset existing at the tree should not be extracted, and only other TC(or PC) should be extracted. The method of extracting the TC(or PC) by carving is an email recovery method that represents an improvement on the existing algorithm.

3 Conclusion and Future work

In this paper shows that deleted email messages from various email files are recoverable. In addition, various recovery methods for each email file are explained.

For Thunderbird, a method of extracting the files by parsing it as a text file is given, while for Outlook Express, the structure of the deleted list exists at the file format, so recovery was possible but only in a partial area. Finally, for Microsoft Office Outlook, after collecting information based on the AMAP and PMAP which are complicated but have the allocation information, and the deleted email item can be extracted using the same method.

To verify our methodology described in this paper, we will develop integrated email analysis software for digital forensics which is supported recovery deleted email items in various email file format.

Acknowledgments. This research was supported by Bio R&D program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (20100020634).

References

1. Yoo, B., Park, J., Bang, J., Lee, S.: A Study on a Carving Method for Deleted NTFS Compressed Files. In: F2GC 2010, vol. 3(1), Future Technology Research Association International, Philippines (2010)
2. Campaign monitor, http://www.campaignmonitor.com/stats/email-clients/
3. Lotus Notes, http://www-01.ibm.com/software/lotus/products/notes/
4. Schloh, A.: OE dbx file format: 'file header', http://OE-Dbx-FileHeader.html in http://oedbx.aroh.de/zip/oedbx.zip/
5. Schloh, A.: OE dbx file format : 'deleted', http://OE-Dbx-Deleted.html in http://oedbx.aroh.de/zip/oedbx.zip/
6. Microsoft: [MS-PST] Outlook Personal Folders File Format (.pst) Structure Specification (2010)
7. Metz, J.: Personal Folder File (PFF) file format specification (2009)

Digital Evidence Discovery and Knowledge Management Issues Concerning Multimedia Computing Devices Utilizing GPS Navigation Services and Social Network Activities

Hai-Cheng Chu[1], K.H. Chang[2], Yi-Da Wang[2], and Jong Hyuk Park[3,*]

[1] Department of International Business,
National Taichung University of Education, 140 Min-Shen Road,
Taichung 40306 Taiwan ROC
[2] Department of International Business,
Tunghai University, No.181, Sec. 3, Taichung Port Rd., Taichung 40704, Taiwan ROC
[3] Department of Computer Science and Engineering,
Seoul National University of Technology, 172 Gongneung-dong 2, Nowon-gu 139743, Korea
hcchu@mail.ntcu.edu.tw, kuo1@thu.edu.tw, rran456@gmail.com,
parkjonghyuk1@hotmail.com

Abstract. Mobile communication devices are commonly adopted in contemporary Internet Communication Technology (ICT) era with unparalleled evolving pace. Regrettably, ICT related cybercrimes are rising in an unprecedented speed. Unquestionably, the detection and prevention of the associate information security incidents becomes much more urgent and unarguable both in public sectors as well as private sectors. In this paper, the utilization of a PDA for GPS navigation and Wi-Fi connection to ubiquitous social networking service provider – Facebook are cross examined concerning the related digital evidence collecting and discovering concerning ubiquitous networks and wireless communications, which provides a framework for the associate digital forensics.

Keywords: mobile device digital forensics, global position system, volatile memory acquisition, Facebook.

1 Introduction

Proverbially, as mobile computing devices (PDA, smart phones, and GPS navigation devices, etc.) become cost competitive and the prevalence of ubiquitous computing infrastructures, the information being deposited in mobile computing devices could be much more precious than those of desktop PC. In the meanwhile, plethora of cybercrime arises accompanied with incomparable progresses of ICT. Mobile communication apparatus will contain some kind of substantive digital trails once has been utilized. In a nutshell, digital evidences could be disclosed in a systematic approach, which could be phenomenally critical in cracking down some cyber

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 123–130, 2011.

crime-related incident. Astonishingly, certain illicit cybercrime syndicates extensively exploit ICT as an avenue to achieve lucrative purposes. All of the aforementioned information security leakages are occurring matchlessly. Consequently, in this paper, we provide a framework of digital evidence collection, analysis, and data mining in case upon receiving a PDA with embedded GPS and Wi-Fi functions. The research discloses the potential pervious GPS routes that indicate the current user might have physically presented at those geographic locations. Also, we would like to take this opportunity to reveal possible Facebook accounts that might be using on the current PDA when the issue of industrial espionage is conducting via Facebook.

By the virtue of digital forensics technology, the spotted PDA with GPS navigation program shut down and Facebook session logout, we still might be able to disclose some digital trails to indicate some routes that the user had been taken / travelled and the Facebook user account(s) on the associate PDA before it runs out of power. Demonstrably, due to the nature of volatile memory, the digital content of the Random Access Memory (RAM) would be disappeared once the battery is drained. Consequently, the paper provides the paradigm as references for practitioners in public sectors or private sectors.

2 Literature Review

Nowadays, due to the innumerable and metamorphic of cybercrimes, Digital Forensics has been extensively utilized to examine cybercriminal cases in both public and private sectors. Substantively, the physical memory of the mobile computing devices is the ephemeral storage region for the operating system and application programs executing on the mobile devices [1, 5, 6, 12, 13]. Consequently, the acquisition and forensics analysis of physical memory of the mobile computing devices play an essential role concerning traditional digital forensics methodologies. Alternatively to state, the RAM of the computing device is the volatile memory, which need additional care regarding the digital evidence collection. When the mobile computing device is powered off for impounding, all the precious information within the volatile memory will be evaporated forever. Facebook is considered to be one of the leading social network service providers, which requires participants to register to Facebook server without the installing procedures. Unfortunately, Facebook results in numerous information security leakages via using the platform. According to U.S juridical systems, the authentication of electronically sored information becomes critical when pursuing the reality in resolving civil disagreements concerning small scale digital devices (cellular phones, smart phones, or PDAs). This stresses the importance of digital forensics for portable electronic communication devices [2, 3, 7, 14]. Generally speaking, a PDA can be synchronized with a desktop PC using synchronization protocols such as *Microsoft ActiveSync*.

For PDAs forensics, basically, there are two categories of data acquisition methods: physical acquisition and logical acquisition. For the former one, the digital forensics expert conducts a bit-stream copy of an entire RAM. For the later one, the digital forensics staff obtains a bit-stream copy of the file directory and files. Evidently, the acquisition of the internal memory could be the most challenging task during the investigation procedures in some cases, which encompasses ROM, RAM and the flash memory [4, 8, 9, 10, 11, 15]. Indeed, the RAM, which is volatile in its nature, stores the executing processes and the related user data in a way that the data will be permanently lost if the battery went down. Invariably, the exploitation of PDAs could be the media or

avenue to commit cybercrimes. Consequently, the forensics of PDAs becomes another newly emerged and hotly debated research field. According to the recently related researches, they indicated that only a few technologies or products are able to perform forensic acquisition on *Pocket PC* platform [9, 15].

3 Volatile Memory Acquisition and the Corresponding Digital Forensics - A Case Review

In this research, the experiment is conducted based on the premise that there is a logistics carrier of a certain Supply Chain Management (SCM) issues a PDA with a Bluetooth GPS receiver accompanied with a navigation software package for every individual driver for daily dispatching services. Specifically, for the purpose of just-in-time delivery, all transportation carriers are equipped with GPS navigation suite, *Papago Ver. G10* [17], which is a popular satellite navigation application program in Taiwan. Based on historical digital records and several digital breadcrumb evidences, the Chief Information Security Officer (CISO) of the logistics carriers highly suspect an industrial espionage incident is occurring within the corporate by a certain driver. Hence, the CISO decides to secretly conduct a digital evidence collection procedures right after the driver completed the delivery task. The information security team took over the PDA without shutting off the mobile device.

Phrase 1

The team applied *Paraben's PDA Seizure Ver. 3.0.3.93 and Microsoft ActiveSync Ver. 4.5* to obtain the image file of the RAM of the current PDA as Fig. 1 shown.

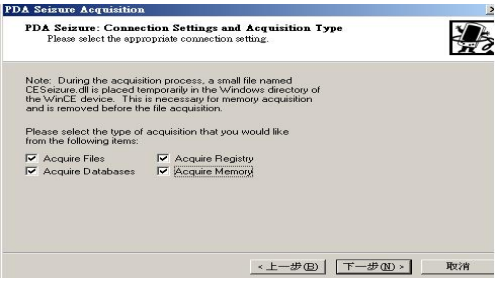


Fig. 1. The team tried to obtain the image file of the RAM of the current PDA

From Fig. 1, the team retrieved the registration code of the GPS navigation suite, *Papago Ver. G10*, *63627541*, as Fig. 2 indicated. The team confirmed that the code is the official authorization code owned by the corporate and no other GPS navigation software packages are installed on the current PDA.

Phrase 2

The team found the PDA has automatically kept records of the navigation routes. Hence, the team saved the records into a file, *092701.TR7*, transforming it into an importable file for Google Map via the conversion web site [16]. The web site is capable of converting **.TR7* files to **.KMZ* and **.GPX* files that can be recognizable in *Google Earth* and *Google Map*.

Phrase 3

The team inspected the transformed GPX file and found the latitude, longitude, and the associate time stamps for some journeys. After scrutinizing the digital data, the team

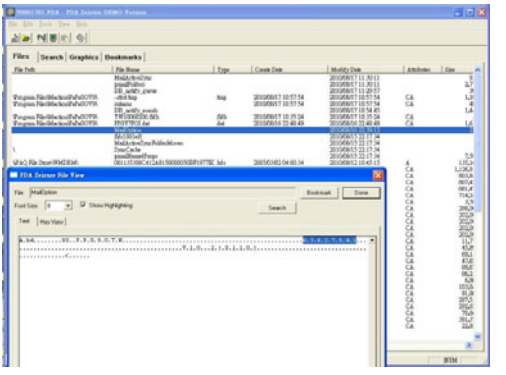


Fig. 2. The registration code of the GPS navigation suite, *63627541*, was disclosed

can initially conclude that the driver had been travelling via three distinct routes. In the meanwhile, the team uploaded the transformed *KMZ* file to *Google Earth* and obtained the corresponding expected results, which means that there are three different routes, which the driver had actually taken. The 1st route on *Google Map* and the digital records as indicated in Fig. 3, Fig. 4 and Fig. 5, respectively. Fig. 3 demonstrated that the driver started as the arrow pointed based on *latitude (24.18351)*, *longitude (120.613515)*, and the *time stamp (2010-09-24 04:44:57 UTC)*. Furthermore, the driver arrived as the arrow pointed based on *latitude (24.081678)*, *longitude (120.542058)*, and the *time stamp (2010-09-24 05:24:57 UTC)*. Fig. 4 and Fig. 5 support the above digital evidence interpretation. Similarly, the 2nd route on *Google Map* and the digital records as indicated in Fig. 6, Fig. 7 and Fig. 8, correspondingly. Alternatively, the 3rd route on Google Map and the digital records as indicated in Fig. 9, Fig. 10 and Fig. 11, individually.

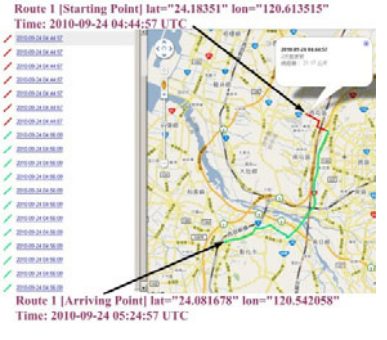


Fig. 3. The 1st route shown on *Google Map* with the starting and arriving *latitudes, longitudes, and* the *time stamps*

Fig. 4. The 1st route shown on Google Map indicating the latitude, longitude, and the time stamp of the starting point

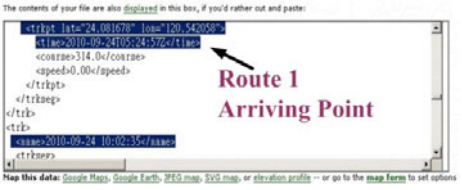


Fig. 5. The 1st route shown on *Google Map* indicating the *latitudes, longitude, and* the *time stamps* of the arriving point

Fig. 6. The 2nd route shown on Google Map with the starting and arriving latitudes, longitudes, and the time stamps shown on Google Map

Fig. 7. The 2nd route shown on *Google Map* indicating the *latitude, longitude, and* the *time stamp* of the starting point

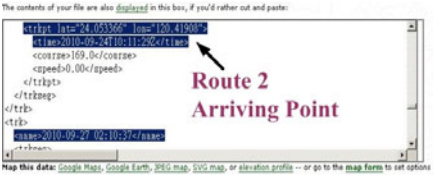


Fig. 8. The 2nd route shown on *Google Map* indicating the *latitude, longitude, and* the *time stamp* of the arriving point

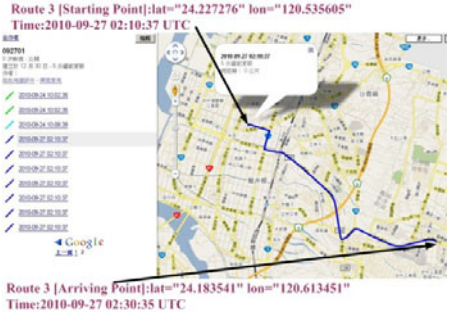


Fig. 9. The 3rd route shown on *Google Map* the starting and arriving *latitudes,longitudes, and* the *time stamps* shown on *Google Map*

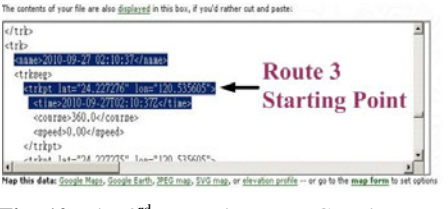


Fig. 10. The 3rd route shown on Google Map indicating the latitude, longitude, and the time stamp of the starting point

Fig. 11. The 3rd route shown on *Google* indicating the *latitude, longitude, and* the *time stamp* of the arriving point

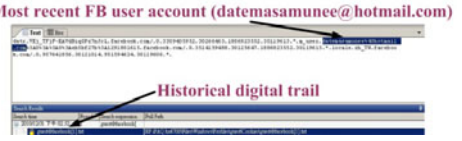


Fig. 12. The most recent Facebook user Map account, datemasamunee@hotmail.com via applying the search keyword, guest @facebook[

Phrase 4

As some employees filed complaints regarding the driver heavily surfs on Facebook during regular office hour for leisure purposes and the CISO of the logistics corporation feels the confidential information might accidentally reveal through this most popular social network channel nowadays. Hence, the CISO decides to disclose previous Facebook users on this current PDA, which is the asset of the corporate.

Phrase 5

Moreover, the team applied the search keyword, *guest@facebook[*, targeting on the image file of the RAM of the PDA and obtained the most recent Facebook user account, *datemasamunee@hotmail.com*, on the current mobile communication device as Fig. 12 demonstrated.

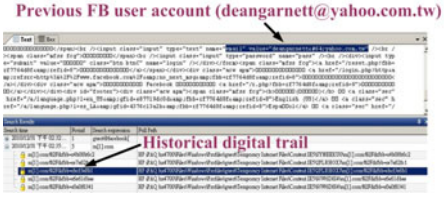


Fig. 13. Via applying the search keyword, *m[1].com*, the previous Facebook user, *deangarnett@yahoo.com.tw* was disclosed as the arrow pointed

Fig. 14. Via applying the search keyword, m[1].com, the previous Facebook user, lambert_ethan@yahoo.com.tw was disclosed as the arrow pointed

Fig. 15. Via applying the search keyword, *m[1].com*, the previous Facebook user, *rran456@yahoo.com.tw* was disclosed as the arrow pointed

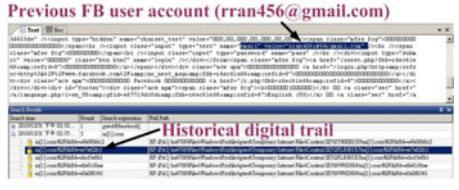


Fig. 16. Via applying the search keyword, m[1].com, the previous Facebook user, rran456@gmail.com was disclosed as the arrow pointed

Phrase 6

Additionally, the team employed the search keyword, *m[1].com*, for the purpose of disclosing the possible other previous Facebook users on this PDA. As Fig. 13~16 illustrated, there were other four previous Facebook user logon on the PDA. The CISO can utilize the obtained Facebook accounts to explore the social network activities via the methodology presented above to discover previous Facebook social network activities, which would not be the focus of this research paper.

4 Summary and Review of the Corresponding Digital Forensics Results

Based on the digital trails being collected, analyzed, and reported, the digital forensics team summarizes the above digital trails as Fig. 17 depicted, which would be the essence of the paper concerning dada mining of the digital breadcrumb trails.

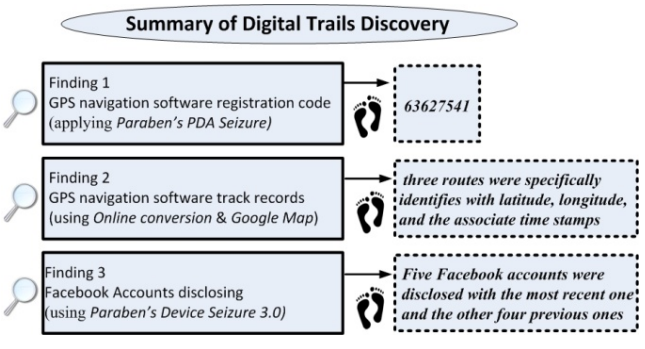


Fig. 17. Summary of the collected digital evidences and the corresponding discovery

5 Conclusion

The digital content within volatile memory, RAM, will be vanished forever one the battery drain. Hence, precious digital evidences need to be collected in time in order to avoid irreversible collecting procedures. In this paper, the PDA is crossed examined via the image of the RAM of the PDA. The digital track file, **.TR7*, was transformed to **.GPX* file, which can be diagnosed concerning latitudes, longitudes, and the time stamps for the distinct trips. Additionally, the Google Map specific demonstrates the three individual trips that the driver of the carrier had taken. Moreover, the most recent and previous Facebook account was disclosed for the purposes of cracking the potential undergoing industrial espionage.

Acknowledgment

The author would like to acknowledge the funding support of NSC (National Science Council) of Taiwan concerning the grant of Project NSC 99-2221-E-142-009.

References

1. Strawn, C.: Expanding The Potential for GPS Evidence Acquisition. Small Scale Digital Device Forensics Journal 3(1), 1–12 (2009)
2. Iqbal, M.U., Samsung, L.: Legal and ethical implications of GPS vulnerabilities. Journal of International Commercial Law and Technology 3(3), 178–187 (2008)
3. Sohn, H., Ro, Y.M., Plataniotis, K.N.: Content Sharing Based on Personal Information in Virtually Secured Space. In: Kim, H.-J., Katzenbeisser, S., Ho, A.T.S. (eds.) IWDW 2008. LNCS, vol. 5450, pp. 388–400. Springer, Heidelberg (2009)
4. Acquisti, A., Gross, R.: Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In: Danezis, G., Golle, P. (eds.) PET 2006. LNCS, vol. 4258, pp. 36–58. Springer, Heidelberg (2006)
5. Brown, R., Pham, B., de Vel, O.: Design of a Digital Forensics Image Mining System. In: Khosla, R., Howlett, R.J., Jain, L.C. (eds.) KES 2005. LNCS (LNAI), vol. 3683, pp. 395–404. Springer, Heidelberg (2005)
6. Chen, H., Chung, W., Xu, J., Wang, G., Qin, Y., Chau, M.: Crime Data Mining: A General Framework and Some Examples. IEEE Computer 37(4), 50–56 (2004)
7. Asher, C., Aumasson, J.-P., Phan, R.C.-W.: Security and Privacy Preservation in Human-Involved Networks. In: Camenisch, J., Kesdogan, D. (eds.) iNetSec 2009. IFIP AICT, vol. 309, pp. 139–148. Springer, Heidelberg (2009)
8. Lee, S., Kim, H., Lee, S., Lim, J.: Digital Evidence Collection Process in Integrity and Memory Information Gathering. In: Proceedings of Systematic Approaches to Digital Forensic Engineering, Proc. IEEE on First International Workshop, pp. 236–247 (2005)
9. Matthew, S., Jill, S.: Recovery of Skype Application Activity Data from Physical Memory. In: International Conference on Availability, Reliability and Security, pp. 283–288 (2010)
10. El-Khatib, K., Zhang, Z.E., Hadibi, N., Bochmann, G.V.: Personal and Service Mobility in Ubiquitous Computing Environments. Wireless Communications and Mobile Computing 4(6), 595–607 (2004)
11. Keenan, T.: On the Internet, Things Never Go Away Completely. IFIP International Federation for Information Processing 262, 37–50 (2008)
12. Andrew, M.W.: Defining a Process Model for Forensic Analysis of Digital Evidence Devices and Storage Media. In: Proceedings of the Second International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE). IEEE, Los Alamitos (2007)
13. Counts, S., Fisher, K.E.: Mobile Social Networking: An Information Grounds Perspective. In: Hawaii International Conference on System Sciences, HICSS (2008)
14. Nance, K., Hay, B., Bishop, M.: Digital Forensics: Defining a Research Agenda. In: Proceedings of the 42nd Hawaii International Conference on System Sciences. IEEE Press, Waikoloa (2009)
15. Walters, A., Nick, J., Petroni, L.: Volatools: Integrating Volatile Memory Forensics into the Digital Investigation Process. In: Black Hat DC (2007)
16. `http://www.gpsvisualizer.com`
17. `http://www.mactiontech.com`

Website Risk Assessment System for Anti-Phishing

Young-Gab Kim and Sungdeok Cha

College of Information and Communication, Korea University,
1, 5-ga, Anam-dong, SungBuk-gu, 136-701, Seoul, Korea
{always,scha}@korea.ac.kr

Abstract. Phishing attacks steal a user's identity data and financial account credentials using social engineering and technical spoofing techniques. Many counter measures have been developed to protect user's sensitive information from phishing attacks. Although most approaches use both website black lists (WBLs) and website white lists (WWLs), these approaches have several weakneksses. This paper presents a novel anti-phishing Website Risk Assessment System (WRAS). WRAS computes a security risk index of website and generates warnings as to the website trustworthiness. Therefore, it can protect inexperienced users against spoofed website-based phishing attacks and exploit-based phishing attempts that may occur from legitimate web pages.

Keywords: Phishing, Pharming, Anti-Phishing, Website Risk Assessment.

1 Introduction

Phishing is a way of attempting to steal user identity data and financial account credentials using social engineering and technical spoofing techniques in an electronic communication. Pharming (e.g., hijacking brand names of banks) is a more advanced phishing attack in that it redirect a website's traffic to another, forged website or proxy servers typically through DNS hijacking, often using crimeware, such as Trojan keylogger spyware [1]. Phishing and pharming attacks frequently occur and diversify, as Internet technologies (e.g., e-commerce, Internet banking, and social networking) evolve. Therefore, preventing a phishing attacker from stealing user's sensitive information (e.g., social security number, credit card number, account username, and password) is a major challenge in Internet security. Much research has been undertaken and many solutions have been proposed to protect user's economic loss and privacy against these phishing and pharming attacks. Most use website blacklists (WBLs) (list of e-mail addresses or IP addresses that originate from known true website) and website white lists (WWLs) (list of known safe website). Although many Internet Service Providers (ISPs) and enterprise security solution vendors use both WBLs and WWLs, these approaches have several weaknesses [2]: First, validity of WBLs is low, because the life cycle of phishing websites is short. Second, it is impossible to discriminate between legitimate and forged websites until the phishing attacks are detected and recorded in WBLs. Furthermore, the WBL- and WWL-based approaches hardly counter the new generation of sophisticated malware phishing attacks, pharming attacks, designed to target certain services. In addition, many of the

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 131–138, 2011.

existing WWL-based approach use only a website URL or IP address to distinguish between legitimate and forged websites. Therefore, the previous work [2] proposed a novel approach, which can quantitatively estimate the security risk of websites based on WWLs to overcome the limitation of WBL and WWL approaches. The proposed method in previous work defined the security risk elements and steps to quantitatively calculate the security risk index of a website.

In this paper, we present a novel Website Risk Assessment System (WRAS) for anti-phishing. WRAS employs a combination of the WWLs and the self-learning phishing filtering techniques that deliver high accuracy and wide coverage of websites. This can detect suspicious websites containing phishing attack and abnormal behavior. It generates a warning if the website is considered untrustworthy. Finally, it can protect inexperienced users against spoofed website-based phishing attacks and exploit-based phishing attempts that may occur from legitimate web pages.

The remainder of the paper is organized as follows. We briefly introduce background and related work in Section 2. We describe our solution, Website Risk Assessment System (WRAS), and provide details about its implementation in Section 3. Finally, we conclude the paper in Section 4.

2 Background and Related Works

Anti-Phishing Working Group (APWG) [1] is the global industry working group to eliminate fraud and identity theft that result from phishing, pharming and email spoofing. APWG is composed of many organizations and security companies. It provides diverse information, such as phishing reports, research data, and resources related to phishing, pharming, and crimeware.

Several research efforts have been made to prevent phishing attacks. In this section, we briefly review some typical approaches, dividing them into two parts: Server-side approach and Client-side approach.

Server-Side Approach. In the server-side approach, server authentication is required to defend against phishing attacks. One of the main reasons why phishing attacks are possible is because e-mails can be spoofed easily. Although spam filters work quite well today, they cannot guarantee that all phishing e-mails are intercepted. As example solutions, which authenticate the sender's e-mail, and prevent the phisher using a hijacked mail address, Microsoft presents the Sender ID Framework [3], and Yahoo uses its own technique called DomainKey [4]. Currently, Yahoo and other industry leaders are in the process of standardizing a technique called DKIM (DomainKeys Identified Mail) [5]. Another authentication approach is to share a secret, such as a password and an image, between server and client. Dhamija et al [6, 7] proposes Dynamic Security Skins, which allows that users visually verify whether the image from the server matches its corresponding local images. Finally, Fu et al. [8, 9] proposes a visual similarity assessment-based antiphising strategy, which uses visual characteristics to identify potential phishing websites and measure a suspicious web pages' similarity to actual sites registered with the system.

Client-Side Approach. In the client-side approach, most solutions are supported as a toolbar, which show different types of security messages to help users detect phishing

websites, built-in the web browser [10]. Chou et al. [11] proposes a framework for client-side defense using a browser plug-in called SpoofGuard that examines web pages and warns the user when a request for data may be part of a spoof attack. It uses domain names, invalid links, URL obfuscation, and images to measure the similarity between a given page and the pages in the caches. Many toolbars, such as Netcraft [12], EarthLink [13], and MS Phishing Filter [14], are designed to detect and prevent phishing attacks. Most of them use WBL and WWL, which depends on phishing reports. As long as a phishing website has not been reported, phishers may steal personal data from visitors to the website. Wu et al. [15] presents Web Wallet, which prevents phishing attacks by forcing users to compare, then confirm before going to a website instead of just confirming. A more comprehensive survey of anti-phishing solutions can be found in [16].

3 Website Risk Assessment System (WRAS)

3.1 Overview and Requirements of WRAS

WRAS is a system integrated into the Internet Explorer (IE) web browser. It checks and evaluates the security risk of a website domain and its web pages, before a user visits a website. That is, WRAS verifies the websites using WWLs and real-time analysis of web pages. A Website Qualified Lists (WQLs), which contains candidate websites for the WWLs, is designed in this paper to complement WWLs. Each candidate website has a score dynamically calculated by submissions from users. While the existing WWL-based solutions only maintain domain-specific information, such as IP address and URL, our approach uses domain-specific information defined in [2] and webpage-specific scores to reduce false alarms. When the webpage score of a candidate website is below the threshold, the website information is moved from the WQLs to the WWLs.

WRAS proposed in this paper follows typical client-server architecture. The system should satisfy the following requirements to efficiently estimate the security risk of website between client and server.

Client-Side Requirements. First, the client-side service in a system should obtain all information about a website the user wishes to navigate. This information includes webpage URL, DNS related information (e.g., Domain-Country, Domain Life, Domain-Age, DNS-Ranking, etc), and webpage analysis data that will be used to evaluate vulnerabilities. Second, the system should provide a user-friendly feedback system to efficiently maintain WWLs and WQLs in the client. Furthermore, the client service should not affect the performance of the client system.

Server-Side Requirements. The server-side service in a system has the same problem as the existing WBL-based solutions. The main problem is server overload caused by the heavy transaction of the database in the server-side. That is, the heavy transaction of WWL and WQL can compromise system performance. Furthermore, high security of the server-side system is required.

3.2 Architecture

WRAS employs typical client-server architecture, as illustrated in Fig. 1. The client side system, called WRAS client, is implemented as a plug-in to the IE web browser. When a user tries to visit a website, the WRAS client makes the first decision whether the website is a phishing website by checking a WWL cache in the client or a WWL DB in the server. If the website is not in the lists, the WRAS client calculates the security risk of the website using WQL cache in the client or WQL DB in the server.

The main goal of the WRAS server is to maintain and update WWLs and WQLs. by analyzing the security risk of the websites. That is, the server computes the security risk index for the websites based on the method proposed in previous work [2]. If a website security risk index (WSRI) is below a threshold, the server registers the website in the WWLs, and it transfers the website information to the WRAS client when the WRAS client requests website information. The following subsections present a more detailed description of WRAS client and server.

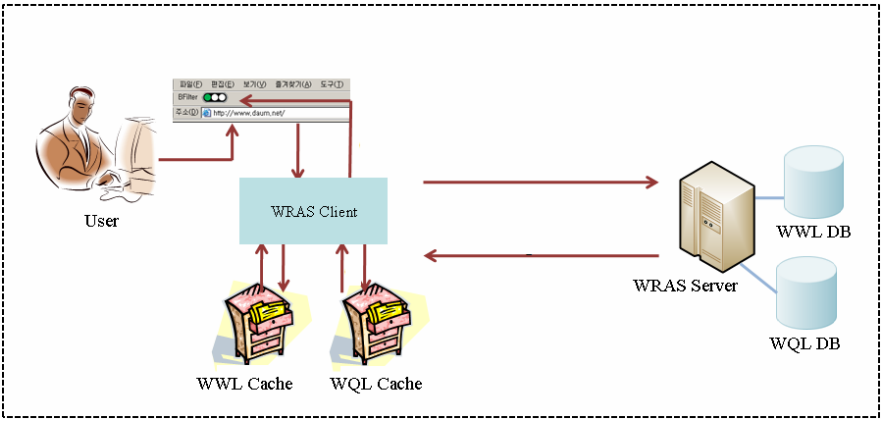


Fig. 1. Overview of WRAS architecture

Fig. 2 depicts the operational flow chart of WRAS: (1) A user inputs a target URL in the web browser. (2) WRAS client checks if the target website's webpage contains an inputtable page. If there is no inputtable page, the WRAS client permits a user to visit the target website. This is based on the premise that the attacker can only steal sensitive information in the inputtable webpage after the user performs a submit. (3) If the target website contains an inputtable webpage, the WRAS client first looks for the requested URL in the WWL cache. The WRAS client feedbacks the security risk index for the target website to the user, if the target URL is located in the WWL cache. (4) If the requested URL is not in the WWL cache, the WRAS client requests the target website information from the WARS server. (5) If the target URL is located in the WWL DB at the WRAS server, the WRAS server sends the target website information; then, the WRAS client analyzes the information. If the target URL is in the WWL cache or WWL DB, and the target website is not a portal website, the WRAS client permits a user to visit the target website (e.g., green light in the toolbar).

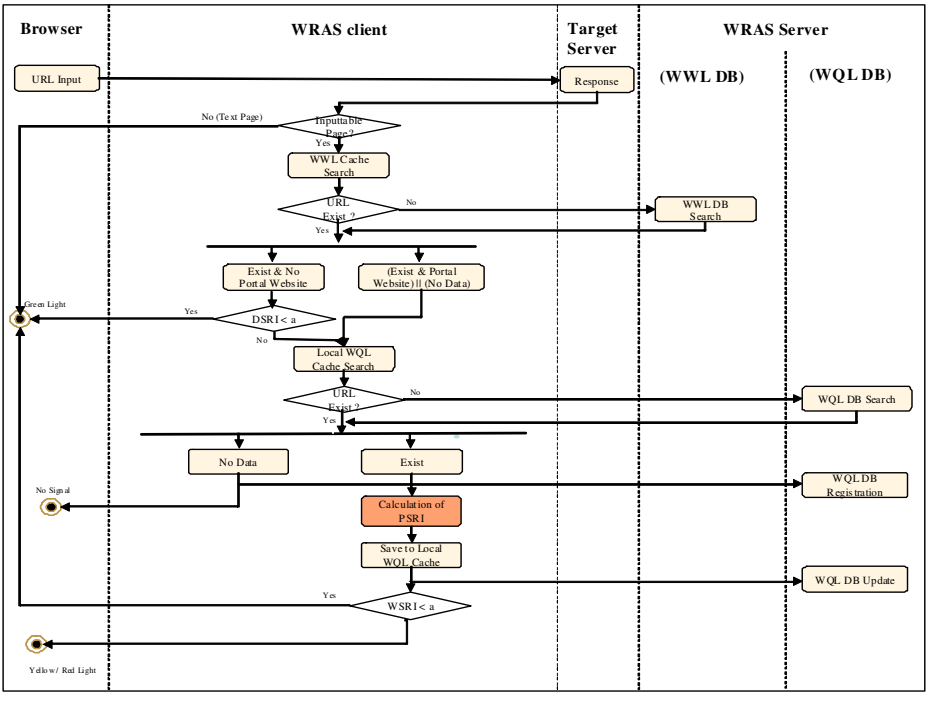


Fig. 2. WRAS Operational flowchart

(6) In the previous step, if the requested URL is not in the WWL DB at the WRAS server, the WRAS client looks for the request URL in the WQL cache. Like the previous steps, the WRAS client calculates the security risk index based on the information in the WQL cache if the target URL is located in the WQL cache. (7) If the request URL is not in the WWL cache, the WRAS client requests WQL DB from the WRAS server. If the target URL is located in the WQL cache or WQL DB, the WRAS client calculates a Page Security Risk Index (PSRI). It saves it to the WQL cache and updates WQL DB in the server. (8) Finally, if WSRI (e.g., WSRI is the value of DSRI+PSRI) is under the threshold, the WRAS client permits a user to visit the target website (e.g., green light in the toolbar).

3.3 WRAS Client

The WRAS client is implemented as a plug-in in the form of a toolbar in Internet Explorer (IE), since a majority of Internet users use this web browser, and the IE interface is an intimate environment for the user. Furthermore, the WRAS client can analyze the request webpage of users by developing a Browser Helper Module (BHO) on Microsoft Windows. BHO is the most popular technique to integrate a plug-in component into IE. The developer has access to the event mechanism of IE using BHOs and can create user interface elements, such as toolbars [10]. Fig.3 shows the WRAS client module plugged-in to IE.

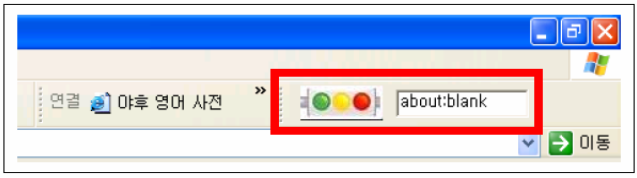


Fig. 3. WRAS client module plugged-in to IE

As depicted in Fig. 3, the WRAS client module is represented with signal lights (e.g., green, yellow, and red light). The signal light is activated by the webpage's security index. That is, the signal light reports the security risk index of the target website calculated by the proposed algorithm depicted in Fig. 2. Furthermore, the WRAS client supports more detailed security index information, as illustrated in Fig. 4.

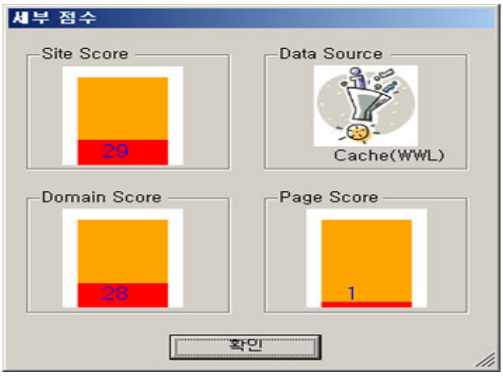


Fig. 4. Example of WSRI detailed information

The WRAS client, as well as the server, maintains WWL and WQL information (e.g., WSRI, Server-Name, Domain-Country, Domain-Life, Domain-Age, Domain-Famous, and DNS Ranking defined in [2]). Whenever a user tries to navigate to a website, the WRAS client checks the domain name (or IP address) in a WWL cache. If the domain name is not in the WWLs, the client proceeds to check the safety of the website, as explained in Section 3.2.

The WRAS client can provide the value of the website security risk directly to user. Furthermore, the performance is more effective than other approaches, such as an application or java applet, because it is a component in the web browser. In this paper, we inherit some pre-defined classes (e.g., UserControl, IObejectWithSite, IDeskBand, IDockingWindow, OIeWindow, IIputObject) by Microsoft to implement the WRAS client.

3.4 WRAS Server

The main functionality of the WRAS server is to maintain WWL and WQL DB by frequently analyzing the security risk of websites. The WRAS server supports the domain-specific information for the target website by an XML Web service whenever the client requests the target information. Furthermore, it helps a website manager control (e.g., register, edit, and delete) WWL and WQL DB, as illustrated in Fig. 5.

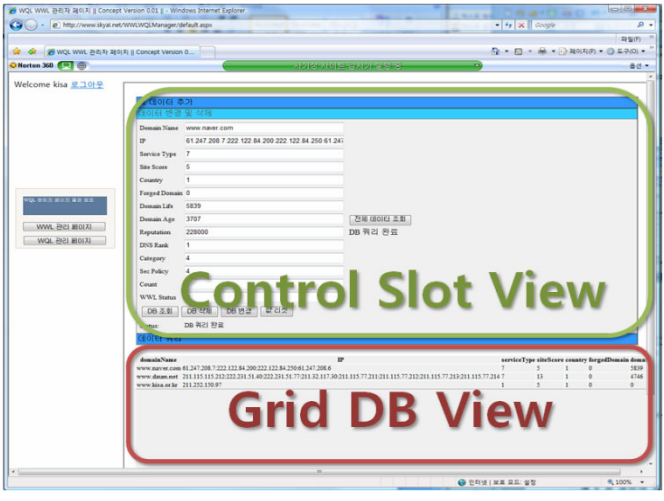


Fig. 5. WRAS server manager

In this paper, the XML Web service with Oracle DB is used to support effective processing and protection of communication data from external access. A popular RDB (Relational Data Base), such as Oracle or MS SQL, guarantee the system stability and security, because they have provided the service to users for a long time. Furthermore, XML Web Service guarantees compatibility between other products, because it operates in existing web environments.

4 Conclusion and Future Work

We presented a novel Website Risk Assessment System (WRAS) for anti-phishing. The system protects a user against phishing and pharming attacks. That is, the WRAS system supports a more precise security risk index for the target website using both WWLs and self-learning phishing filtering techniques. Thereby, inexperienced users can prevent phishing websites from transferring sensitive information.

The algorithm to compute the security risk index in the client and the heavy transaction of the database in the server are critical in influencing the effectiveness and efficiency of the system. Although our proposed system protects users from phishing sites, the detection algorithm and its efficiency should be improved.

Acknowledgement. This research was supported by the National IT Industry Promotion Agency (NIPA) under the program of Software Engineering Technologies Development.

References

1. Anti-Phishing Working Group (APWG) (2011), http://www.antiphishing.org
2. Kim, Y.-G., Cho, S.-H., Lee, J.-S., Lee, M.-S., Kim, I.H., Kim, S.H.: Method for evaluating the security risk of a website against phishing attacks. In: Yang, C.C., Chen, H., Chau, M., Chang, K., Lang, S.-D., Chen, P.S., Hsieh, R., Zeng, D., Wang, F.-Y., Carley, K.M., Mao, W., Zhan, J. (eds.) ISI Workshops 2008. LNCS, vol. 5075, pp. 21–31. Springer, Heidelberg (2008)
3. Microsoft, Sender ID Framework Overview (2011), http://www.microsoft.com/mscorp/safety/technologies/senderid/default.mspx
4. Yahoo: Yahoo! Anti-Spam Resource Center (2008), http://antispam.yahoo.com
5. Mutual Internet Practices Association, DomainKeys Identified Mail, DKIM (2011), http://www.dkim.org
6. Dhamija, R., Tygar, J.D.: The Battle against Phishing: Dynamic Security Skins. In: Proc. of the 2005 Symposium on Usable Privacy and Security (SOUPS 2005), pp. 77–88 (2005)
7. Dhamija, R., Tygar, J.D.: Phish and Hips: Human Interactive Proofs to Detect Phishing Attacks. In: Proc. of the Second International Workshop, pp. 127–141 (2005)
8. Fu, A.Y., Wenyin, L., Deng, X.: Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover's Distance (EMD). IEEE Transactions on Dependable and Secure Computing 3(4), 301–311 (2006)
9. Liu, W., Deng, X., Huang, G., Fu, A.Y.: An Antiphishing Strategy Based on Visual Similarity Assessment. IEEE Internet Computing, 58–65 (2006)
10. Raffetseder, T., Kirda, E., Kruegel, C.: Building Anti-Phishing Browser Plug-Ins: An Experience Report. In: Proc. of third international workshop on Software Engineering for Secure Systems, SESS 2007 (2007)
11. Chou, N., Ledesma, R., Teraguchi, Y., Boneh, D., Mitchell, J.C.: Client-side Defense against Web-Based Identity Theft. In: Proc. of 11th Annual Network and Distributed System Security Symposium, NDSS 2004 (2004)
12. NetCraft (2011), http://www.netcraft.com
13. EarthLink (2011), http://www.earthlink.com
14. Microsoft, Anti-Phishing Technology (2011), http://www.microsoft.com/mscorp/safety/technologies/antiphishing/
15. Wu, M., Miller, R.C., Little, G.: Web Wallet: Preventing Phishing Attacks by Revealing User Intentions. In: Proc. of Symposium On Usable Privacy and Security (SOUPS 2006), pp. 102–113. ACM Press, New York (2006)
16. Emigh, A.: Online Identity Theft: Phishing Technology, Chokepoints and Countermeasures. ITTC Report on Online. Identity Theft Technology and Countermeasures (2005)

Delivering Real-Time Bus Tracking Information on Mobile Devices

Bashir Shalaik and Adam Winstanley

National University of Ireland Maynooth
{bsalaik,adamw}@cs.nuim.ie

Abstract. The recent technological advances in mobile communication, computing and geo-positioning technologies have made real-time transit vehicle information systems an interesting application area. In this paper we present a transit application system which displays the transit information on an OpenStreetMap (OSM) web interface and delivers this information on the Google Android mobile device. The content is in the form of predicted arrival/departure times for buses at user-selectable geographic locations within a transit region. This application uses the real-time information such as current location and timestamp of both bus and users to provide bus route information. The public interface provides a graphical view which used to display and updates the vehicle locations and to allow users to see routes, stops and moving buses. The mobile device provides the user with the expected arrival/departure time of the next bus at the bus stop based on the user's current position.

Keywords: Real-time data, OSM, AVL, Android, Openlayers.

Introduction

In transportation information, different mappings interfaces are used to provide cartographical means for displaying, monitoring, and improving transit vehicles performances. In a pilot project between National University of Ireland Maynooth (NUIM) and Blackpool transit (Great Britain) a prototype of a working web-based real-time bus tracking transit system is developed using OpenstreetMap(OSM). In real-time bus tracking systems data is collected in real-time and transmitted to a central server for analyzing and extracting transit information. Computer software technologies such as Ajax is used to update the map-based display without interrupting the users by switching pages or screens as they view the map[1]. The developed transit system pro-vides services that take advantages of modern technologies to displays bus Arrival/departure time information on the platform of Google Android mobile devices. Doing so will enable passengers to enquire about bus-arrival/departure time to a selectable geographic locations of their interest.

For Bus Arrival/departure time estimation, three prediction models (namely, a historical-data-based model, a multiple linear regression model and one-dimensional Kalman filter model) were implemented and their performance is evaluated using the Mean Absolute Percentage Error (MAPE) [2] [3]. In this paper, the process for

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 139–147, 2011.

downloading data to platform of the Google Android device and mobile interfaces were designed using the Android operating system for mobile devices and java programming language. The utility of displaying bus arrival/departure times to selectable geographic locations can be applied to both a fleet management context and a bus information system environment. The developed system provides real-time information about bus routes and bus locations for those how have a mobile device with internet accessibility. They can link to the web site and get the current transit information such bus arrival time to the nearest bus stop.

1 Related Works

Real-time arrival information for bus, subways, light rail, and other transit vehicles are displayed in a significant number of cities worldwide at places such as rail stations, transit centers, and major bus stops. With the possibility that real-time transit information will not be available on a public display at every stop, the smart mobile devices are being used to help manage the complexity of using transit information. Whether it is a simple phone or SMS interface, or a more complex native mobile application, these systems can provide schedules, routes and real-time arrival information. Google Transit, which was started as a Google Labs project in December of 2005, is now directly integrated into the Google Maps product and provides interfaces to Google Transit which are exist on a variety of mobile devices, making use of location sensors such as GPS and WiFi localization on the device to improve the usability of the transit application. Various mobile-phone-based transit information systems have been developed to provide users with transit information.

The Intelligent transportation system research group at university of Washington has developed a real time-system for predicting bus-arrival time, based on access to transit agency data. The prediction times made available to the traveling public via web site known as MyBus. The usability of public transit system can be enhanced by providing good traveler information system. OneBusAway [4] is a set of transit tools focused on providing real-time arrival information. This application made use of increased availability of powerful mobile devices and the possibility of displaying transit data in machine readable format. In OneBusAway systems, transit information such as bus arrival time to a particular bus stop is displayed on internet-enabled mobile devices. In [5] the usage of a transit vehicle information system that delivers estimated departure times for a large transit fleet is described. Due to the physical restriction of mobile devices which affects the user interaction and data presentation, the WML, has been introduced as the new language for WAP-enabled device.

In transit trip planner (TTP) system known as Trnasitr [6]. The system provides the shortest paths between any two points within the transit network using the real-time information provided by a third party bus arrival prediction system, relying on GPS equipped transit vehicles. Users submit their origin and destination points through a map-based iPhone or through a Java script enabled web browser. Services such as 511.org and Google Transit allow users to plan public transit trip by generating routes based on static schedule data where as with the proposed Transidriod system a Dynamic transit information is received via web services. In [7] a mobile public transportation information service was developed to provide map-based information

of the nearest mass rapid transit station, the nearest bus stop of the bus route chosen by the user, that can take the user to his/her chosen destination. The developed systems can deliver a map marked with the nearest mass rapid transit station on a Nokia 6600 cell phone. In [8] a transit information system was developed to implement a bus arrival time predictor on a Google Android mobile device. The developed system provides relevant bus routes information with arrival time to users in order to explore the possibility and capability of various sensors and GPS on the device. In the developed system, the users' current location is collected and, together with a stored bus schedule, a bus arrival time is calculated on the server and was displayed on the mobile device on a built-in Google Maps display.

2 Real-Time Bus Tracking Systems

Intelligent transportation systems provide the technology to enables people to make smarter travel choice. Better transit information allows users to make better decision about their travel options. One of the first online bus tracking systems, BusView, was developed at the University of Washington [9]. Nowadays, many public bus services provide on-the-fly information to their users, including the current locations of buses and the predicted arrival times at bus stops. These buses typically use the Global Positioning System (GPS) for positioning and use wireless communication such as radio or GSM/GPRS for communicating their position to a central server. Figure (1) shows the architecture of a real-time vehicle tracking system. Real-time prediction of

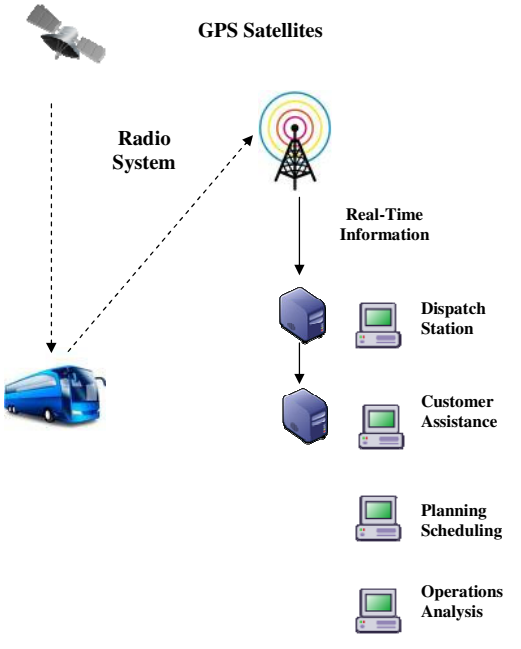


Fig. 1. Real-Time Bus Tracking System's Architecture

accurate bus arrival times has been studied in literature for a couple of decades In the real-time bus tracking system, Information on expected arrival and departure times is difficult to estimate for buses which share their journeys with other traffic. Different methods are used for predicting bus arrival time, some researchers use simple statistical/mathematical models, e.g. prediction according to deviation from the schedule [10]. Kalman filter and more sophisticated artificial intelligence and machine learning algorithms have also been used [11] [12]. However, no single predictor had yet been developed that presented itself to be universally accepted as the best. In this project, three bus arrival models were tested, namely a historical-data-based model, a multiple linear regression model and a one-dimensional Kalman filter model. These models have fewer parameters and can be adaptive to real-time updating data. The Automatic Vehicle location (AVL) data is used to track vehicle location while the other vehicle data is used to predict arrival time to a certain bus stop or the selected transit area along the route. Figure (2) shows the tracking and predicting components of real-time bus tracking system.

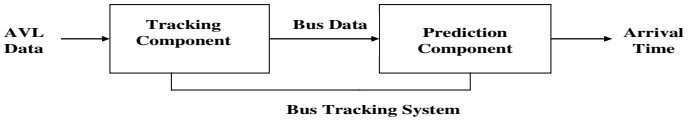


Fig. 2. Tracking and Predicting Components of real-time bus tracking system

In order to evaluate the performance of the three models, the Mean Absolute Percentage Error (MAPE) was used as a measure of the absolute difference between the predicted and the observed values.MAPE represents the average percentage difference between the observed value and the predicted value [13]. Table (1) shows MAPE values of the three prediction models on the same sample test data. Historical data-based model outperformed the other models due to the stability of traffic pattern in the area of the interest (Blackpool city).

Table 1. MAPE Values of Bus Arrival Times Prediction Models

Model	MAPE
Historical Data Model	13 %
Kalman Filter Model	20 %
MLR Model	29 %
Average	20.6%

3 System Implementation

In Blackpool real time bus tacking system and to increase the satisfaction of among transit users, the system delivers transit information via standard web browser and

mobile devices. The system uses off-the-shelf GPS/GPRS integrated units programmed to transmit location at regular intervals (45 seconds approximately) while the vehicle is in motion. The data is stored on a server and is then visualised through a standard web browser to show views representing current locations of vehicles in close-to-real-time.

3.1 Web Interface Development

The system uses web technologies such as JavaScript, MySQL, XML, PHP and Ajax. The position of the bus along with the timestamp and bus details is sent to the server using GPRS. The remote server inserts the data into a MySQL database. An interactive public interface was developed to allow user more interaction with the transit system Figure (3) shows the public interface of Blackpool transit system with updated vehicle location on an OSM interface developed using OpenLayers.

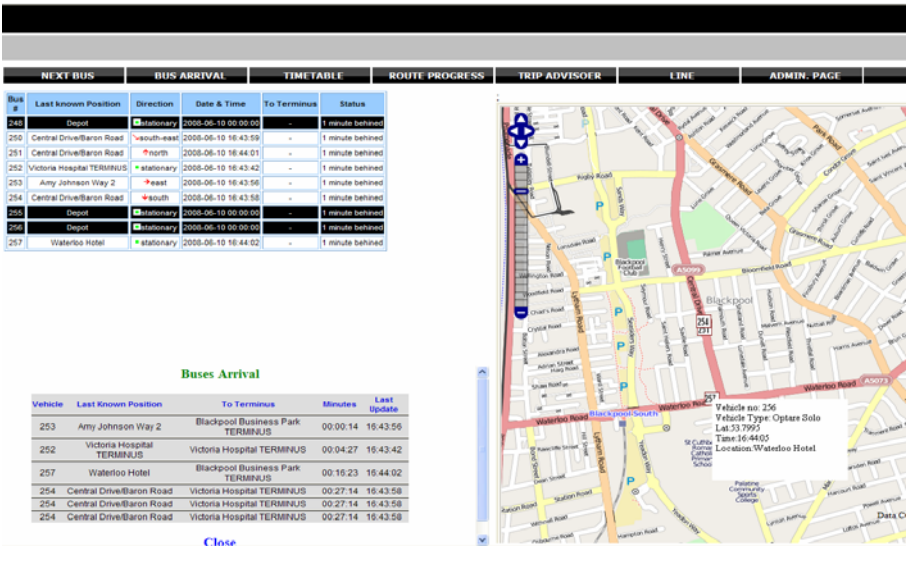


Fig. 3. The Public interface showing updating textual display plus moving locations on OSM

Ajax technology is used to continually update the backend database and update the map display without interrupting the user by switching pages or screen as they view the map. Fig (4) shows the Ajax techniques for updating the map-based display in the browser. The powerful combination of OpenLayers and OSM for the display spatial data is providing a very strong Information Technology (IT) support for transit data utilization.

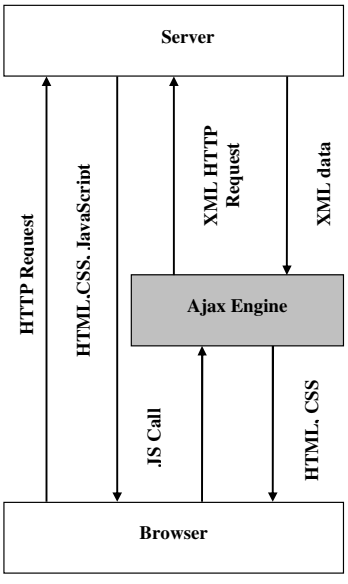


Fig. 4. Ajax techniques for updating map-based display on browser

3.2 Mobile Interface

The wireless communication technology is designed to utilise existing internet protocols and standards. A URL address is used for the identification of a resource, and HTTP is used as a protocol between WAP gateways and content server. Wireless content can be served using existing web server software.

The smart mobile phone interface was developed using the open source android operating system for mobile devices. In this project an HTC Magic smart phone that runs on the Android operating system was used to communicate with the transit project server. The mobile device uses a HTTP protocol to connect with the MySQL database on server. Data between Android and PHP could be coded in JavaScript Object Notation (JSON) format which has built-in JSON functions in both languages. To view arrival/departure time on the mobile device, the user selects his/her preferred destination; the transit system collects his current location from built-in GPS. The user's current location along with the destination selected by the user is sent using the HTTP protocol to the server by appending it to the URL which connects to the server. The URL from the mobile device is transmitted and the response is displayed on the mobile screen. The nearest bus stop to his/her current location is suggested and Bus Arrival time is displayed on the device. Figure (5) shows a sequence diagram of the interaction between users, mobile Devices ad the web.

Figure (6) shows the initial mobile interface on Android emulator of the proposed system for the Blackpool transit system.

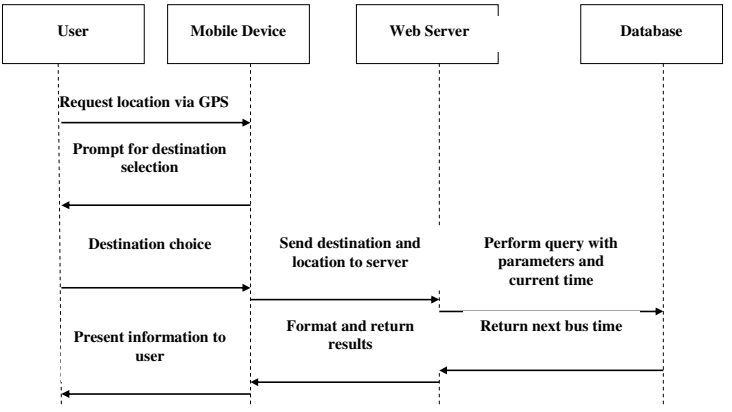


Fig. 5. A sequence diagram showing the interaction between users, mobile Devices ad the web

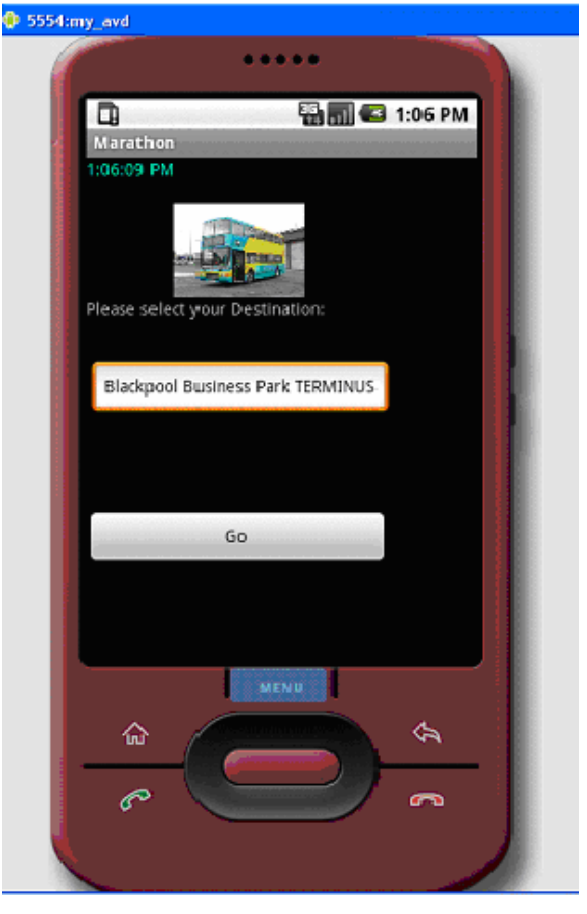


Fig. 6a. Shows a desire distention selected by user

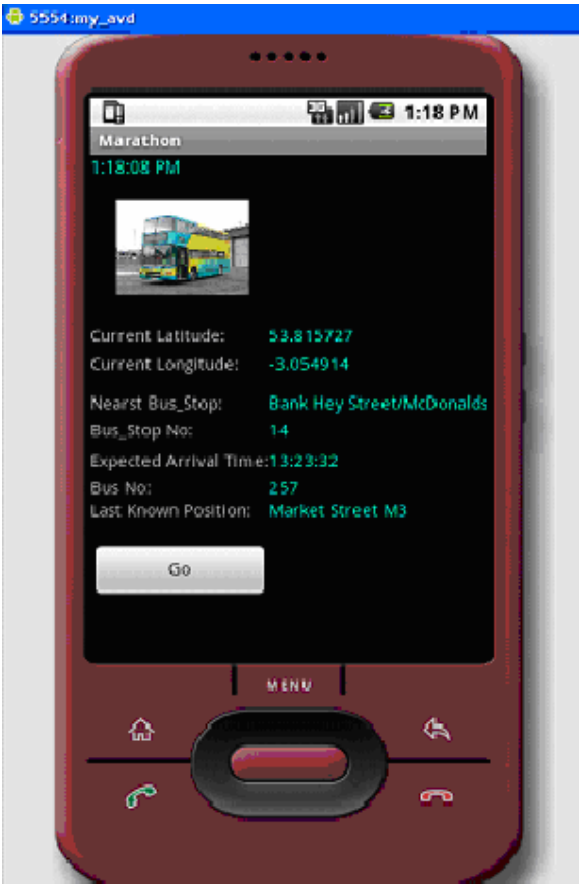


Fig. 6b. Shows the response from server

4 Conclusions and Future Work

In this work we have shown that transit information collected in real time can be shown on OSM for tracking and monitoring purposes. Internet enabled mobile phones can receive real-time transit information. Android software for smart mobile phones offers the ability to overcome the physical restriction of interface design on mobile phones. In this work three different bus arrival prediction models were applied, testes show that the historical data-based model over preformed Kalman filter model and multiple regression model. Real-time transit information was delivered on Android mobile devices to enable transit users making a better traveling choice and save their waiting times at bus stops. If the user is provided with the accurate arrival information, the reliability of the service will increase and the users will shift from the private mode. To further improve bus arrival time prediction accuracy, transit data from other sources can be incorporated into the predictor algorithms. Future work on this project includes development of a feature which alerts a user when bus is a specified number of minutes away.

Acknowledgements

Thanks are due to Blackpool Transport Ltd for facilitating the project and particularly to Oliver Howarth for his comments and feedback. Research presented in this paper was funded by a Strategic Research Cluster grant (07/SRC/I1168) by Science Foundation Ireland under the National Development Plan. One author is supported by a PhD studentship from the Libyan Ministry of Education.

References

1. Zepeda, J.S., Chapa, S.V.: From Desktop Applications Towards Ajax Web Applications. In: proceedings of the 4th International Conference on Electrical and Electronics Engineering (ICEEE 2007), Mexico City (2007)
2. Patniak, J., Chien, S., Baldikas, A.: Estimation of Bus Arrival Times using APC Data. Journal of public transportation 7(1) (2004)
3. Predic, B., Rancis, B.: Automatic Vehicle Location in Public Bus Transportation System. In: Proceedings of the 11th International Conference on Computers (WSEAS) Crete Island Greece (2007)
4. Brian, F., Kari, E., Alan, B.: OneBusAway: Behavioral and Satisfaction Changes Resulting from Providing Real-Time Arrival Information for Public Transit. In: Proceedings of CHI 2010, USA (2010)
5. Maclean, S.D., Daily, D.J.: The use of Wireless Internet Services to Access Real-Time Transit Information. In: Proceedings of IEEE Intelligent Transportation Systems Conferences, USA (2001)
6. Jariyasunant, J.: Mobile Transit Trip Planning with Real-Time Data. In: Proceedings of (TRB 2010) Annual Meeting, USA (2009)
7. Shwu, J., Gong, H., Shian, H.: Location-Aware Mobile Transportation Information Service. In: Proceedings of the 2nd International Conference on Mobile Technology, Applications and Systems, Guangzhou (2009)
8. Lam, K.: Bus Arrival Predictor on the platform of Google Android , http://lbms03.cityu.edu.hk/studproj/cs/2009cslkh140.pdf
9. Dailey, D., Fisher, J.: BusView and Transit Watch: an Update on Two Products from the Seattle Smart Trek Model Deployment Initiative. In: Proceedings of the 6th World Congress Intelligent Transport System, USA (1999)
10. Wei-Hua, L., Zeng, J.: Experimental Study of Real-Time Bus Arrival Time Prediction with GPS Data. Publication of Transportation Research Board of the National Academies, USA
11. Dailey, D., Wall, Z.R.: An Algorithm and Implementation to Predict the Arrival of Transit Vehicles. In: Proceedings of the 2nd IEEE Conference on Intelligent Transportation Systems, USA (2000)
12. Jeong, R., Rilett, R.: Bus arrival time prediction using artificial neural network model. In: Proceedings of the 7th IEEE Conference on Intelligent Transportation Systems (2005)
13. Wikipdedia The free Encyclopedia, http://en.wikipedia.org/wiki/Mean_absolute_percentage_error

Camel-Vehicle Accidents Mitigation System: Design and Survey

Khaled Ragab[1], Mohammed Zahrani, and Asrar Ul Haque

College of Computer Sciences and Information Technology, King Faisal University, Hofuf, Saudi Arabia Kingdom
{kabdultawab,zahrani,ahaque}@kfu.edu.sa

Abstract. Animal-vehicle collisions (*AVC*) affect human safety, property and wildlife. Furthermore, the number of collisions with large animals worldwide and especially in the Saudi Arabia Kingdom has increased substantially over the last decades. The negative effects of *AVC* and the increase in collisions prompted the initiation for designing a deployable and intelligent *Camel-Vehicle Accident Avoidance System* (CVAAS) using *global positioning system* (GPS) technology. *CVAAS* can be classified as an Intelligent Transportation System (*ITS*). The use of GPS technology in this kind of application is a *novel idea*. This article provides a detailed discussion in there related literature review. Moreover, it discusses the high-level design of the *CVAAS*.

Keywords: GPS, Animal vehicle-collision avoidance technologies.

1 Introduction

Reducing animal-vehicle accidents across roadways are significant issues to consider in highway construction for human safety, economical, and ecological reasons. In the Saudi Arabia Kingdom (*KSA*), hundreds of Camel-Vehicle accidents are reported every year causing numerous deaths and loss of property running into billions of Saudi Riyals. Summaries of traffic accident data show that more than 600 camel-vehicle accidents occur each year [1]. Similarly, the total number of reported animal-vehicle accidents in United States (*US*) is approximately 300,000 per-year [20]. In Europe and Canada moose and deer have been shown to be a considerable problem on the road [15]. *AVC* is not only a traffic problem in *KSA* but also considered a major safety problem in the US, Japan, and Europe [4]*, Stout et al. 1993*). In *KSA*, usually camels that are found near highway are domestic camels because the owners like to live close to highway for transportation facility. These camels move across highways looking for water and food, and during mating season. Camels are very hard to detect by vehicle drivers especially in the night time and results in severe accidents if a collision occurred. According to *Al-Ghamdi* and *AlGadhi* [1] study, the most frequently involved animal in *AVC*'s is camel; it is estimated that *97%* of all reported *AVC*'s were camel related. More than *90%* of these accidents occur at night, between

[1] On leave from Ain Shams University, Cairo, Egypt.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 148–158, 2011.

dusk and dawn [22]. These accidents cause a lot of damage to the environment, economy and social life such as significant economic loss, human injuries and/or fatalities, loss of valuable wildlife, and damage to properties as seen in figure 1. Langley [25] examined risk factors involved with fatal *AVC*'s in the *US* from 1995 to 2004 and found that *89.5%* occurred on rural roads, *64.8%* in darkness, *85.4%* on straight sections of road, *91.1%* occurred in dry weather conditions, and *28%* of the victims were motorcyclists.

Fig. 1. Camel-Vehicle accident

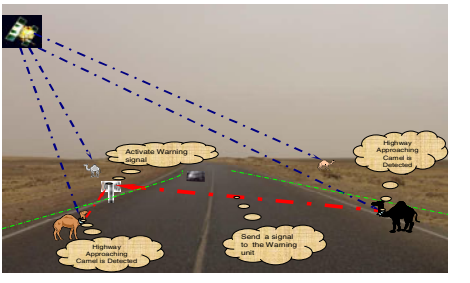


Fig. 2. The Warning System being Activated as Camels approach the Highway

More efforts need to be done to reduce the number of *AVC*'s. Most researches have attempted to deal with the *AVC*, but neither unique solutions nor efficient results have been found. Many kinds of animal detection and warning systems are used around the world to indicate presence of animals on highways to avoid accidents. Animal detection systems are divided into three main categories namely *Road-based*, *Vehicle-based*, and *Animal-based*. Detailed discussion will be undertaken in the literature review. This paper proposes a design of a ***Camel-Vehicle Accident Avoidance System*** (*CVAAS*) using *global positioning system* (GPS) technology. The use of GPS technology in this kind of application is a novel idea. Use of GPS receivers has increased tremendously for navigation purpose, in tracking animals [33], [35], in sensors networks, and many other applications. GPS receiver can be obtained for a reasonable price of around 20-50$. Thus, the use of programmable GPS devices in *CVAAS* is a novel and feasible solution. *CVAAS* system is an *animal-based* system that identifies the presence of a camel on or near the highway and then sends out its position to the *Dedicated Short-Range Communication* (*DSRC*) transmitter. Consequently, the *DSRC* transmitter forwards the camel position to a *DSRC* receiver mounted on a warning system. The signal will activate the warning system to warn the vehicle drivers to slow down in order to avoid collision with the camel. Figure 2 illustrates such a scenario.

The remainder of this paper is organized as follows. Section 2 illustrates and classifies the current animal-vehicle collision avoidance technologies. A discussion of these technologies has been presented in section 3. Section 4 briefly introduces the design of *CVAAS*, the proposed novel system. Section 5 draws a conclusion of this paper.

2 Current *AVC* Technologies

Worldwide numerous technologies have been used in attempts to reduce *AVC*. This section presents a brief review of the methods widely used to reduce *AVC* on most of the world's highways. The techniques used in the past to reduce *AVC* were classified into three categories *roadway-based*, *animal-based*, and *vehicle-based* techniques. The *roadway-based technique* developed for and dedicated to the highways. It includes roadway fencing, underpasses, overpasses, roadway reflectors, warning signs, infra-red, microwave RF, etc. Roadway-based detection systems, however, are designed to inform all drivers, regardless of what equipment their vehicle may or may not have. The second category, animal-based includes the technologies that were mounted in animals to reduce the *AVC*. The third category, vehicle-based includes the technologies that are equipped into vehicles to reduce the *AVC*.

2.1 Roadway-Based Technologies

2.1.1 Roadway-Based Conventional Techniques

Fences. They have been installed to keep animals away from the road [7], [6]. Roadway fencing is one the famous conventional techniques used to reduce *AVC*. It is the only method used to stop camels from coming on the KSA's highways. *Ward* [36] signified that a 2m high big-game fence is effective in reducing vehicle collisions involving deer. Fencing is extremely expensive because they have been combined with wildlife crossing structures such as *underpasses* and *overpasses* that enable animals to move freely along both sides of the highways. Fencing must be inspected frequently and repaired to original condition to be successful at reducing collisions because animals quickly exploit breaks in the fence [10]. Apparently, deer continually test fencing, making a good maintenance program necessary [36].

Warning Signs. They warn drivers of high large animals crossing locations are the most common approach to reducing *AVC* [28]. *Romin* and *Bissonette* [31] suggested that deer crossing signs may be effective if drivers would reduce their vehicle speed. Lighted, animated deer-crossing warning signs were evaluated in Colorado. *Pojar* concluded that drivers' speeds were initially slightly reduced, but after the drivers got used to the animated sign, it lost its effect in reducing the drivers' speeds [27].

Highway lighting. Most of the *AVC* occurred from sunset to sunrise. It was expected that highway lighting enhance drivers' night vision and reduce *AVC*. Highway lighting did not affect drivers' behavior or animal crossings-per-accident ratios [30]. Thus, increased highway lighting was not effective at reducing *AVC*.

2.1.2 Roadway-Based Detection Systems

Animal detect technologies detect large animals as they approach the road. When an animal is detected, signs are activated that warn drivers that large animals may be on or near the road at that time. Vehicle detect technologies operate on a slightly different principle as they detect vehicles, not the animals. They detect vehicles or trains, not the animals. Once a vehicle or train is detected large animals are alerted through a range of audio and visual signals from stations placed in the right-of-way

[18]. Briefly, this section will list different technologies that have been used to develop animal-detect and vehicle detect techniques.

Infrared sensors. They were designed and installed in seven sites in *Switzerland* to detect deer within *30-100*m radius on both sides of the road. Once a deer was detected LED signs with a deer symbol were activated and stayed on for *45* seconds to alert the drivers [23]. This technique produced false detections because of strong winds and warm engines of passing vehicles. Moreover, the broken sensors, loss of power due to snow covered solar panels and broken lamps in the warning signs caused additional problems. Similarly, the *Flashing Light Animal Sensing Host* (*FLASH*) was designed to detect mule deer in Wyoming, USA. It also used a series of infrared sensors [14]. More than *50%* of the detections through *FLASH* system were false. This was due to frost on the sensors, birds feeding on carrion in the crossing area, and snow thrown by passing snowplows.

Microwave radar sensors. In *Finland*, they were designed and installed to detect large animal movements up to *50* m in distance within a *60°* horizontal angle. When an animal was detected, LED message signs with an animal symbol were turned on and remained on for two-three minutes. To verify the presence of animals a video camera is installed. In addition, to distinguish animal from other moving objects such as rain or rain spray, the system was programmed to only detect objects moving towards the sensors at a speed greater than 0.8m/s. This technique produced false detection in spring when the snow melted and the water warmed on the pavement, spray from passing vehicles triggered the system.

Laser sensors. In *2000* an animal detection system was installed in Washington, USA. It consisted of *two* lasers, one placed on each side of the road, two standard deer warning signs, two smaller rectangular signs that read "When Flashing," and two solar-powered red flashing beacons. When the laser beam was broken the lights were switched on. The lasers operated on batteries with a one-week lifespan while the red strobes were solar powered. The sighting of the lasers proved difficult, partly because of the distance between the sensors. Sunlight heating up of the plastic boxes holding the laser equipment may have caused problems with the sighting of the laser. False detections caused the batteries to drain quicker than anticipated. Similarly, in October 2002, an animal detection system was installed along US 97A, near *Wenatchee*, *Washington*. It used ***laser beams*** to detect deer. If deer stays there longer than one minute, the warning signals were turned off, and drivers are no longer warned of its presence [37]. The laser beams could only be used for short distances on straight sections of roadway. Anything that broke the beam triggered the warning, including birds, dogs, mail trucks and snow plow curls. Perfect alignment was critical (high maintenance costs). Even the sun could trigger the beacons depending on the time of year as sunrise and sunset angles changed.

Microwave technology. An animal detection system consisted of series of transmitters and receivers. It was installed in 2002 along the highway in Montana. Each transmitter sent a uniquely coded, continuous microwave RF signal to its intended receiver [34]. The transmitters and receivers were mounted about 120cm above the ground. The system produced a large number of false detections for several causes such as snow spray.

A vehicle detect system has been installed in *April*, *2002* in *Canada*. It consists of a small cabinet with electronics, sensors for vehicle detection, and an animal warning device. The units are powered by solar panels and batteries. When no vehicles are present the system is not active. Once vehicles are detected, units in the roadside are activated that alert deer through a variety of noise and light signals [21]. Many kinds of animal have been shown to adapt to disturbance if this is not accompanied by an immediate and real threat. Therefore, the audio and visual signals produced by the stations in the right-of-way may not scare the animals away from the road once they have been exposed to it for a certain time. Additionally, such system is not well suited for high traffic flows since the animal warnings would be running continuously in such locations.

2.2 Animal-Based Technologies

The animal based technologies to mitigate *AVC* used different types of collars fasten with the animal to trigger a warning system such as blinking signals. They are classified as reflective collars and radio collars.

Reflective collars. In *British Columbia*, *Canada*, the ministry of environment conducted a method to reduce *AVC*. In 2006, they put collars with reflective tape on a number of animals to increase their visibility to drivers. In addition, a major company *Aramco* in *KSA* distributed around *3000* reflective collars to the camels' owners in *Al-Ahsa*. These collars are not efficient to reduce the *AVC* because vehicles must be close enough to ensure that the collars are visible which defeats the whole purpose of avoiding accidents. Moreover, the reflective materials of the collars will disappear over time.

Radio Collars. Multiple of projects utilized radio collars since 1999 up to now. A system was installed along a 4,827m long section of Hwy 101, near *Sequim*, on the *Olympic Peninsula,Washington*. In 1999 about 10 percent of the elk herd was radio collared [5]. An effort was made to radio collar lead cows, but this was not always possible. Receivers placed along the road scan for the frequencies of the individual radio collars 24 hours per day. When the radio-collared individuals come within about 400m of the road, the receivers that pick up the signal activate the flashing beacons. As a consequence, the animals without a radio collar are only detected if radio-collared animals accompany them. Therefore, the system only works well for highly gregarious species. The radio-collar system requires re-collaring effort. The batteries of the radio collars usually run out after several years, and then they must be replaced.

GPS Collars. It is a valuable tool for documenting the movements of large, wide-ranging animal kinds. Recently, GPS collar has been instrumental in monitoring large mammals use of highways and wildlife underpasses in Arizona [26]. Using data gathered from GPS collar, [26] were able to identify spatial patterns in bighorn sheep movement relative to a key section of US 93. Based on GPS collar data, the authors were able to make informed recommendations regarding placement of wildlife-engineered crossing structures on US 93. Dodd et al. [8] used GPS collars to assess permeability of SR 260 to elk through successive phases of reconstruction, which included widening the highway, integrating wildlife crossing structures, and

implementing ungulate-proof fencing. Gagnon et al. [11,12] were able to determine how patterns in traffic flow affected elk crossing and distribution in the vicinity of SR 260; the authors found that although high traffic volumes greatly affected elk crossings, seasonality and proximity to quality habitat also strongly affected elk behavior. To the best of our knowledge, most of the systems that used GPS collars only to monitor large animal movements for the sake of recommending the placements of wildlife-engineered crossing structures on highways.

2.3 Vehicle-Based Technologies

The vehicle based technologies to avoid *AVC* can be broadly classified into two major groups: warning whistles (e.g deer whistles) and infrared detection systems. They would not depend on the installation of any roadside equipment. Deer whistle were introduced as early as late 1970s, [24]. Air activated deer whistles, mounted on the fronts of vehicles, allegedly produce ultrasonic frequencies and/or audible sounds from the wind rushing through them. These sounds are supposed to scare away animals. It has been observed that given the masking effect of road and vehicle noise, however, it is unlikely deer would be able to hear the whistles Romin and Dalton 1992 [32]. In addition, there is no evidence that audio signals affect animal behavior [2] and habituation to sounds has been observed [32], [29]. However, the infrared detectors inform drivers when a large animal is detected within a certain range from the sensors attached to the vehicle (e.g., [3], [16], [17]). The range should be sufficient to allow for the driver to stop the vehicle before impacting the detected animal. As an option on the *Cadillac DeVille* an infrared sensor, mounted in the front grille, picks up heat energy from a person or an animal. The image is projected onto a monochromatic display on the lower part of the driver's side of the windshield. Hot objects appear white and cool objects appear black in that image. Some drivers have noted that objects are difficult to see and appear fuzzy due to the field of view that is too limited to be useful. Others have complained of headaches after only one hour of use.

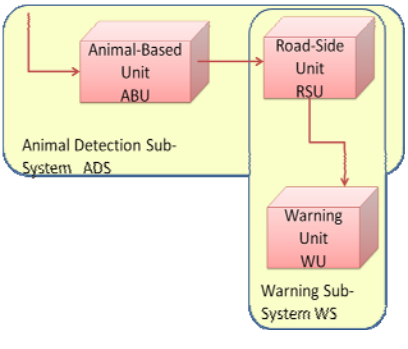


Fig. 3. CVAAS: Block diagram

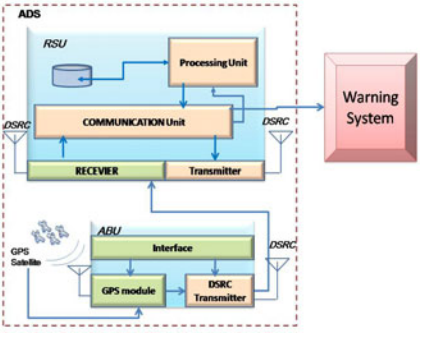


Fig. 4. CVAAS components

3 Discussions

The systems to reduce *AVC* encountered some technical problems and maintenance issues. More importantly they experienced false positives and false negatives. The false positive occurs when the warning system is activated even if there is no animal. Whereas, false negative occurs when there is animal but the warning system is not activated. The *AVC* systems with break-the-beam sensors may experience false positive detections due to falling branches in forest, especially in strong winds or snow spray from snowplows, etc. Broken sensors, loss of power due to snow-covered solar panels, and broken lamps in the warning signs may caused false positive. False negative may occur due to curves, slopes not covered by sensors and insufficient warning time [19]. False positives may cause drivers to eventually ignore activated signs [13] and false negatives present drivers with a hazardous situation. It is of immense importance that any system designed to reduce or avoid *AVC* should ensure minimal number of false positives and false negative. *CVAAS* aims to address these false detection problems through employing the novel idea of using a programmable GPS device which gives accurate positioning of an animal. In the following section, this paper discusses the high level design of *CVAAS*.

4 CVAAS: Design

The design of the *CVAAS* consists of two sub-systems: Animal Detection sub-Systems (*ADS*) and Warning sub-System (*WS*) as shown in figure 3. *ADS* includes two units: Animal-Based Unit (*ABU*) and Road-Side Unit (*RSU*). *ABU* is attached to the animal and consists of *GPS* receiver, *DSRC* transmitter, and interface as shown in figure 4. The *European Telecommunications Standard Institute* (*ETSI*) decided to allocate frequency band from 5875 to 5905 MHz for *ITS* [9]. Similarly, we decide to utilize *DSRC* transmitter and receivers that operate with 10MHz band 5885 to 5895 MHz in *CVAAS*. They operate with 33dBm[2]@10MHz transmit power that enable transmitters and receivers to reach communication distance range from 500m to 1000m.

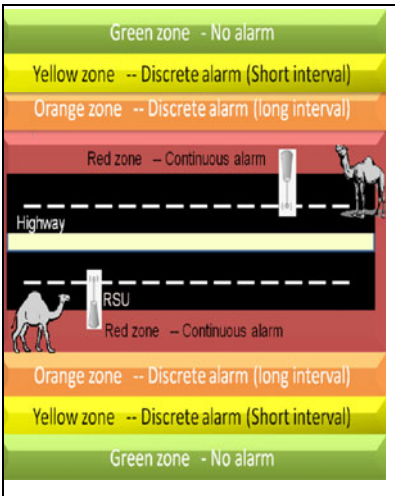
GPS receiver operates to capture key data such as animal's position, velocity, acceleration, heading, etc. The transmitter forwards that key data to the *RSU*. The *ABU*'s interface grants the ability to update the system parameters of both *GPS* receivers and

Table 1. Components in the CVAAS's block diagram

Unit	onfiguration Data	Input	Output	Functions (Tasks)
ABU	GPS fixing time Transmitting interval	PS signals	amel's key data (position, elocity, heading)	Determine camel's key data. Transmit camel's key data
RSU	ones definition	amel's key ata	/arning (activate) message.	Receive camel's key data Forward appropriate warn message to *WU*.
WU	larm's frequency overage area to send SM essage.	ctivate essage	lashing light and alarming, SMS message to drivers.	ctivate warning to drivers.

[2] *dBm* is the measured power referenced to one milliwat.

DSRC transmitter such as the frequency of key data transmission, positioning times based on animal behavior (e.g. more frequent during activity, less frequent when relaxing), packet payload size and message life time. Road-Side units are organized along the highway. Each *RSU* consists of *DSRC* transmitter and receiver, communication unit, processing unit (*Application processor*) and storage area. It stores the highway map for a distance of 100-300m around the *RSU*, and the description of the dangerous zones around it as shown in figure 5. Table 1 outlines the CVAAS's components and their functions.




```
RSU_activate(){
 while(1){
 receive(key_data);
 activate_message->zone=match (key
 data);
if (activate_message->zone != green)
    send(activate_message);
} }

Warning_setup(){
zone= green; // default value
Alarm = silent;
while(1) {
 receive(activate_message)
  switch (activate_message->zone){
  case red: Alarm= continuous; break;
  case orange: Alarm = long_discrete;
  break;
  Case yellow: Alarm=short_discrete;
  break;
  default: Alarm = silent; break;
} } }
```

Fig. 5. Description of dangerous zones

The *RSU*'s receiver gets the key data from *ABU*'s transmitter. The communication unit forwards the received key data to the *RSU*'s processing unit. The processing unit executes a thread that runs the *RSU_activate()* procedure. As soon as, the *RSU_activate()* procedure receives the key data that matches with the description of the dangerous zones it takes the decision to send *activate message* to the *WS*. The *activate message* includes the classification of dangerous zones. The *WS* executes a thread that runs the *Warning_setup()* procedure. When the *Warning_setup()* procedure receives the activate message, it identifies the degree of hazardous and setups the alarming period. For example, red-zone is the most dangerous zone that includes bi-direction lanes and stripes around it with range10-20 meters. If the key data received from the *ABU* matched with the definition of the red-zone then the *RSU* activates the *WS* to produce continuous alarm until receives different data key from the *ABU*. Otherwise, the *WS* will not be activated whereas the key data matched with the definition of the green-zone.

Currently, we are developing a simulation of the *CVAAS* to identify the efficient values of the system's parameters of both *GPS* receivers and *DSRC* transmitter such as the frequency of key data transmission, positioning times and number of RSUs/Km, etc. This simulation will provide some recommendations for system deployment. It enable us to select efficient parameters that considers the system's power consumption and the system's deployment cost.

5 Conclusion

This paper introduced a survey of the developed animal detection and warning systems. It provided recent and numerous reviews of the worldwide technologies that have been used in attempts to reduce *AVC*. Moreover, this paper introduced the design of the camel-vehicle accident avoidance system *CVAAS* in *KSA*. The *CVAAS* took a first comprehensive steps toward a system that will detect camels on the highway and warn drivers as well. The innovation of the *CVAAS* is the careful and intelligent use of the programmable GPS device to detect the camel position, direction and movement. Moreover, *CVAAS* classifies the dangerous zones that enables the warning system to adapt the alarming period.

Acknowledgment. The authors thanks the King Faisal University for sponsoring this project also acknowledge the academic support from our CCSIT College.

References

[1] Al-Ghamdi, A.S., AlGadhi, S.A.: Warning signs as countermeasures to camel-vehicle collisions in Saudi Arabia. Accident Analysis and Prevention 36(5), 749–760 (2004)
[2] Bender, H.: Deterrence of kangaroos from roadways using ultrasonic frequencies: efficacy of the Shu Roo, University of Melbourne, Dept. of Zoology, Road Traffic Authority of South Wales and Transport South Australia, Australia (2001)
[3] Bendix Commercial Vehicle Systems, Bendix Xvision system service data (2002), http://www.bendix.com/downloads/195160.pdf
[4] Bruinderink, G.W.T.A., Hazebroek, E.: Ungulate traffic collisions in Europe. Conservation Biology 10, 1059–1067 (1996)
[5] Carey, M.: Addressing wildlife mortality on highways in Washington. In: Proc. Int. Conf. on Ecology and Transportation, pp. 605–610 (2001)
[6] Clevenger, A.P., Chruszcz, B., Gunson, K.: Highway mitigation fencing reduces wildlife-vehicle collisions. Wildlife Society Bulletin 29 (2001)
[7] de Molenaar, et. al.: Effectiviteit van wildspiegels: een literatuurevaluatie, IBN rapport 362. Instituut voor Bos- en Natuuronderzoek, Wageningen, The Netherlands (1998)
[8] Dodd, et. al.: Evaluations of measures to minimize wildlife-vehicle collisions and maintain permeability across highways: Arizona Route 260, Final Report to Arizona Dept. of Transportation, JPA 01-152 JPA 04-024T (2007c)
[9] (2008), http://www.etsi.org/WebSite/NewsandEvents/2008_09_Harmonized standards_ITS.aspx

[10] Foster, et al.: Use of underpasses by Florida panthers and other wildlife. Wildlife Society Bulletin 23, 95–100 (1995)
[11] Gagnon, et al.: Traffic volume alters elk distribution and highway crossings in Arizona. The Journal of Wildlife Management 71(7), 2318–2323 (2007)
[12] Gagnon, et al.: Using Global Positioning System Technology to Determine Wildlife Crossing structure placement and evaluating their success in Arizona. In: Int. Conf. on Ecology and Transportation, USA (September 2009)
[13] Gordon, K.M., Anderson, S.H.: Motorist response to a deer-sensing warning system in western Wyoming. In: 2001 Proc. Int. Conf. on Ecology and Transportation, pp. 549–558 (2001)
[14] Gordon, et al.: Evaluation of the FLASH (Flashing Light Animal Sensing Host) system in Nugget Canyon. Univ. of Wyoming, Wyoming (2001)
[15] Haikonen, H., Summala, H.: Deer-vehicle crashes: Extensive peak at 1 hour after sunset. American Journal of Preventative Medicine 21(3), 209–213 (2001)
[16] Hirota, et al.: Low-cost infrared imaging sensors for automotive applications. In: Valldorf, J., Gessner, W. (eds.) Advanced Microsystems for Automotive Applications, pp. 63–84 (2004)
[17] Honda Motor Co., Ltd. Intelligent night vision system able to detect pedestrians and provide driver cautions (2004), http://www.all4engineers.de/preview.php?cms=andlng=enandalloc
[18] Huijser, M.P., McGowen, P.T.: Overview of animal detection and animal warning systems in North America and Europe. In: Proc. of the Int. Conf. on Ecology and Transp., pp. 368–382 (2003)
[19] Huijser, et al.: Animal Vehicle crash mitigation using advanced Technology, Phase I: Review, Design and Implementation, Report SPR(076), Western Transportation Inst. Montana State Univ
[20] Wildlife-vehicle collision reduction study. Report to congress, U.S. Department of Transportation, Federal Highway Administration, Washington D.C., USA (2008)
[21] IRD. Wildlife warning system. IRD (International Road Dynamics), Saskatoon, SK, Canada (2002),http://www.irdinc.com/english/pdf/sys/safety/Wildlife_0202.pdf
[22] Sullivan, J.M.: Relationships between Lighting and Animal-Vehicle Collisions, Technical report UMTRI-2009-35, The University of Michigan Transportation Research Inst (2009)
[23] Kistler, R.: Wissenschaftliche Begleitung der Wildwarnanlagen Calstrom, WWA-12-S. Schlussbericht. Infodienst Wildbiologie & Oekologie, Zürich, Switzerland (1998)
[24] Knapp, K.K., et al.: Deer-vehicle crash countermeasure toolbox: a decision and choice resource, F Report DVCIC-02. Midwest Regional Univ. Transp. Center, Wisconsin, USA (2004)
[25] Langley, R.L., Higgins, S.A., Herrin, K.B.: Risk factors associated with fatal animal-vehicle collisions in the United States. Wilderness and Environmental Medicine 17(4) (2006)
[26] McKinney, T., Smith, T.: US93 Bighorn Sheep study: Distribution and trans-highway movements of Desert Bighorn Sheep in northwestern Arizona, Final report to Arizona Department of Transportation. JPA04-032T / KR04-0104TRN (2007)
[27] Pojar, T.M., Prosence, R.A., Reed, D.F., Woodward, R.H.: Effectiveness of alighted, animated deer crossing sign. Journal of Wildlife Management 39, 87–91 (1975)
[28] Putman, R.J.: Deer and road traffic accidents: options for management. Journal of Environmental Management 51, 43–57 (1997)

[29] Ujvari, M., Baagoe, H.J., Madsen, A.B.: Effectiveness of acoustic road markings in reducing deer-vehicle collisions: a behavioral study. Wildlife Biology 10, 155–159 (2004)
[30] Dalef, F.: Reed, Effectiveness of highway lighting in reducing deer-vehicle collisions. Journal of wildlife management 45(3), 721–726 (1981)
[31] Romin, L.A., Bissonette, J.A.: Deer-vehicle collisions: status of state monitoring activities and mitigation efforts. Wildlife Society Bulletin 24, 276–283 (1996)
[32] Scheifele, et al.: Analysis and effectiveness of "deer whistles" for motor vehicles: frequencies, levels, and animals threshold responses. Acoustics Research Letters Online 4, 71–76 (2003)
[33] Rutter, S.M., Bras, R.: Zootec. The integration of GPS, vegetation mapping and GIS in ecological and behavioural studies, vol. 36 (2007), doi:10.1590/S1516-35982007001000007
[34] STS. Roadway animal detection system. Operator's manual and installation guide, RADS-00-007-OM, Revision 4. Scottsdale, Arizona, USA (October 16, 2002)
[35] Turner, L.W., et al.: Global Positioning Systems (GPS) and Grazing Behavior in Cattle Livestock Environment VI. In: Proc. of the 6th Int. Symp., Louisville, Kentucky, USA, pp. 640–650 (2001)
[36] Ward, A.L.: Mule deer behavior in relation to fencing and underpasses on Interstate 80 in Wyoming, Transport, Research Record 859:8-13, National Research Council, Washington (1982)
[37] What we are doing to reduce vehicle/Wildlife collisions, Washington State Department of Transportation (2010), http://wsdotblog.blogspot.com/2010/02/what-were-doing-to-reduce.html

A Hierarchical Architecture for Semantic Sensing Information Representation

Rajani Reddy Gorrepati[1], Hyeon Park[2], Dong-Hwan Park[2], and Do-Hyeun Kim[1]

[1] Dept. of Computer Engineering, Jeju National University, Jeju, Republic of Korea
rajani_reddi@hotmail.com, kimdh@jejunu.ac.kr
[2] Dept. of Office, Electronics & Telecommunication Research Institute, Republic of Korea
hpark@etri.re.kr, dhpark@etri.re.kr

Abstract. Wireless sensor networks are used in various applications in several domains for measuring and determining physical phenomena and natural events. Wireless sensor networks can support to observe characteristics of physical objects and features of natural incidents. This paper proposes a hierarchical architecture of semantic information for heterogeneous sensor data representation in vehicle. This architecture includes 5 layers which have a physical layer, an event layer, a semantic layer, an awareness layer, and service layer. This architecture provides various real-time sensing data such as velocity, acceleration, temperature and steering wheel position in vehicles. Information of this architecture supports to detect sudden unexpected events such as accidents on the vehicle environment. We develop ontology description for sensor data and use the Protégé for processing which includes querying and inference over sensor data.

Keywords: Context-aware, Vehicle, Semantic sensor information.

1 Introduction

Acoustic sensors, such as microphone array, it can collect an acoustic signals to identify the type and localize the position of a moving vehicle. Acoustic sensors can be used in sensor network applications such as traffic monitoring, tracking vehicles and driver condition [1]. Almost every moving vehicle makes some kind of noise; the noise can come from the vibrations of the running engine, bumping, and friction of the vehicle tires with the ground and wind effects and soon. When vehicle travelling at different speeds, under different road conditions, or with different acceleration, will come emits at different noise patterns.

Vehicle noise is a kind of stochastic signal. The recording microphone array is fixed, and the vehicle running conditions usually short sound duration can be recorded [2]. Acoustic vehicle classifications mainly focus on signal processing and pattern recognition techniques. The vehicle generated the different sounds recorded by microphone array. In a conventional pattern recognition framework features are extracted from the sensor data and these features are filtered according to perceived information content, prior to use in classification.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 159–166, 2011.

The noise emitted from a rolling time as a result of its interaction with the road surface. The tire noise is the two types of vehicles noises: vibration and air noise [3]. The vibration noise is caused by the contact between the tire threads and the pavement texture. The air noise is generated by the air being forced out of the rubber blocks of a tire [4]. In the driving direction of the car, the road and the tire forms a geometrical structure that amplifies the noise generated by the tire-road interaction. Inside and outside vehicle sensor information systems of the complexity are already common place for voice communications.

Inside and outside vehicle sensor information systems of the complexity are already common place for electronic engine management, voice communications and parking sensor systems. Monitoring driver behavior of steering, braking, accelerating and atmospheric conditions may allow parking sensor system to give early warning of dangerous driving situations. We have equipped a vehicle with several sensing, computing and communicating systems. The contextual information is about the car, driver and the environment. We measure various air gases with the sophisticated sensing equipment. Control unit of the vehicle to extract velocity, acceleration, temperature, steering wheel position etc. A GPS sensor gives the geographical position of the vehicle accuracy. The context driving, it consists of the environment to all objects and to maintain traffic conditions.

Rules can be used for capturing the semantic relationships and dependencies between ontological properties or between ontologies and other domain predicates. RDF, OWL and SWRL are semantic representation languages with high degree of expressiveness that are adequate for modeling ontologies and rules. Ontologies and rules can be reasoned by logic inference engines. All these logics mechanisms can be utilized for generating new knowledge that can be used for labeling unknown objects.

In this paper, we propose a hierarchical architecture for sensor information description using sensor data so that it becomes meaningful in vehicle sensor networks. We describe each information of layers using sensor data representation and attached semantics to it. We study specifically the hierarchical architecture information processing service enabler through a semantic and context information and demonstrate that this can be achieved using ontology technologies. This architecture supports to process sensor data so that high level information can be extracted and shared among different applications.

The organization of this paper is as follows: Section 2, we introduce the related work in section 2. In section 3, we propose a hierarchical architecture for semantic sensor information for vehicle. And this section covers sensor information description and semantic information representation reasoning. The concluding remarks are given in section 4.

2 Related Works

Many research works of academic and industrial communities has been made on cars using several sensors. The driver plays an important role in smart cars. Machine learning dynamical graphical models, such as HMM [5], Gaussian Mixture Modeling (GMM) [6] and the Bayesian network can be applied for modeling and recognizing driver behaviors [7]. Reichardt establishes a cooperative and multi-level transport

virtual sub-center by vehicle-to-vehicle communication and vehicle-to-center communication [8]. It provides local situation information, traffic and infrastructure status information via a dedicated infrastructure to support vehicle communication link. Rusconi designs an intelligent cooperative system, which provides real-time information from other vehicles in the vicinity and roadside equipment to improve driver's responses [9]. Hoch develops BMW Online and driver assistance systems, supporting lane change warning and parking assistant [10]. Mercedes-Benz is developing an intelligent driver assistance system that utilizes stereo cameras and radar sensors to monitor the surroundings around the car [11]. Daniel illustrates the concept of context widgets with the beginnings of a widget library and they developed for sensing presence, identity and activity of people and things [12].

Wang establishes two papers of computer vision to intelligent vehicle highway systems. First paper developed of a system for monitoring traffic scenes using video in-formation. This information is to detect quickly incidents of stalled vehicles and accidents. Second paper, It is developed vision as a sensor technology for vehicle control [13]. Nicholas True proposes a camera-based system would use computer vision algorithms for detecting vacant parking spaces. This algorithm uses a combination of car feature point detection and color histogram classification to detect vacant parking spaces in static overhead images [14]. Yun established the portable stolen vehicle recognition system is designed for police officers who work on the roadside, so the convenience is very important in design of the system operation. This system also works real-time outdoor experimental and it can work on the road side to recognize the license plate and verify if it is stolen vehicle or not [15].

3 A Hierarchical Architecture for Semantic Sensor Information in Vehicle Sensor Networks

The environment of the car context provides sensing capabilities including light sensors, barometric pressure, photo-sensitive light, humidity, temperature, and microphone sensors. The parking sensors system was used to estimate the distance between two adjacent cars, and to immediately detect the accidents.

A microphone sensor was used to receive the vocal command from a driver including weather, music on, music off and parking space. It was defined for a driver to access services by voice rather than by hand.

In this figure 1, we show the five layers architecture. The hierarchical architecture has been divided into physical sensor layer, event layer, semantic layer, awareness layer and service layer. The object from the physical layer is to make the semantic layer with the context-awareness layer and transfer the latter to the service. Moreover, the context layers guarantees the independence of a service and device and mediates the context information provided by a sensor to a service. If user turns on the sensing device, the context provides which are suitable from the device.

Many vehicles include in-vehicle communication systems, such as a cell phone connection systems and voice driven navigational systems. These systems may work with the vehicle's sound system to provide the driver with hands free communication capability while in the vehicle [16]. Many such communication systems must detect audible information from the driver or another user within the vehicle. One or more

microphones fixed in a position that focuses on the expected area of a driver's head detect the audible information. Provided the driver's head is in the expected area, focusing the fixed microphone on the expected area reduces undesirable outside noise and facilitates retrieving a quality sound signal from the driver. That is, focusing the microphone on the source of desired sound reduces picking up undesirable noise. Since the position of the driver fluctuates depending on, for example, the driver's height and seated position, fixed microphones cannot focus on too small an area. That is, relative to the fixed microphones, the position of the driver's head may move between a range of heights and distances. Accordingly, the fixed microphone must balance receiving a quality sound signal with accommodating different driver positions. This architecture of the system is shown in figure 2.

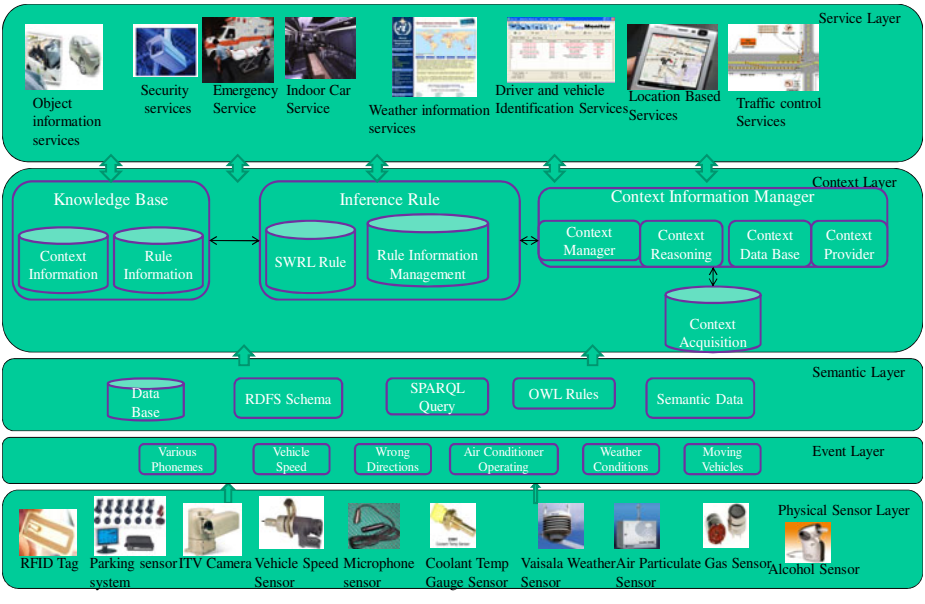


Fig. 1. Hierarchical architecture of semantic sensor information in vehicle

A semantic information representation method for positioning a microphone includes determining the position of an occupant in a passenger compartment of a vehicle. The example method directs the microphone toward the position. Another method for positioning a microphone includes generating at least one signal corresponding to at least one measurement and determining a position of an occupant in the passenger compartment of a vehicle using the at least one signal. This method adjusts at least one microphone based on the position. This method may include adjusting a function of an occupant safety system based on the position.

Our system consists of the sensing devices in the physical layers. The data provided from sensor consists in semantic layers, obtained semantic data divided into aware-ness data and finally related services provided by the service layer. The parking sensor system is attached 6 sensors (2 sensors in front and 4 sensors in rear bumper)

giving good coverage of both front and rear end of your vehicle when parking. Wireless LED digital display with Beeper and volume ON/Off switch [17, 18]. This digital dis-play is attached inside the vehicle and also it shows the exact distance display indicating orientation of obstacles and safe/warning/stop zone and also attached control box in the back side of the vehicle. This system is a safety device that accurately detects all objects, as children, pets, another autos etc. that are behind your vehicle while you back up. The beeping as well as wireless LED digital display inside the car warns the driver of the danger and prevents a potential accident.

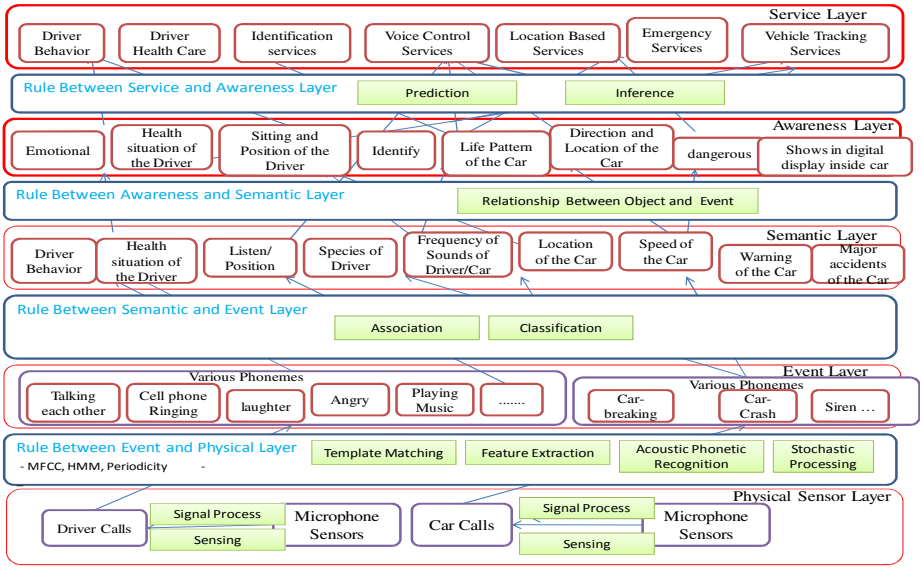


Fig. 2. Semantic information representation based on microphone in vehicle

This system function is to measure the velocity of vehicles travelling inside a tunnel, highway roads, bridges etc. It calculates the distance between two vehicles from the velocities and difference in transmit time off from the two vehicles. If the distance between two vehicles is shorter than a previous distance then it displayed a warning message to the driver on the information board. This architecture of the system is shown in figure 3.

This system is automatically activated when you engage reverse gear. Small sensors are located at the rear of the vehicle, they send and receive ultrasonic waves, which bounce off obstacles and alert you to their presence. There is an internal buzzer that gradually increases in frequency as you approach the object and shows you the distance in the LED display. This system can support services such as object directory service, object information service, location based services and secret services. This system can be reduced number of accidents in the traffic.

The vehicle speed sensors is mounted on the output shaft of the transmission sends electrical pulses to the computer, pulses which are generated through a magnet that spin a sensor coil [19]. When the vehicle speed increases, the frequency of the pulses

also increases. The automobile's engine contains a speed sensor. This speed sensor automatically sends the information to the computer as to how fast the car is traveling at the moment of driving. The engine's speed sensor is craftily de-signed to be able to record the rate at which the vehicle's speed sensor is spinning. The speed sensor is made up of a toothed metal disk that is basically mounted onto the vehicle speed sensor and a stationary detector that functions to cover a magnetic coil through which the current passes through. A wave of pulses in the current is created. The computer readily calculates the speed at which the engine is traveling through by means of the numbers of the vehicle speed rotation. These rotations determine the frequency of the pulses as created from the engine speed sensor. The vehicle speed sensors are also measured speed of the vehicle, acceleration, distance and brake testing and also find out the position of the moving vehicle.

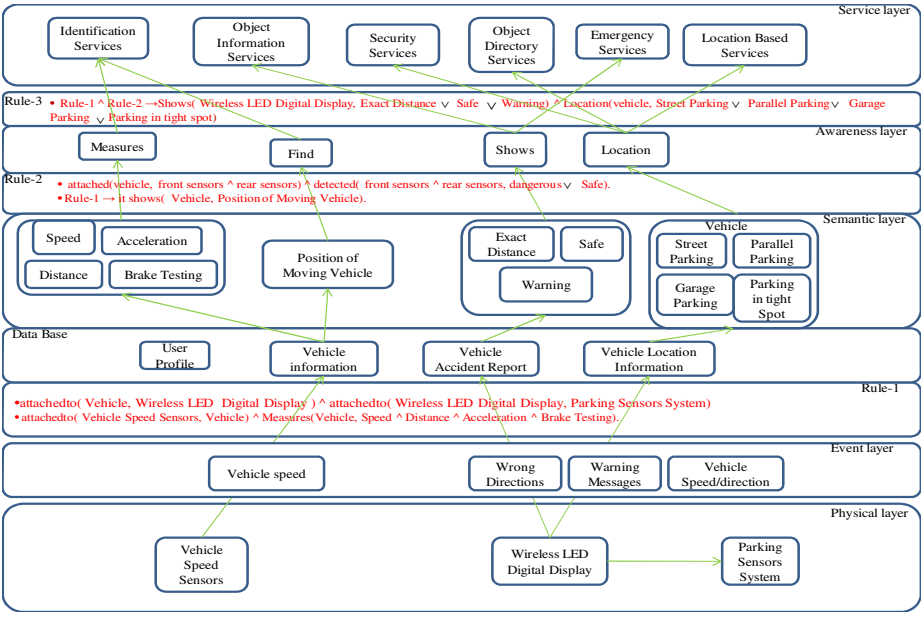


Fig. 3. Semantic information representation based on speed sensor and parking sensor system in vehicle

We implement and construct taxonomy an ontology development tool, called protégé for building and editing the ontology. The knowledge representation language for modeling the various data types of physical sensor, Event data is OWL-DL. We manually add classes to the ontology by creating event data and physical sensor classes and all their sub-classes. The class hierarchy is called manually created classification hierarchy and is shown in figure 4.

As a validation tool, we used Pellet because of its strong reasoning capabilities and interoperability with protégé. The manually created class hierarchy is fed to Pellet whose main responsibility is to automatically compute the inferred class hierarchy (called asserted ontology) based on the description of classes and relation-ships.

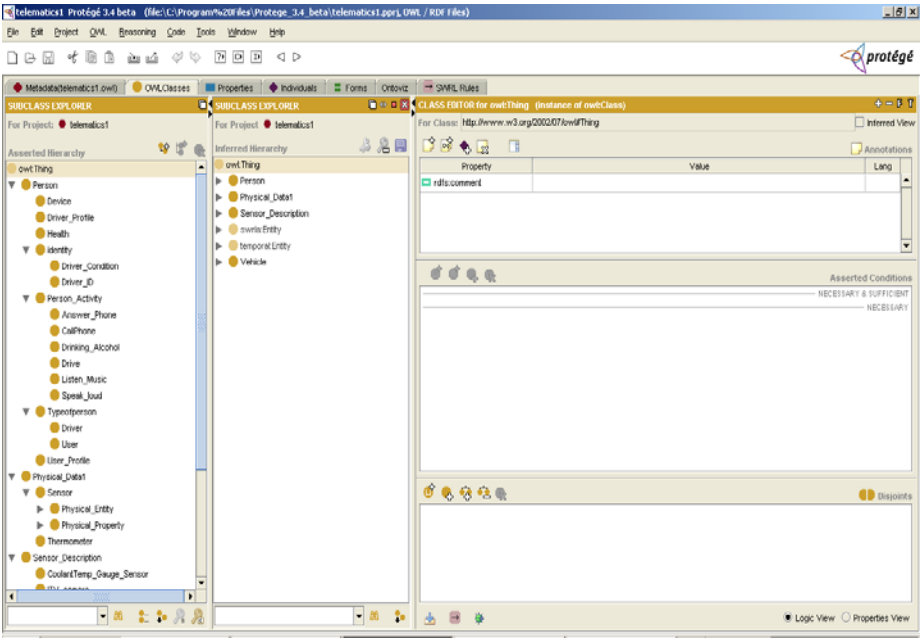


Fig. 4. Ontology structure of hierarchy semantic information modeling using protégé

4 Conclusions

This research attempts to build vehicle environments from the view of context awareness. We propose a hierarchical architecture for sensor information description using parking sensor system, microphone, and speed sensor in a vehicle sensor networks. The proposed architecture using microphone is possible to detect sudden unexpected events such as accidents. And, the proposed architecture using parking sensor system and speed sensor provides safety measures and offering prompt and accurate traffic information. Accordingly, the semantic information in vehicle sensors is related to physical environments and the ontology includes the description of weather information, accidents information, traffic information, inside and outside car information and the road safety conditions etc.

Acknowledgments. This work was supported by the Industrial Strategic Technology Development Program funded by the Ministry of Knowledge Economy(MKE, Korea) [10033915, Adaptive fusion technology for large-scale sensor node based intelligent surveillance system]. Corresponding author; DoHyeun Kim (e-mail: kimdh @jejunu.ac.kr.).

References

1. Forren, J.F., Jaarsma, D.: Traffic monitoring by tire noise. In: Proc. IEEE Conf. on Intelligent Transportation System, Boston, pp. 177–182 (1997)
2. Sandberg, U.: Tyre/road noise - myths and realities. In: Proceedings of the 2001 International Congress and Exhibition on Noise Control Engineering, Hague, Netherlands, pp. 27–30 (2001)
3. Wellman, M.C., Srour, N., Hills, D.B.: Acoustic feature extraction for a neural network classifier Army Research Laboratory, Tech. Rep. ARL-TR-1166 (1997)
4. Thomas, D.W., Wilkins, B.R.: The analysis of vehicle sounds for recognition. Pattern Recognition 4(4), 379–389 (1972)
5. Oliver, N., Pentland, A.P.: Graphical Models for Driver Behavior Recognition in a Smart Car. In: Proc. lEEE Intell. Vehicles Symp., pp. 7–12 (2000)
6. Miyajima, C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K., Itakura, F.: Driver modeling based on driving behavior and its evaluation in driver identification. Proc. IEEE 95(2), 427–437 (2007)
7. Kumagai, T., Akamatsu, M.: Prediction of human driving behavior using dynamic bayesian networks. IEICE- Trans. Inf. Syst., E89-D (2), 857–860 (2006)
8. Reichardt, D., Miglietta, M., Moretti, L., Morsink, P., Schulz, W.: CarTALK 2000: safe and comfortable driving based upon inter-vehicle-communication. In: IEEE Intell. Vehicle Symp., vol. 2(17-21), pp. 545–550 (2002)
9. Rusconi, G., Brugnoli, M.C., Dosso, P., Kretzschmar, K., Bougia, P., Fotiadis, D.I., Salgado, L., Jaureguizar, F., de Feo, M.: I-WAY, intelligent co-operative system for road safety. In: IEEE Intell. Vehicles Symp., vol. 13(15), pp. 1056–1061 (2007)
10. Hoch, S., Althoff, F., Rigoll, G.: The Connected Drive Context Server-flexible Software Architecture for a Context Aware Vehicle. In: Advanced Microsystems for Automotive Applications, Springer, Berlin (2007)
11. Mercedes-Benz to Present Driver Assistance Systems of Tomorrow at the ITS World Congress (2007), http://e-carzone.com/news/auto_5911Shtml
12. Daniel, S., Anind, K.D., Gregory, D.A.: The Context Toolkit: Aiding the Development of Context-enabled Applications. In: Proc. SIGCHI Conf. on Human Factors in Computing Systems, pp. 434–441 (1999)
13. Wang, F.Y., Zeng, D., Yang, L.Q.: Smart cars on smart roads: an IEEE intelligent transportation systems society update. IEEE Perv. Comput. 5(4), 68–69 (2006)
14. True, N.: Vacant parking space detection in static images. University of California
15. Chung, Y.-C., Chuang, Y.-J.: The design of portable stolen vehicle detector. Institute of computer science and information engineering, National Taiwan Normal University
16. Haddad, K., Benoit, V.: Capabilities of a beamforming technique for acoustic measurements inside a moving car. In: International Congress and Exposition on Noise Control Engineering (2002)
17. Wireless Parking Sensors, http://www.ehow.com/how-does_5033713_wireless-parking-sensors-work.html
18. LED digital display, http://www.stee.stengg.com/newsrm/Vol21No2/Solutions1.pdf
19. Vehicle Speed Sensors, http://www.ehow.com/how-does_4963080_does-cars-speed-sensor-do.html

An Integrated Process Method of Sensing Data for Multi-sensor Networks

Jun-Pil Boo[1], Do-Hyeun Kim[1], Hee-Dong Park[2], and Paul Roe[3]

[1] Faculty of Telecommunication & Computer Engineering, Jeju National University, Jeju-do, 690-756, Korea
{bip0389,kimdh}@cheju.ac.kr
[2] Department of Information & Communication, Korea Nazarene University, Cheonan-city, ChoongNam, 330-718, Korea
hdpark@kornu.ac.kr
[3] Microsoft QUT eResearch Centre, 126 Margaret Street Level 4, Queensland University of Technology, GPO Box 2434 BRISBANE QLD 4001
p.roe@qut.edu.au

Abstract. Recently, SWE (Sensor Web Enablement) working group of OGC (Open General Consortium) is carrying out a research on providing real-time sensor data and video collected in sensor networks through the Internet Web. However, the research does not deal with mobile objects such as cars, trains, ships, and persons. Therefore, we present an integrated processing method of video, sensor data, and location information collected in mobile sensing nodes for tracking and surveillance of mobile objects. The proposed method integrates and processes heterogeneous data from GPS devices, sensor networks, and video devices. Based on the proposed method, we designed and implemented a combine adaptor which receives the context data through a common interface and creates combined messages after parsing and queuing the data. We verified the proposed method works well with our integrated adaptor. The proposed method is expected to be helpful to develop various tracking and surveillance based on location.

Keywords: **M**obile object, Integrated processing, GPS, Integrated data processing, Location based services.

1 Introduction

Recently, various ubiquitous applications are being developed based on GPS and sensor networks. GPS is mainly used for location-based applications, and sensor networks are deployed for collecting context information such as temperature, humidity, atmospheric pressure, and so on. Additionally, many researches on sensor web technology are carried out by OGC (Open Geospatial Consortium), Microsoft, etc. The sensor web or sensor map technology is for displaying and monitoring sensor data repositories connected to the Internet or World Wide Web. The current Sensor Web research focuses on providing real-time sensor data and video collected in sensor

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 167–174, 2011.

networks. Meanwhile, more studies are required on mobile sensor web technology dealing with data from mobile sensor networks composed of cars, ships, persons, and so on.

This paper proposes an integrated processing method for video and sensor data based on mobile node's location in mobile sensor networks. The proposed method integrates and processes heterogeneous data from GPS, sensors, video devices to provide information of moving objects in mobile sensor networks. Based on the proposed method, we designed and implemented a combine adaptor which provides a common interface for the heterogeneous data and creates combined messages after parsing and queuing them.

The rest of this paper is organized as follows: First, we review some of related work in Section 2, then Section 3 presents the proposed algorithm that processed heterogeneous data such as GPS, video, and sensor data. Finally, the conclusion is given in Section 4.

2 Related Works

LBS (Location-Based Service) using GPS as an information service, accessible with mobile devices have the ability to make use of the geographical position of the mobile devices. LBS include services to identify a person or object with location information. The location information can be provided to users with other context data like sensor data. LBS technologies include LDT (Location Determination Technology), mobile communication technology, LEP (Location Enabled Platform), and LAP (Location Application Program) [1,2].

A WSN (Wireless Sensor Network) is a wireless network composed of spatially distributed sensors, actuators, RFIDs, etc. to provide users with context information such as temperature, humidity, pollutants, or pressure [3, 4].

Sensor web is an environment in which users can discover and monitor various sensors, instruments, image devices, and sensor data repositories through the Internet Web. SWE standardization has been in progress by OGC to provide common bases for sharing and utilizing heterogeneous sensor data. SWE standards include SensorML, Observation & Measurement, sensor observation service, and web notification. Sensor web can be used for real-time detection and early warning system for forest fires, earthquakes, tidal waves, and so on [5,7].

Our previous research investigated a general architecture in vehicular sensor networks, and the problem of intermittently connecting from cars or other mobile devices to the Internet. This work did not address multimedia sensors issues, such as a camera, a microphone, multiple sensor, etc[8,9].

3 Location-Based Integrated Processing Algorithm

We propose an integrated processing algorithm for video and sensor data based on location information of mobile objects. The proposed method creates a combined frame, which is composed of sensor and GPS data of a mobile object, and then transmits the frame to applications or middleware. Fig. 1 shows the processing procedures of video and sensor data based on location. The proposed method provides

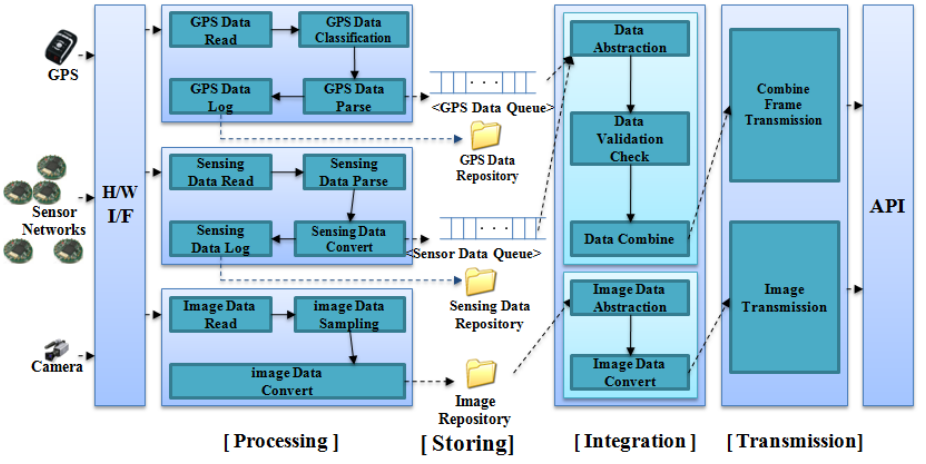


Fig. 1. Video and sensor data integration in the proposed algorithm

a common communication interface to receive the heterogeneous data from cameras, sensors, and GPS. After processed by the processing module, each data is transferred to data integration module and applications in turn.

3.1 GPS Data Processing

GPS data processing module receives and processes real-time GPS data. This module includes GPS data reader for receiving GPS data through a common input interface, GPS data classifier, GPS data parser/converter, and GPS data logger. Additionally, GPS data queue and repository are also required to store the data.

As shown in Fig. 2, the GPS data processing sequence is repeated periodically. The GPS data processing module communicates with GPS receivers through a common communication interface and makes the GPS data reader read primitive GPS data in buffering. Because the primitive data are composed of various data, it is required to extract necessary data from the primitive data by using the classifier. The extracted data are parsed and converted to numerical data which in turn will be transferred to data integration module. The processed GPS data is buffered or stored in GPS data queue and repository as a temporary variable. The primitive GPS data can be used again by using log information, not through GPS receiver.

3.2 Sensor Data Processing

Sensor data processing module receives and processes real-time context data generated by sensor nodes. Like GPS data processing module, the sensor data processing module is composed of sensor data reader, parser/converter, and logger. Sensor data queue and repository are also added to the sensor data processing module. Fig. 3 shows flow chart of the sensor data processing procedure. Unlike GPS data, necessary information is extracted from primitive sensor data by analyzing data fields of the primitive sensor data format. Then, the extracted data are converted to numerical data by sensor data converter. The processed sensor data with a format of temporary variable are buffered or saved in sensor data queue and repository.

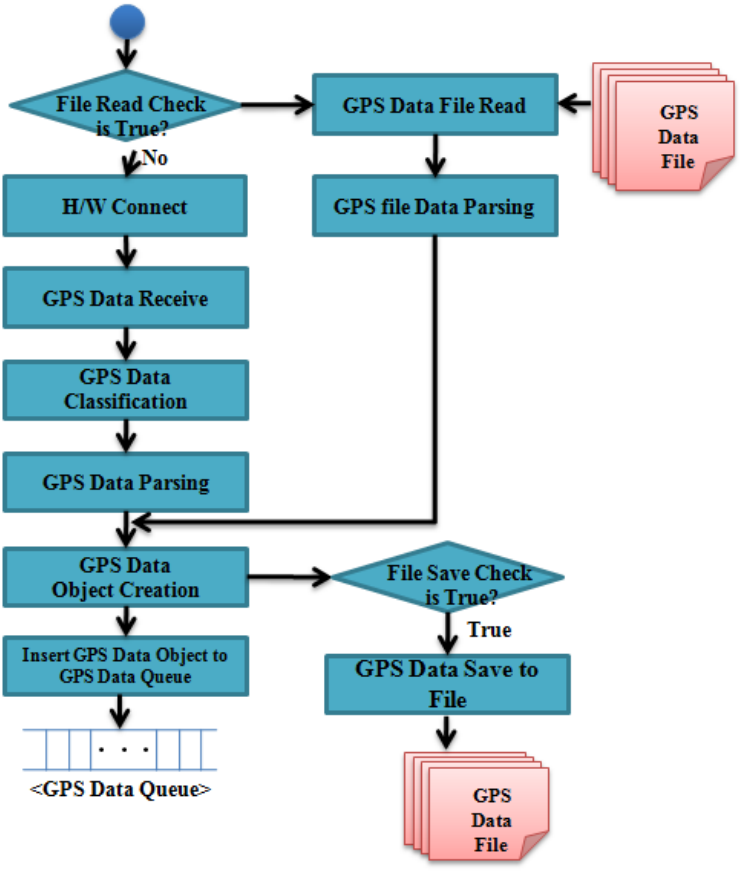


Fig. 2. Flow chart of GPS data processing

3.3 Video Data Processing

Video data processing module receives video data from cameras through the common interface, and processes the data. This module consists of video data reader, video data sampler, and converter. Queue and repository for the video data also added to the module. The video data sampler periodically extracts a sample image from the video data. The sample image is given a sequence number by the converter.

3.4 Combined Frame Creation and Transmission

The proposed algorithm makes up a combined frame composed of GPS and sensor data, and then transmits the frame to middleware or applications by transmission module. The combined frame includes ID information of a sensor node. As mentioned earlier, video data have different characteristics from GPS and sensor data. Therefore,

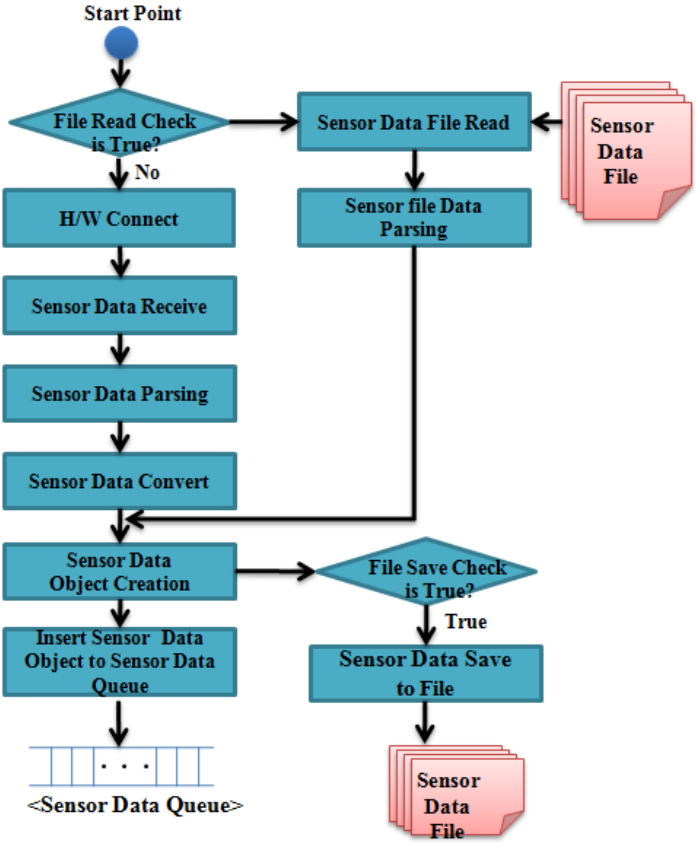


Fig. 3. Flow chart of sensor data processing

video data cannot be a constituent part of the combined frame. Instead, the proposed algorithm fuses video data and the combined frame by providing both data in parallel The consistency and synchronization between the video data and the combined frame can be acquired with the ID information of the combined frame.

Fig. 4 shows the flowchart of creating and transmitting a combined frame, and transmitting an image. Before creating a combined frame, as shown in Fig. 4, validation check for GPS data and sensor data is performed. Fig. 5 shows the procedure of data validation check in detail. To make a combined frame, the data validation check validates if GPS data or sensor data exists or not, after checking errors of each data. Depending on the results of data validation check, the proposed algorithm defined three types of message modes. The message mode becomes one of header fields of a combined frame to indicate which of GPS and sensor data, the combined frame. Then we use a voltage collected each sensor node for verifying the reliability of sensor data.

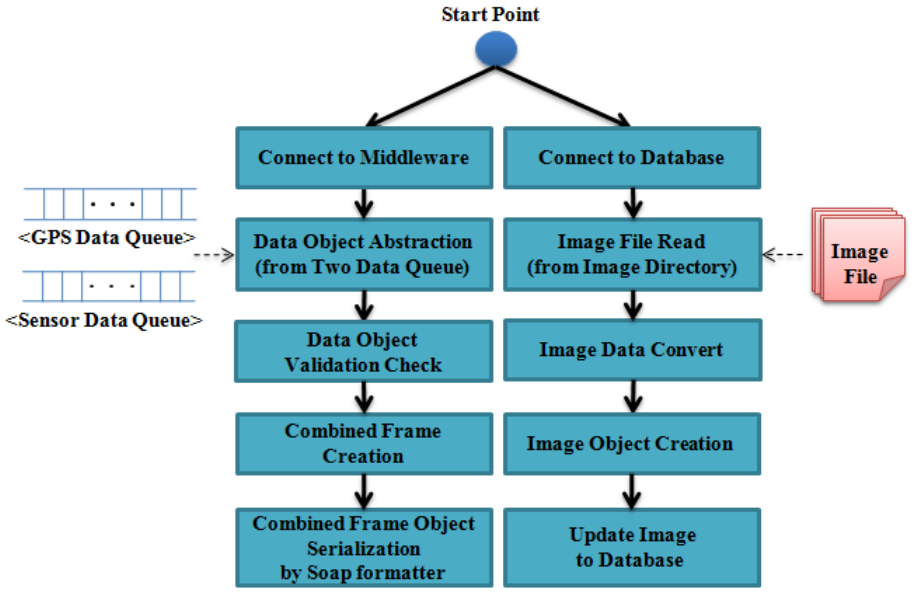


Fig. 4. Flow chart of combined frame creation and transmission, and image transmission

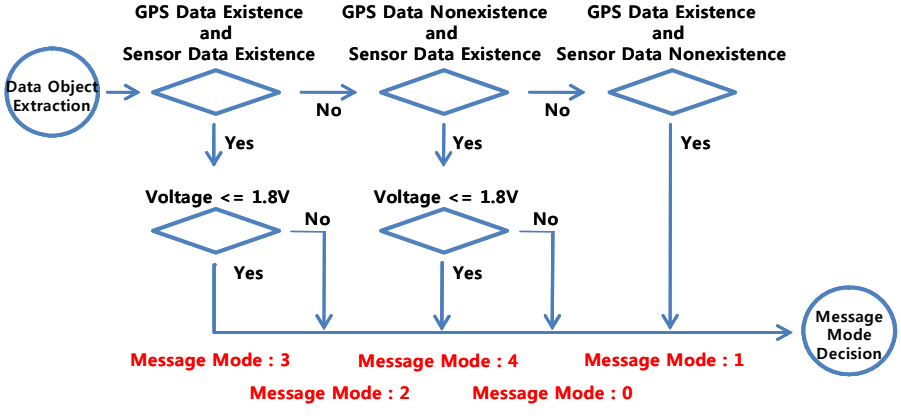


Fig. 5. Data validation check

As mentioned earlier, each of GPS and sensor data is saved as a temporary variable with a predefined format. The format of GPS data variable is composed of latitude, longitude, altitude, GPS time, direction indicator (N/S and E/W), the number of satellites used, status, and others. And the format of sensor data variable is composed of group ID, channel, and node ID, temperature, humidity, power intensity, and others. Fig. 6 shows the structure of a combined frame. The combined frame integrates GPS data and sensor data with header fields such as node ID, Region ID, and message mode. The node ID is provided by applications to identify mobile

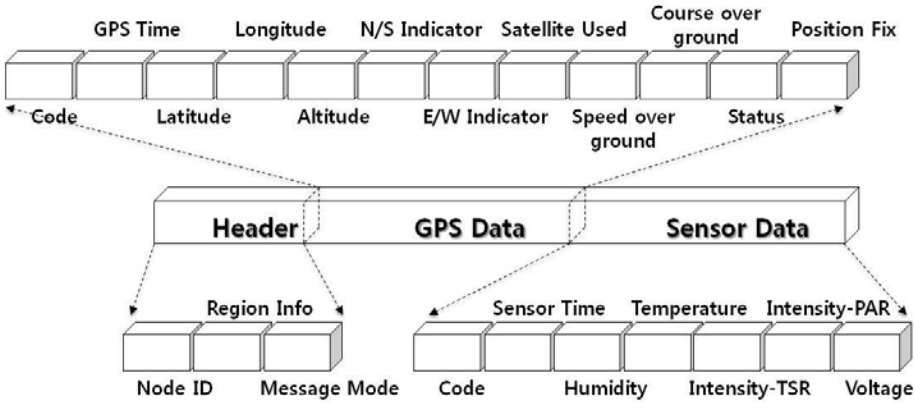


Fig. 6. Architecture of a combined frame

objects which the data of a combined frame are about. And the region ID indicates in which region the mobile object exists.

4 Conclusions

Sensor Web technology provides real-time sensor and video data collected in a sensor network through Internet web. Yet, it doesn't deal with mobile sensor networks. Therefore, this paper proposes an algorithm of location-based video and sensor data integration for mobile sensor networks. The proposed algorithm receives, processes, and integrates heterogeneous data such as GPS, video, and sensor data. We verified the proposed algorithm by implementing an integrated adaptor which has functions of receiving GPS, video, and sensor data through a common interface, parsing and queuing them, and creating and transmitting a combined frame composed of GPS and sensor data.

Acknowledgments. This work was supported by the Industrial Strategic Technology Development Program funded by the Ministry of Knowledge Economy(MKE, Korea) [10033915, Adaptive fusion technology for large-scale sensor node based intelligent surveillance system]. This research was supported by the regional technology innovation program of the Ministry of Knowledge Economy, Republic of Korea. Corresponding author; DoHyeun Kim (e-mail: kimdh@jejunu.ac.kr.).

References

1. Wang, S., Min, J., Yi, B.K.: Location Based Services for Mobiles: Technologies and Standards. In: IEEE International Conference on Communication (2008)
2. Küpper, A.: Location-Based Services: Fundamentals and Operation. Willey, West Sussex (2005)

3. Römer, K., Mattern, F.: The Design Space of Wireless Sensor Networks. IEEE Wireless Communications 11, 54–61 (2004)
4. Culler, D.: Overview of Sensor Network. Computer (2004)
5. Chu, X., Buyya, R.: Service Oriented Sensor Web. In: Mahalik, N.P. (ed.) Sensor Network and Configuration: Fundamentals, Standards, Platforms, and Applications. Springer, Germany (2007)
6. Bott, M.: OGC® Sensor Web Enablement: Overview and High Level Architecture. Open GIS Consortium (2006)
7. Open Geographic Information System, http://www.opengis.org/
8. Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Shih, E., Balakrishnan, H., Madden, S.: CarTel: A Distributed Mobile Sensor Computing System. In: Proceedings ACM SenSys (2006)
9. Eriksson, J., Balakrishnan, H., Madden, S.: Cabernet: Vehicular Content Delivery Using WiFi. In: Proceedings of the 14th ACM international conference on Mobile computing and networking (2008)

Relevance Feedback for Surveillance Video Retrieval at Object Level

Thi-Lan Le

MICA, Hanoi University of Science and Technology,
Hanoi, VietNam
Thi-Lan.Le@mica.edu.vn

Abstract. Object retrieval plays more and more important role as the number of video surveillance systems and the amount of stored data drastically increase. We address in this paper the specific part of retrieving objects of interest within surveillance video sequences problem: relevance feedback. In order to allow users to interact with retrieval system, we propose two relevance feedback methods at object level. These methods take into account appearance as well as temporal aspects of moving objects in surveillance video sequences. That feature is the main difference between our work and previous works. Experimental results on real surveillance video sequences captured in a metro station have proved the performance of two proposed methods.

1 Introduction

We address in this paper the problem of retrieving objects of interest within surveillance video sequences. Nowadays, such solutions for content-based retrieval of surveillance data are ever more required as the number of video surveillance systems and the amount of stored data drastically increase. The main reasons for searching video surveillance data are forensic, i.e. to look for video evidence after an incident occurred. Figure 1 shows the position of surveillance indexing and retrieval in a complete video surveillance system. Videos coming from cameras will be interpreted by the video analysis module. This module is divided into two levels: object level (consisting of object detection, object classification, object tracking in general) and event level (consisting of event recognition). There are two modes for using the output of analysis module. In the first mode, the corresponding alarms are sent to security staffs to inform them about the situation while in the second one the analyzed results are stored in order to be used in the future. Video indexing and retrieval relate to the second mode. Corresponding two levels of video analysis module, surveillance video indexing and retrieval approaches can be grouped into two categories: object level category and event level one. The reader is invited to read the paper [1] for more information.

We have previously proposed a general framework [2], [3] for video surveillance indexing and retrieval. This framework is based on the hypothesis that videos are partially indexed thanks to existing work in video analysis and video surveillance such as object tracking and event recognition. The proposed framework allows to retrieve both objects and events of interest. However, relevance feedback is not

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 175–182, 2011.

considered in the previous version of this framework. In this paper, we propose two new methods for surveillance indexing and retrieval at object level.

The rest of this paper is organized as follows: In section 2, we focus on analyzing existing works for relevance feedback surveillance video indexing and retrieval. Section 3 aims at describing the proposed approach. Experimental results on real surveillance video sequences are presented and analyzed in section 4. We conclude and discuss future work in section 5.

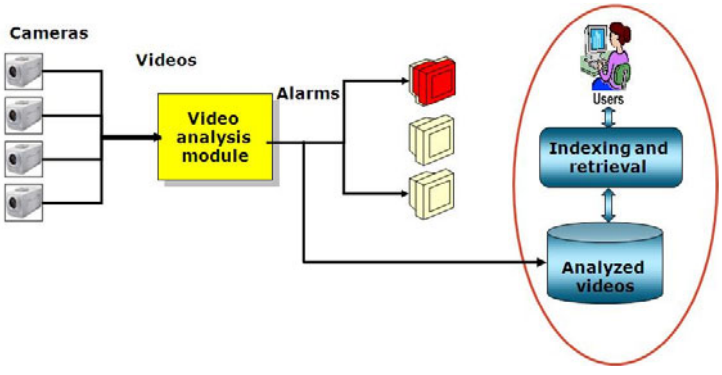


Fig. 1. Surveillance video indexing and retrieval position in a complete video surveillance system. Videos coming from cameras will be interpreted by the video analysis module.

2 Related Works

The relevance feedback is a technique that allows the system to learn feedback from users and adapt to better meet the requests of users [4]. While many works have been proposed for information retrieval in general and image retrieval in particular based on relevance feedback [4], very few works have been done for surveillance video retrieval [1] because of the difficulty in object spatial-temporal aspect modeling. In general, objects in surveillance video sequences appear in the scene in a certain time (or certain frames). At each frame, we can capture an instance of their appearance. Up to now, there are two main significant works have been presented: the work of Meessen et al. [5] and that of Chen et al. [6].

Meessen et al. [5] present a relevance feedback method combining multiple instance learning (MIL), support vector machine (SVM) and active learning. In this work, relevance feedback method is considered as a multiple instance problem because when the user is asked to judge entire video frames, while the learning is achieved on the object regions they consist of. Based on the fact that most existing solutions assume the user is looking for only one single region, the authors have proposed a system called "MILES- Multiple-Instance Learning via Embedded Instance Selection" to solve the multiple-instance problem without any assumption on the number of region to consider for the target class. According to the authors, this solution is suitable for surveillance video because in surveillance video the present of several objects in a frame might be required to define target surveillance events. The advantage of this work is that video surveillance indexing and retrieval module does

not require much effort of video analysis module. The video analysis module needs to provide only object detection. However, this work does not take into account object temporal aspect and object tracking results. While Meessen et al. [5] work with relevance feedback method at object level, Chen et al. [6] introduce a method at event level. Given a set of raw videos, the objects are tracked and the corresponding trajectories are modeled and recorded in the database. Several spatio-temporal event models such as single vehicle and two vehicle events are then constructed. In the learning and retrieval phase, with the top returned sequences, the user provides feedback to the relevance of each video sequence. For learning and retrieval, the neural network for time series prediction is adapted to fit the specific needs of event identification for video data. The authors show their experiments on live transportation surveillance videos. In order to use this work, the video analysis module has to support object detection, object classification and object tracking. The main drawback of this work is that object appearance is not considered.

As we explain, previous works for relevance feedback surveillance video retrieval do not take into account a special feature of moving objects in surveillance video sequences: variation of object appearance (appearance aspect) during a certain time (temporal aspect).

3 Proposed Approach

Before describing our proposed approaches, we define here several notations and definitions.

Definition 1: Object region is a region determined by a minimal bounding box in a frame where object is detected.

Objects in video surveillance are physical objects (e.g. people, vehicles) that are present in the scene at a certain time. In general, they are detected and tracked in a large number of frames. Consequently, an object is represented by a set of object regions. Due to errors in object detection, using all these object regions for object indexing and retrieval is irrelevant. Moreover, it is redundant because of the similar content between object regions. The object representation requires a method enabling to choose the most relevant and representative object regions. We call this method, the representative object region detection method. In the previous work [3], we have proposed two representative object region detection methods based on Kmean and agglomerative clustering. In this paper, we use results of these methods to represent moving objects.

Definition 2: Moving object representation: A moving object O in surveillance video sequence is represented as $O = \{B_j, w_j\},\ j = 1...N$

Where B_j and w_j are the object region determined by the representative object region detection method and its weight. The greater the weight is, the more important the role of object region is. With this representation, we can take into account both temporal and appearance aspect of objects because all possible appearance features such as color, texture can be extracted in these object regions.

Definition 3: Moving object matching is an operator that allows to compute the similarity between two given objects.

In [3], we have proposed a moving object matching method based on EMD (Earth Mover's Distance) and covariance matrix. In this work, we use this method for calculating object similarity.

$\Omega = \{O_i^t\}, i = 1..N_{obj}$: set of N_{obj} objects in the database.

$R^t = \{O_i^t\}, i = 1..N_{obj}$: retrieval results at t^{th} iteration, $R^0 \in \Omega$: retrieval results at the 0^{th} iteration. These results are retrieval results without relevance feedback. The results in R are sorted in descending order of their similarity with query object. In a retrieval system, in general, the system shows only the first M results in R^t: we use $R^t(M)$ to indicate the first M results in R^t.

$O_i^{P,t}$: object judged by user as positive,

Relevance feedback method performance is evaluated by two factors:

- The number of relevant objects in $R^t(M)$ must greater than that in $R^{t-1}(M)$
- The number of iteration must be as small as possible

We present here two relevance feedback methods: a method based on object representation and a method based on SVM. The main idea of the first method is that we can create a new query object from a set of positive object regions determined by user. This query object can be a good representation of user wish. The main idea of the second method is we modify the similarity value of objects in the database and query. In other term, the first method bases on query modification and the second one bases on similarity measure modification. The input and output of these methods are the same. The input can be an object or an image by example and the output is the iteration where user decides to finish the interaction (t_{finish}) and the corresponding results R^t_{finish}. Pseudo code of these methods is shown in Table 1.

Table 1. Pseudo code of the proposed methods

Method based on object representation	Method based on SVM

Begin
Step 1: do object matching to obtain $R^0(M)$
while do
Step 2: ask user to judge results $R_i^{P,t-1} = \{O_i^{P,t-1}\}$
Step 3: modify object of query
Step 4: do object matching to obtain $R^t(M)$
until user decides to finish
End
	Begin
	Step 1: do object matching to obtain $R^0(M)$
	while do
	Step 2: ask user to judge results $R_i^{P,t-1} = \{O_i^{P,t-1}\}$
	Step 3: train SVM
	Step 4: recalculate object similarity to obtain $R^t(M)$
	until user decides to finish
	End

The main difference between two methods is in steps 3 and 4. With the method based on object representation, based on user judgment we create an intermediate query object $O_I = \{B_i, w_i\}, i = 1...M_p$ with $w_i = \frac{1}{M_P}$ where B_i is the positive object regions and M_p is the number of positive results. In step 4, the system computes similarity between this intermediate object and the objects in the database. In the method based on SVM, in the step 3, from M_p positive object regions, we create a training vector for one-class SVM by extracting covariance matrix feature [3] in each object region. Since after traing one-class SVM, the probability of a new sample is determined by $f(x) = \text{sgn}((w.\phi(x)) - \rho)$, the similarity of object O_j with query is then set to $f(O_j) = w.\phi(O_j) - \rho$ in step 4.

4 Experimental Results and Discussion

We can evaluate performance of relevance feedback approaches by evaluation measures. However, it is difficult compare the performance of different approaches even if they work on the same database and the same set of queries. Since retrieval methods based on relevance feedback interact with the user thus their performance at each iteration depends heavily on two factors: results and the number of results judged by the user at this iteration. In some cases, we can fix the number of judged results at each iteration. However the first M results of different approaches consist of different elements. From this analysis, we analyze in this section experimental results of our proposed. In order to validate the retrieval results, we adopt the evaluation measure: the Average Normalized Rank (ANR). It is defined as follows:

$$ANR = \frac{1}{NN_{rel}}\left(\sum_{i=1}^{N_{rel}} R_i - \frac{N_{rel}(N_{rel}+1)}{2}\right) \quad (1)$$

where N_{rel} is the number of relevant results for a particular query, N is the size of the tested set, and R_i is the rank of the i^{th} relevant results. The ANR measure is in the range 0 (good retrieval) to 1 (bad retrieval), with 0.5 corresponding to a random retrieval.

For evaluating the proposed approach performance, two surveillance video sequences named CARE_1 and CARE_2 coming from the CARETAKER (Content Analysis and REtrieval Technologies to Apply Extraction to massive Recording) project are employed. These videos depict human activity in a metro station in Italy. They are captured from the same scene by different cameras. After applying VSIP platform of PULSAR team for analyzing video [2], 770 and 810 people are detected and tracked in CARE_1 and CARE_2 respectively. In this experiment, we take objects of CARE_1 as query objects of CARE_2 as objects in the database.

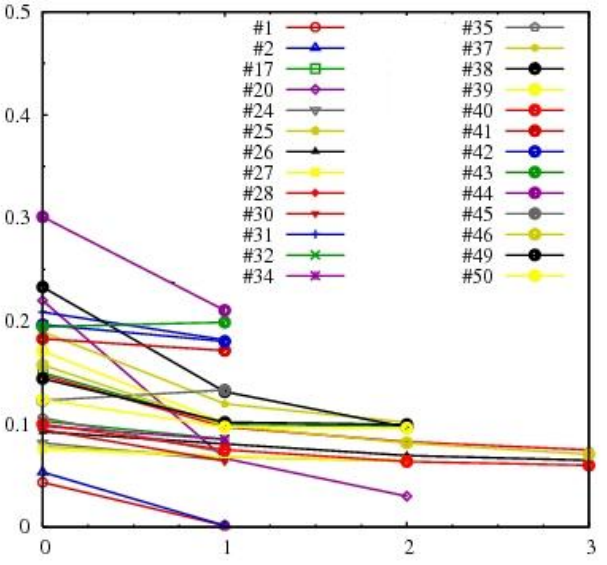


Fig. 2. ANR obtained with the method based on object representation for 50 query objects coming from video CARE_1. The vertical axis is ANR while the horizontal axis indicates the iteration.

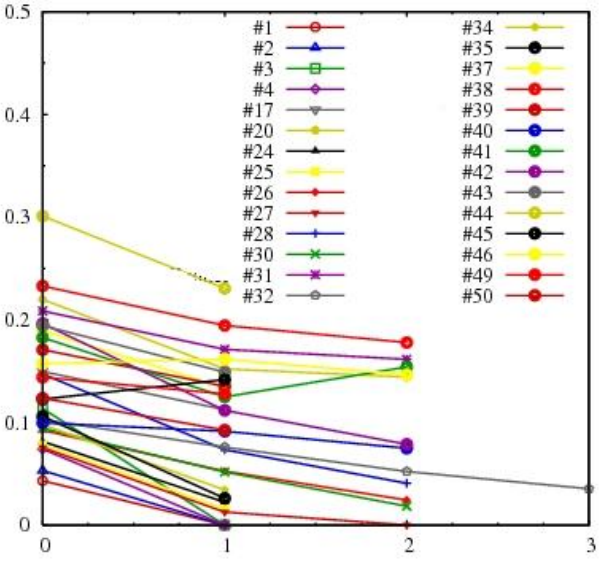


Fig. 3. ANR obtained with the method based on SVM for 50 query objects coming from video CARE_1. The vertical axis is ANR while the horizontal axis indicates the iteration.

We propose a method for performing and evaluating relevance feedback automatically. Firstly, we fix the number of returned results (in this paper M= 100) and the maximum number iterations (in this paper, this value is set to 5). For each iteration, the system chooses automatically relevant objects as positive results based on label information. A relevance feedback retrieval session will finish if it meets one of two criteria: (1) the number of relevant results in the working iteration is smaller than that in the previous iteration (the retrieval quality does not improve); (2) the number of iterations is bigger than maximum number (the system asks a lot of user feedback).

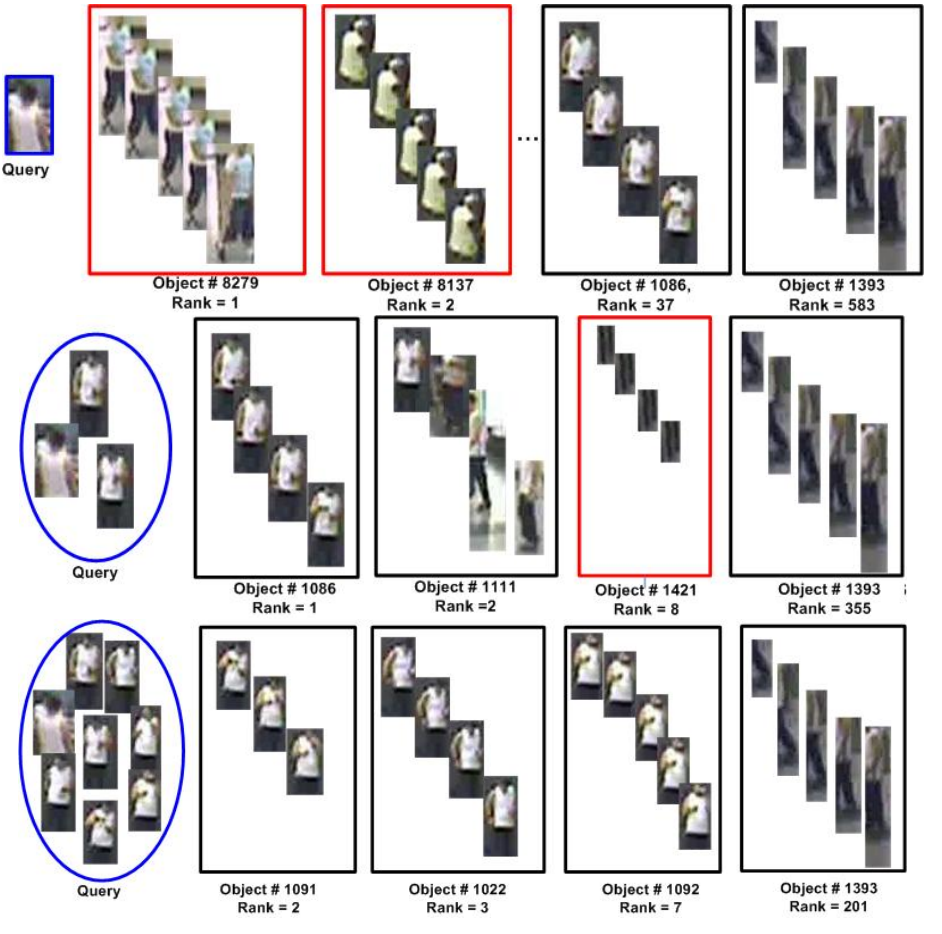


Fig. 4. Retrieval results without relevance feedback (the first row, after one iteration (the second row) and after two iterations (the third row)

Figures 2 and 3 show ANR at each iteration for 50 queries obtained by using the method based on object representation and method based on SVM respectively. As we can see, in general the obtained ANR decreases after each iteration, this means

that retrieval quality is improved. Our experimental results of two methods on the same dataset show that the method based on SVM outperforms the method based on object representation because the ANR of the former method decreases faster than that of the later one. However, the number of relevant results and therefore that of judged results at each iteration are sometimes small. This can cause sample lacking problem for training SVM. Figure 4 illustrates retrieval results without relevance feedback and after one and two iterations. The rank of relevant results decreases after each iteration.

5 Conclusions

We proposed in this paper two relevance feedback methods at object level for surveillance video. As analyzed in section IV, results with relevance feedback are better than those without relevance feedback. Relevance feedback for the video surveillance is equivalent to image indexing and retrieval at region level because an image (a set of regions) is equivalent to a moving object (a set of object regions detected in several frames). However, it is worth to note that for image indexing and retrieval, an image is judged as positive, it can contain only a or few positive regions while for video surveillance, a moving object is positive, all object regions are positive. We can use the most important object region or all the object regions. However, the relevance feedback method presented in this paper is short term relevance feedback. The system forgets the knowledge learned in each search session. This work can be extended by taking into account long-term relevance feedback.

References

Le, T.-L., Boucher, A., Thonnat, M., Brémond, F.: Surveillance video retrieval: what we have already done? In: Third International Conference on Communications and Electronics, ICCE 2010 (2010)

Le, T.-L., Boucher, A., Thonnat, M., Brémond, F.: Surveillance video indexing and retrieval using objet features and semantic events. International Journal of Pattern Recognition and Artificial Intelligence, Special issue on Visual Analysis and Understanding for Surveillance Applications 23(7), 1–37

Le, T.-L., Thonnat, M., Boucher, A., Brémond, F.: Appearance based retrieval for tracked objects in surveillance videos. In: Proceeding of the ACM International Conference on Image and Video Retrieval, Santorini, Fira, Greece, July 08 - 10, pp. 1–8 (2009)

Rui, Y., Huang, T.S.: Relevance feedback techniques in image retrieval. Principles of Visual Information Retrieval, 219–258 (2001)

Meessen, J., Desurmont, X., Delaigle, J.F., De Vleeschouwer, C., Macq, B.: Progressive Learning for Interactive Surveillance Scenes Retrieval. In: IEEE International Workshop on Visual Surveillance (VS 2007), pp. 1–8 (2007)

Chen, X., Zhang, C.: An Interactive Semantic Video Mining and Retrieval Platform—Application in Transportation Surveillance Video for Incident Detection. In: Perner, P. (ed.) ICDM 2006. X. Chen and C. Zhang, vol. 4065, pp. 129–138. Springer, Heidelberg (2006)

Identifying the Common Elements of a Commercial Website by Semantic Ontology

Tirayu Songvetkasem

Hawaii Pacific University,
Honolulu, Hawaii, United States
tsongvet@my.hpu.edu

Abstract. The main objective of this research was to examine the structure and content of a commercial website as well as identify the common elements between these websites by using the ontology concept. A semantic approach was employed, in order to describe the structure and classify the content in a commercial website. The objectives of each website were examined by professionals from advertising agencies who have more than one decade of experience in building a commercial website. The ontology of each website was developed based on website's objectives. Common elements were found in the results of the experiment. However, there was a wide range of properties in a same type of element.

Keywords: Ontology, semantic web, website ontology, website element, commercial website.

1 Background and the Purpose of the Study

A commercial website is a majority tool that companies use for providing the information of their products and services on the internet [1][3][13]. Despite the fact that the trend of using the online channel as a marketing tool has dramatically increased in the recent years, the failure rate of using this channel is still high [4]. A wide range of modern user interface designs, various development technologies and numerous marketing strategies have been applied to develop commercial websites, aiming to increase the effectiveness and ability to draw the attention from its target groups [5][9]. However, many researchers discovered that the main objectives of most commercial websites are alike, which are: providing the information of companies' products, services, or marketing campaigns [1][3][6][13][14].

Many researchers conducted the experiments to examine a commercial website [7][8][13]. On the other hand, no research uses an ontology to classify the elements of a commercial website was found. The purpose of this study was to examine the structure and content of a commercial website as well as identify the common elements among these websites by using a semantic and ontology approach. The research hypotheses of this study were: (1) there are common elements in commercial websites and (2) a semantic and ontology approach can be used to improve the ability to classify the structure and content of a commercial website. If the hypotheses is true, then a semantic approach can improve the ability to explain the structure and content of a commercial website. Moreover, an ontology approach can improve the performance of other measuring tools such as web analytics.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 183–187, 2011.

2 Methodology

The sample data in this study were selected from commercial websites developed by multinational companies that were currently online at the time of the study. After that, the objectives of websites were examined by experts. Then, the ontology of each website were constructed based on the objectives from experts. In the analysis part, the ontology of each website and the common elements among websites were examined.

A. Websites in the Study

In this study, 384 commercial websites were selected. All of these websites were developed by enterprise and multinational companies. The researcher used a convenient approach to choose the websites from the list of companies in the Fortune 500. Some websites were selected from the favorite website awards (FWA) [11]. FWA is a widely recognized award program that contains high quality websites, which were developed by advertising agencies and leading brands around the world, including Ogilvy, Young and Rubicam, McCann, Mindshare, DDB, Adidas, Apple, Toyota, Honda, Mercedes Benz, Walt Disney, and Google. In this award, experts used several criteria to evaluate the quality of candidate websites. Key criteria that FWA used for evaluating these websites were creativity, originality, design, content, and personality. In the study, the four key criteria that the researcher used for selecting the website were:

- The website must be developed by a company in the Fortune 500 company list [2] or a company that has annual of revenue more than one billion dollars.
- The website must be developed by a multinational company only.
- The websites are online at the time of the study.
- The website must be in the English language only.

B. Experts

Five experts from leading advertising agencies were invited to examine the objectives and construct the ontology of each website in this study. These experts were working at Mindshare, Ogilvy, McCann, Chuo Senko, and Young and Rubicam. All of them have more than one decade of experience in developing a commercial website for many enterprise organizations such as Intel, Disney, Nestle, Hewlett Packard (HP), Epson, Toyota, Honda, Sony, Nokia, LG, and Samsung. Each of them usually manages more than 100 commercial websites per year.

C. Ontology

In this study, the researcher developed the ontology in the Web Ontology Language (OWL) standard [15]. The Protégé-OWL editor were used to construct the ontology of each website. The Protégé-OWL editor was developed by Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine [12]. Protégé is a free and open-source platform that is widely used for developing ontology. The researcher developed the ontology based on the guidelines of Taniar and Rahayu [10]. Two concerns were employed in the development process; first, the ontology must meet stakeholders' needs; second, the ontology must be flexible and be able to change in the future.

3 Data Analysis and Results

In the experiment, the experts categorized the commercial websites into two categories: marketing campaign website and product website. A marketing campaign website is a website that was developed for very specific objectives; most of these website were created for raising company's awareness. Some marketing campaign websites were not related to a product at all. A product website is a website that was developed for providing the detailed information of a product, product's promotion. Some of these websites were used as a sale channel. According to the information from experts, the success of a marketing campaign website is normally measured by the number of website's participants, while the success of a product website is usually measured by the product's sale numbers and website's pageview. It was found in the results that common elements existed in the high level of ontology in both types of commercial websites. Table 1 shows the list of elements that were found in the high level of a commercial website ontology.

Table 1. The List of Elements in the High Level of a Commercial Website Ontology

Element Name	% of existence in a marketing campaign website	% of existence in a product website	% of existence in a commercial website
Membership	50	26	31
Product	63	100	92
Website objective	75	89	85
Sale channel	0	53	31
Online store	0	37	27
Store location	0	32	23
Current promotion	0	16	12
Company information	0	26	23
Social Network Site's link	63	79	77
E-mail sharing tool	25	21	23
Multimedia gallery	75	63	69
Picture	50	53	54

In the experiment, product element existed in all product websites and still had high percentage of existence in a marketing campaign website. Website objective element, social network site's link element and movie element also had a high number of existence in a commercial website. However, in the lower level, there was a wide range of properties in one element. For instance, in the product element of each website ontology, one website contains movie gallery that describes the information of the movie, which can be classified as the information of a product, while another website contains sale information such as sale channels and store location. In sale channel element, one website contains world map element that has the information and coordinates of the stores of a product, while another website contains online store element that provides an online purchasing system for online users.

4 Conclusions and Discussion

In this study, the researcher examined the structure and content of commercial websites by using semantic and ontology approach; the ontology of each website was created and two types of commercial websites were found: a marketing campaign website and product website. A marketing campaign website was developed to serve very specific objectives, and a product website was developed for providing the promotion and information of a product. Moreover, some of these website were used as a sale channel. According to the experts, Two different sets of criteria were used for measuring the success of these two website types. A marketing campaign website was measured by the number of participants in a website, while a product website was measured by the product's sale numbers and a website's pageview.

Each website has a unique way for presenting its information. Although, various types of marketing strategies and interface designs were employed in the development of commercial websites, common elements were found in the high level of a commercial website ontology. However, same element in two commercial websites might contain two total different set of properties, which made it very difficult to create a standard ontology for a commercial website. For instance, the commercial websites of Saab, Mitsubishi Motors and Mercedes Benz provided the information of the same type of product, which was car, but these three websites contained different sets of properties in the product element of each website.

It was shown in the results that an ontology can be used for classifying the content in a commercial website as well as its structure. Moreover, ontology can provide the relation of website sections in a website's structure, while a traditional sitemap cannot. The researcher encourages the further research to continue identifying the rules or standard guidelines for creating an ontology of a commercial website as well as a method for improving the website measuring instrument by using the ontology concept.

Acknowledgment

Tirayu Songvetkasem has more than one decade of experience in a website development for many enterprise companies such as Walt Disney, Samsung, Toyota, Honda, LG, Nestle, Sony, and etc. He was in the management position for five years, providing the professional consulting services of online marketing strategies and commercial websites to various advertising agencies and many recognized

organizations such as Walt Disney, Honda, Samsung, Mindshare, Chuo Senko, Young and Rubicam, McCann and Cheil.

References

[1] Angelides, M.C.: Implementing the Internet for Business: A Global Marketing Opportunity. International Journal of Information Management 17, 405–419 (1997)

[2] Cable News Network. FORTUNE 500 annual ranking of America's largest Corporations, http://money.cnn.com/magazines/fortune/fortune500/2010/full_list/

[3] Dholakia, U.M., Rego, L.L.: What Makes Commercial Web Pages Popular? An Empirical Investigation of Web Page Effectiveness. European Journal of Marketing 32, 724–736 (1998)

[4] Forrester Research Institute, http://www.forrester.com/Product/CoverageArea/0,4674,5,00.html

[5] Gan, Q.: Mining the web to improve search engine performance. Polytechnic University (unpublished)

[6] Luther, W.M.: The Marketing Plan: How to Prepare and Implement. In: AMACOM, 3rd edn. (2001)

[7] Madnick, S., Siegel, M.: Seizing the opportunity: Exploiting Web aggregation. MIS Quarterly Executive 1, 35–46 (2002)

[8] Nel, D., Van Niekerk, R., Berthon, J., Davies, T.: Going with the flow: Web sites and customer involvement. Internet Research: Electronic Applications and Policy 9, 109–116 (1999)

[9] Roitman, H., Carmel, D., Yom-Tov, E.: Maintaining dynamic channel profiles on the web. In: Proceedings of the VLDB Endowment, vol. 1 (2008)

[10] Taniar, D., Rahayu, J.W.: Web semantics and ontology (2006)

[11] The FWA Favorite Website Awards, http://thefwa.com/

[12] The Protégé Ontology Editor, http://protege.stanford.edu/

[13] Thelwall, M.: Effective websites for small and medium-sized enterprises. Journal of Small Business and Enterprise Development, vol. 7, pp. 149–159 (2000)

[14] Saaksjarvi, M.: In Search of Business Success on the Web. The Dilemma of Defensive Functionality. In: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences (1999)

[15] Web Ontology Language, http://www.w3.org/2004/OWL/

Computational Techniques in Political Language Processing: AnaDiP-2011

Daniela Gîfu[1] and Dan Cristea[1,2]

[1] "Alexandru Ioan Cuza" University, Faculty of Computer Science,
16, General Berthelot St., 700483, Iaşi
{daniela.gifu,dcristea}@info.uaic.ro
[2] Institute for Theoretical Computer Science, Romanian Academy - Iaşi branch,
2, T. Codrescu St., 700481, Iaşi

Abstract. This paper presents a computational method, AnaDiP-2011, based on natural language processing (NLP) techniques for the interpretation of the political discourse. The application considers the 2009 presidential campaign in Romania. The concept behind this method is that the manner in which individuals speak and write betrays their sensibility. Our investigation is intended to give support to researchers, specialists in political sciences, political analysts and election's staff, being helpful mainly in their social exploration of the electoral campaigns in their intend to measure reactions with respect to the developments in the political scene.

Keywords: political discourse, natural language processing, elections, semantic analysis, presidential, journalist.

1 Introduction

The motivation for our study relies on the need for objectivity in the interpretation of the political discourse, situated at the intersection of three important symbolic spaces: the political space, the public space and the communicational space (Wolton, 1998), as well as on the need to measure to what extend a discourse can influence its direct receptor, the electorate, and in what ways. The political language, essential in building a discourse (Perlmutter, 1999), requires an interdisciplinary approach (Şoitu, 2001) in which linguistics should co-ordinate with the rhetorical science, the communication sciences, the political sciences and the sociology.

Among many attributes the political discourse has, we were interested in the lexicon and its interpretation in a range of semantic coordinates. The final objective of our research is to develop a computational framework able to offer to the researchers in political sciences, to political analysts, to mass-media, to the public at large (interested to consolidate their options before elections), and, why not, even to politicians themselves, the possibility to measure different parameters of a written political discourse. Based on these parameters they should be able to appreciate certain aspects characterising the author of the discourse, as shading lights on his/her personality, or on the way he/she perceives the society or only some levels of it, as

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 188–195, 2011.

well as on his/her persuasive arsenal, etc. This can be done provided we will be able to link the statistical values outputted by the computational tool onto facts about the author of the message or the reality he/she is depicting. We were aware that the interpretation of numerical findings the program outputs should be validated by human experts in order to become facts. Part of our research, as reported in (Gifu, 2010), was concentrated on this type of human validation.

Analysis of this type are not new. LIWC-2007 (www.liwc.net), for instance, does a similar type of investigation over texts. The software we developed, AnaDiP-2011 (Analysis of Discourse in Politics), while copying some features from LIWC-2007, includes many new functionalities. It offers the possibility to analyze efficiently large bodies of text and to characterize them quantitatively and qualitatively, the results having to be as close as possible to the analysis made by a human expert. The system offers a global perspective over the political discourse, as well as a punctual one.

The paper is structured as follows. Section 2 shortly describes the functionality of the software and the associated resources for the Romanian language. Then, section 3 discusses an example picked up from the 2009 presidential campaign in Romania, and sections 4 presents some conclusions.

2 The Software and Resources

Although, as mentioned, the functionality of AnaDiP-2011 is inspired by LIWC-2007, there are important differences between the two platforms. LIWC-2007 is basically counting words and incrementing counters of all their declared semantic classes, when they are discovered in the input text. In the lexicon, words can be given by their long form, as a complete string of characters, or abbreviated, in which case the sign '*' plays the role of the universal jolly-joker, replacing any character. For each text in the input, LIWC produces a set of tables, each displaying the occurrences of the word-like instances of the semantic classes defined in the lexicon, as sub-unitary values. For one semantic class, such a value is computed as the number of occurrences of the words corresponding to that class divided by the total number of words in the text. It remains in the hands of the user to interpret these figures. And there is no support for considering lexical expressions.

We will refer now to the way in which AnaDiP-2011 organises the lexicon and how it counts words. The software performs part-of-speech (POS) tagging and lemmatisation of words. This is why the lexicon can now be declared as a collection of lemmas having the POS categories: verb, noun, adjective and adverb. As seen, we leave out the pronouns, numerals, prepositions and conjunctions, considered to be semantically empty. An entry of the lexicon has the form: <lemma> <POS> <sem-list>, where <sem-list> is a list of semantic classes. This means that the same lemma can appear with more than one POS and, if needed, with different semantic interpretations.

Although the introduction of lemmatisation and POS tagging makes useless the 'any' operator, we have kept it in the definition of the lexicon entries. The user has the possibility to either define an entry as a <lemma><POS><sem-list> triple, as explained above, or as <word-root>(*) - <sem-list>, with the significance that the root can be ended with a '*' sign, no POS is defined and any word matching the root

during analysis will increase the counters of all semantic classes belonging to the <sem-list>.

The second range of differences between the two platforms stays in the user interface. In AnaDiP-2011, the user has an easy to interact interface, offering a lot of services: opening of one or more files, displaying the file/s, modifying/editing and saving the text, functions of undo/redo, functions to edit the lexicon, visualising the mentioning of instances of certain semantic classes in the text, etc. Then, the menus offer a whole range of output visualisation functions, from tabular form to graphical representations and to printing services. Figure 1 shows a snapshot of the interface during a working session.

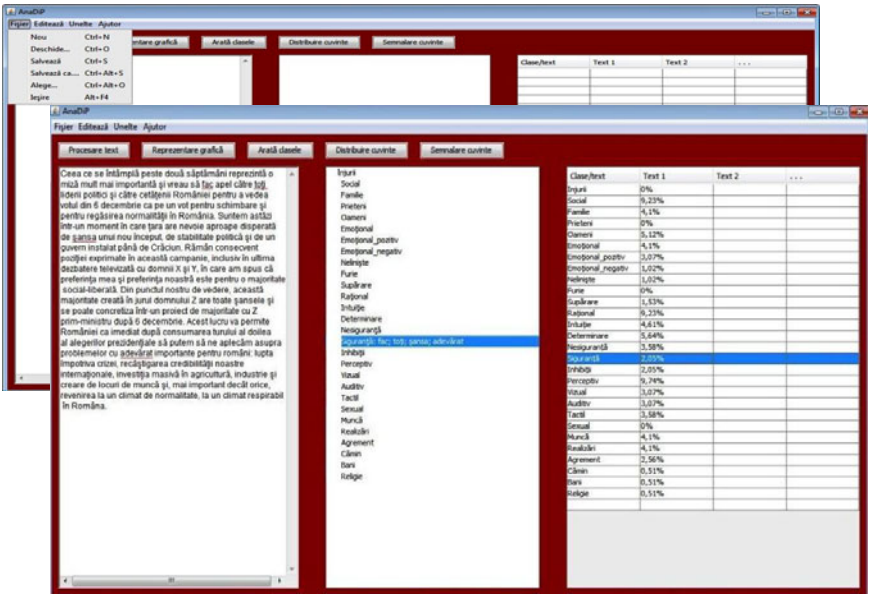


Fig. 1. The AnaDiP-2011 interface: in the left window appears the selected text, on the middle window – the classes, and in the right window – the words and their respective counted percentages. By selecting some word on the right or a class in the middle window, the corresponding occurrences in the text are highlighted in the left window.

Finally, another important development was the inclusion of a collection of formulas which can be used to make comparative studies between different subjects. In section 3 we will present an example.

The Romanian lexicon of Ana-DiP-2010 contains now approximately 6000 words and roots and 28 semantic classes. We have upgraded lexicon by importing from DEX-online (www.dexonline.ro), the greatest public online dictionary for Romanian. The semantic classes in AnaDiP-2011 are partially placed in a hierarchy. In the future we plan to align this hierarchy with WordNet (Fellbaum, 2001), for languages which support this type of linguistic resource.

A special section of the lexicon includes expressions. An expression is defined as a sequence: <root-list> => <sem-list>, in which <root-list> is a list of roots of words, therefore each optionally followed by the '*' sign. Because, in principle, a root can also be a numerical value and the semantic classes in <sem-list> are indicated by numbers, to separate the roots section from the semantic categories section we had to place a special sign (=>). Each time a sequence of words matching the <root-list> is recognised in the text, the counters associated with the semantic classes in the <sem-list> are increased.

As an example, for the root `plictico*`[1], the following classes are assigned: 6 = Affect, 8 = Negative_emotion and 11 = Sadness. Whenever the word or a variant of the word is detected in the input file, all three counters, corresponding to the classes mentioned, are incremented.

The 28 semantic classes included now in AnaDiP-2011 (Table 1) have been selected to fit optimally with the necessities to interpret the political discourse of the presidential campaign in Romania, in 2009. We were mainly interested to determine those political attitudes which were able to influence the voting decision of the electorate. However, the user can define at his/her will these classes and the associated lexicon.

Table 1. The 28 semantic classes currently included in the AnaDiP-2011 package

code	Class in Romanian (English)	Description and examples in RO (EN)
1	`Injurii` (`Swear`)	Words that harm the reputation of someone: *afurisit, cretin, jigodie* (*damned, idiot, cur*)
2	`Social` (`Social`)	Words about social appurtenance: *preşedinţie, eveniment, celebrare* (*presidency, event, celebration*)
3	`Familie` (`Familiy`)	*clan, părinte, fiu* (*clan, parent, son*)
4	`Prieteni` (`Friends`)	*camarad, coleg, companion* (*comrade, colleague, companion*)
5	`Oameni` (`Humans`)	*adult, personalitate, ţăran* (*adult, personality, peasant*)
6	`Emotional` (`Affect`)	Words with emotional effect: *abătut, curajos, impasibil* (*depressed, courageous, impassive*)
7	`Emotional_pozitiv` (`Positive_emotion`)	Words with positive effect: *ademenitor, plăcut, popular* (*alluring, nice, popular*)
8	`Emotional_negativ` (`Negative_emotion`)	Words with negative effect: *agresiv, cinic, trândav* (*aggressive, cynic, slothful*)
9	`Nelini te` (`Anxiety`)	*alarmant, tensionat, bănuitor* (*alarmed, jumpy, suspicious*)
10	`Furie` (`Anger`)	*brutal, ofensator, irascibil* (*brutal, offensive, irascible*)
11	`Sup rare` (`Sadness`)	*îngrijorător, revoltător, trist* (*alarming, revolting, sad*)
12	`Raţional` (`Rational`)	*inventiv, discernământ, precaut* (*inventive, discernment, precautious*)
13	`Intuiţie` (`Intuition`)	*ager, explicativ, realist* (*astute, explicative, realist*)
14	`Determinare` (`Cause`)	*deductiv, insinuant, manipulator* (*deductive, insinuate, manipulator*)
15	`Nesiguranţă` (`Uncertainty`)	*vag, confuz, ambiguu* (*vague, confused, ambiguous*)

[1] From *plicticos*, *plicticoasă*, *plicticoşi*, *plicticoase* (*boring*).

Table 1. (*continued*)

16	Siguranţă (Certainty)	*exact, hotărât, categoric* (*precise, resolute, absolutely*)
17	Inhibiţie (Inhibition)	*absent, descurajat, şovăitor* (*absent, discouraged, giddy*)
18	Perceptiv (Percept)	Words related to perceptions: *auzi, observa, palpabil* (*hear, observes, tangible*)
19	Vizual (See)	*holba, lumină, orb* (*stare, light, blind*)
20	Auditiv (Hear)	*acustic, tăcut, surd* (*acoustic, silent, deaf*)
21	Tactil (Feel)	*ascuţit, fragil, estompat* (*sharp, fragile, blurry*)
22	Sexual (Sexual)	Words with a sexual tonality: *pipăi, dezmierdare, încins* (*caress, endearment, hot*)
23	Muncă (Work)	Work specific words: *Calificare, department, organizaţii* (*qualification, department, organization*)
24	Realizări (Achievements)	Words that reveal human achievements: *facultate, succes, competenţă* (*college, success, skills*)
25	Agrement (Leisure)	Words from the recreational domain: *dans, cazinou, camping* (*dance, casino, camping*)
26	Cămin (Home)	Words related to house: *cameră, decor, mansardă* (*room, decoration, attic*)
27	Bani (Money)	Words related to economy and finance: *avere, cont, licitaţie* (*wealth, account, auction*)
28	Religie (Religion)	Words with a religious tonality: *Biblie, divinitate, creştin* (*Bible, divine, Christian*)

Keeping in mind the remarkably sophisticated and time consuming process in which the Romanian version of the dictionary LIWC-2007 was acquired, we knew that some decisions have to be taken in order to optimize its content while also diminishing the influence of English, evident if a simple EN-onto-RO translation process would have been applied. The development of the lexicon was done in several phases:

- first the LIWC-2007 English terms belonging to the 28 classes previously mentioned, considered meaningful for this type of analysis, were translated, retaining only the Romanian words which had senses in accordance with the corresponding classes;
- then, words / roots in each class were sorted alphabetically;
- then we have reconsidered each class in part, eliminating words that could introduce ambiguities and including synonyms. We have done this activity with a class of master students in Computational Linguistics[2]. Then, the work done by students has been once again validated by both authors. By working within one semantic class at a time, we were able to easily recognise classification mistakes and correct them. Also, the alphabetical ordering offers the possibility to operate certain roots optimizations, by exploiting the use of the jolly-joker *;
- then a close look in a part of the documents. The collection was done by one of the authors. When a competent reader does this, normally do pop-up important words.

[2] In 2010, in their 1st year at the Faculty of Computer Science, the "Alexandru Ioan Cuza" University of Iaşi.

They were checked against the DEX-online lexicon by using '*' to include more possibilities and the resulting matches were included in the AnaDiP-2011 lexicon.

- then, the dictionaries assigned to all classes were merged, and the obtained list was sorted again alphabetically. Multiple occurrences were removed (by clashing together the lists of classes, and leaving only one instance of a class in the list corresponding to one root/word);
- then, when the root/word notation (including or not the use of '*') was seen to give rise to unwanted ambiguities, the <lemma><POS> notation was used instead;
- finally, starting from words, expressions were introduced;
- as mentioned already, in a future stage, we will use the words we have now in the lexicon as seeds in a trial to enrich it automatically, by making use of DEX-online. We study now strategies to exploit the synonymy relation and the definitions.

3 Analysis of the Political Discourse during the 2009 Romanian Presidential Elections

For the elaboration of preliminary conclusions over the presidential elections process, conducted in the period October – December 2009 in Romania, we collected, stored and processed electronically, in three different stages (one month before, the 1st tour and, finally, the 2nd tour of the election's campaign), political texts, i.e. editorials published by three national publications having similar[3] profiles, as well as political speeches (in both oral and written form), all belonging to four candidates for the function of President[4].

The monitored written media corpus was processed directly with AnaDiP-2011. The speech records were previously manually transposed onto text and then they followed the same processing as the written texts. In essence, the program receives the input from one or more text files, and counts occurrences of words belonging to its defined classes. The user can notice directly the mentioned semantic classes (and the corresponding frequencies), as the words belonging to a selected class appear underlined in the left screen. The user can choose a type of graphical representation ("function", "pie" or "columns"), which give intuitive visual perceptions on which the interpretation of discourse data can be performed more conveniently.

Apart from simply computing frequencies, the system can also perform comparative studies. The assessments made are comprehensive over the selected classes because they represent averages on collections of texts, not just a single text.

---

[3] The three newspapers have been *Evenimentul Zilei*, *Gandul* and *Ziua* (www.mediapres.ro), which are known to have a common profile: national dailies of general information, tabloids with a circulation of tens of thousands of copies per edition, each. The newspapers were monitored on their websites: *Evenimentul zilei* – www.evz.ro, *Gandul* – www.gandul.info, *Ziua* – www.ziua.ro.

[4] The four candidates are: Crin Antonescu (PNL), Traian Băsescu (PDL), Mircea Geoană (PSD+PC) and Sorin Oprescu (independent). The information sources considered have been: www.crinantonescu.ro, www.pnl.ro, www.basescu.ro, www.pdl.org.ro, www.presidency.ro, www.mirceageoana.ro, www.psd.ro and www.sorinoprescu.ro.

AnaDiP-2011 provides a library of comparative functions, with 2 to 4 different input streams of data. One stream can be either a presidential candidate, or a newspaper, or only one discursive approach on a certain topic delivered at a certain moment in time by the traced author. To exemplify, one type of graphics considered for the interpretation was the one-to-one difference, as given by Formula (1), included in the AnaDiP-2011 Mathematical Functions Library:

$$Diff_{x,y}^{1-1} = average(x) - average(y) \quad (1)$$

where x and y are two streams; *average(x)* and *average(y)* are the average frequencies of x and y over the whole stream, and the difference is computed for each selected class. Since a difference can lead to both positive and negative values, these particular graphs should read as follows: values above the horizontal axis are those prevailing at the candidate / daily x versus the candidate / daily y, and those below the horizontal axis show the reverse prominence. A zero value indicates equality.

To exemplify, we present below a chart with two streams of data, representing the texts of the final TV debate between the two remaining presidential candidates in the election race, the second tour of voting.

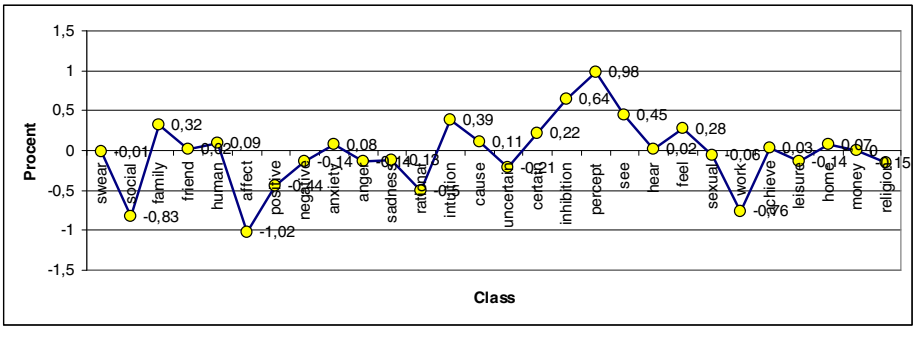


Fig. 2. The average differences in the frequencies for each class after processing the final TV debate, December 4, 2009, between the candidates: Traian Băsescu versus Mircea Geoană.

Our experience shows that an absolute difference value below the threshold of 0.5% should be considered as irrelevant and, therefore, ignored in the interpretation. In the graphical representation of Fig. 2, we compare Traian Băsescu (TB – a Democrat Liberal Party candidate) against Mircea Geoană (MG – the Social Democrat Party), by applying Formula (1): as such, if a value of a class is positive (above the horisontal axis) it shows the prevalence of TB over MG in that class, and if it is negative (bellow the horisontal axis), it shows the prevalence of MG over TB in that class. This should be interpreted as follows: TB's discursive intervention are much more `perceptual` than MG's, therefore touching directly onto human common sense, being also rather careful (`Inhibition`), while his opponent, MG, has touched more onto social issues (`Social`), quite normal, seen his affiliation with Social Democrats, had an emotional attitude (`Affect`), and with an obvious interest towards the working aspects (`Work`).

AnaDiP-2011 was compared on a sample corpus against LIWC-2007. We found some differences in the semantic classes for which updates to the lexicon have been performed. Many of the conclusions found by the program have been confirmed by political commenters. Moreover, the program helped also to outline distinctive features which brought a new, and sometimes even unexpected, vision upon the discursive characteristics of the presidential candidates, of the columnists and, last but not least, of the Romanian voters, at the end of 2009.

4 Conclusions

Currently, the NLP domain has reached a scientific and technological maturity that makes it useful in the activities carried out by researchers in socio-humanity fields, and initiatives like CLARIN (www.clarin.eu) show that a consistent bridge between these two communities can be built.

We believe that AnaDiP-2011 has a range of features that make it attractive as a tool to assist political campaigns. It can also be rapidly adapted to new domains and to new languages, while its interface is user-friendly and offers a good range of useful functionalities.

In the future we intend to include a word sense disambiguation module in order to determine the correct senses, in context, of those words which are ambiguous between different semantic classes belonging to the lexicon, or between classes in the lexicon and outside the lexicon (in which case they would not have to be counted).

Acknowledgments. The AnaDiP-2011 software has been developed by Iulia Scutariu from the Faculty of Computer Science, "Alexandru Ioan Cuza" University of Iaşi. The master students in the 1st year of Computational Linguistics at the Faculty of Computer Science of the "Alexandru Ioan Cuza" University of Iaşi have participated in the initial phases of the acquisition of the lexicon. In order to perform this research the first author received financial support from the POSDRU/89/1.5/S/63663 grant.

References

1. Fellbaum, C. (ed.): WordNet, An Electronic Lexical Database. MIT Press, Cambridge (2001)
2. Gîfu, D.: The Discourse of the Written Press and the Violence of Symbols, PhD thesis, Faculty of Philosophy and Political Studies, "Alexandru Ioan Cuza" University of Iaşi (2010) (in Romanian)
3. Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linquistic Inquiry and Word Count – LIWC 2001. Erlbaum Publishers, Mahwah (2001)
4. Perlmutter, D.D.: The Manship School guide to political communication. Louisiana State University Press, Baton Rouge (1999)
5. Şoitu, L.: Philosophy of Communication, Iaşi, Institutul European (2001)
6. Wolton, D.: Les contradictions de la communication politique. CNRS, Paris (1995)

An Efficient Approach for Video Action Classification Based on 3D Zernike Moments

I. Lassoued[1], E. Zagrouba[1], and Y. Chahir[2]

[1] SIIVA Unit of Research, Institut Supérieur d'Informatique
Ariana, Tunis
[2] Laboratoire GREYC, Université de Caen
14050 Caen Cedex, France
lassoued.imen@yahoo.fr, ezzeddine.zagrouba@fsm.rnu.tn,
youssef.chahir@info.unicaen.fr

Abstract. Action recognition in video and still image is one of the most challenging research topics in pattern recognition and computer vision. This paper proposes a new method for video action classification based on 3D Zernike moments. These last ones aim to capturing both structural and temporal information of a time varying sequence. The originality of this approach consists to represent actions in video sequences by a three-dimension shape obtained from different silhouettes in the space-time volume. In fact, the given video is segmented in space-time volume. Then, silhouettes are extracted from obtained images of the video sequences volumes and 3D Zernike moments are computed for video, based on silhouettes volumes. Finally, least square version of SVM (LSSVM) classifier with extracted features is used to classify actions in videos. To evaluate the proposed approach, it was applied on a benchmark human action dataset. The experimentations and evaluations show efficient results in terms of action characterizations and classification. Further more, it presents several advantages such as simplicity and respect of silhouette movement progress in the video guaranteed by 3D Zernike moment.

Keywords: Actions classification, Zernike moments, LS-SVMs, 3D silhouette shape.

1 Introduction

Audiovisual contents volumes don't stop growing. In fact, the problem is how to navigate and to look for exactly these contents within large collections. So, video indexation and retrieval by automatic content analysis is now one of the major goals in information systems. In this context, it is very important to extract automatically high-level information which can describe the semantic content of the given video. Indeed, many works were interested by the events or actions in video for several applications such as: sports analysis, visual surveillance or human computer interaction. Action classification consists to classify videos based on action of the object detected in a given video. In this paper, we present a new approach of video action classification based on 3D Zernike moments. This last ones are computed using

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 196–205, 2011.

silhouettes volumes. The remainder of this paper is organized as follows: the next section presents an overview of existing methods of action recognition and classification. In the third section, the proposed approach based on 3D Zernike moment and Least Square Support Vector Machine is detailed. The experimental results and evaluations are shown in section4. Finally, conclusion and perspectives are drawn in section5.

2 Related Works

The existing approaches in video action recognition can be classified in three main categories based on the used classification descriptor.

2.1 Optical Flow, Gradient or Intensity Based Methods

There are many existing works in action recognition which are based on global features such as optical flow, gradient histogram and intensity. Zelnik-Manor and Irani [1] used marginal histograms of spatiotemporal gradients at multiple temporal scales to cluster video events. Wu [2] develop an algorithm to automatically extract figure-centric stabilized images from the tracking system. They also propose to use Decomposed Image Gradients (DIG), which can be computed by decomposing the image gradients into four channels, to classify the person's actions. Efros et al. [3] proposed a descriptor based on blurred optical ?ow measurements, and applied it to recognize actions on ballet, tennis and football datasets. Dollar et al. [4] proposed to characterize behaviors through spatiotemporal feature points, in which a behavior was described in terms of the types and locations of feature points present. For this class of methods, the recognition results depend greatly on the recording conditions.

2.2 Feature Tracking Based Methods

Many activity recognition methods are based on feature tracking in either 2-D or 3-D space [5,6]. Rao and Shah [7] proposed an approach based on the trajectory of a tracked hand to differentiate between actions. Song et al. [21] used spatial arrangements of the tracked points to distinguish between walking and biking.Song and al [8] proposed to represent an action by 40 curves derived from the tracking results of five body parts using a cardboard people model. 3-D information is also used to establish motion descriptors based on positions, angles and velocities of body parts [9], [10]. Ali and Aggarwal [10] used the angles of inclination of the torso, the lower and upper parts of legs as features to recognize activity. Feature tracking is complex due to the large variability in the shape and articulation of the human body. In particular, perfect limb tracking is not yet well solved.

2.3 Silhouette-Based Methods

Many researches in action classification are now oriented to silhouette-based methods [11]-[12]. In fact, human actions can be characterized as motion of a sequence of human silhouettes over time. Gorelick et al[13] regard human actions as a three-dimension shape induced by the silhouettes in the space-time volume and utilizes

properties of the solution to the Poisson equation to extract space-time features such as local space-time saliency, action dynamics, shape structure and orientation. In [14] an action is viewed as a temporal sequence of local shape-deformations of centroid-centered object silhouettes, i.e., the shape of the centroid-centered object silhouette tunnel. Each action is represented by the empirical covariance matrix of a set of 13-dimensional normalized geometric feature vectors that capture the shape of the silhouette tunnel. Kellokumpu et al. [15] proposed a human activity recognition method from sequences of postures. A SVM was used for posture classification and then the discrete HMMs were used for activity recognition. Sminchisescu et al [16] recognized human motions based on discriminative conditional random field (CRF) and maximum entropy Markov models (MEMM), using image descriptors combining shape context and pairwise edge features extracted on the silhouette. For human silhouette extraction from videos, it is easy for current vision techniques, especially in the imaging setting with fixed cameras. So, the method that we present here directly relies on moving silhouettes.

3 Proposed Approach

The main goal of the proposed approach is to detect specific actions in videos. The contribution consists to use efficient 3D Zernike moments features to classify actions. In our approach, the human action in a video sequences is represented by a space-time shape in the space-time volume. These shapes are induced by a concatenation of 2D silhouettes in the space-time volume and contain both the spatial information about the pose of the human figure at any time as well as the dynamic information. The general architecture of the proposed approach is decomposed to several stages. In fact, given an input video, it is segmented to obtain images volumes. These last ones are used to construct image foreground and to obtain 3D silhouettes. Then, 3D Zernike moments are calculated for each obtained silhouette series. Finally, the LS-SVM classification method is applied to obtain the actions classes found in the input video. Figure 1 presents the different steps of the proposed approach.

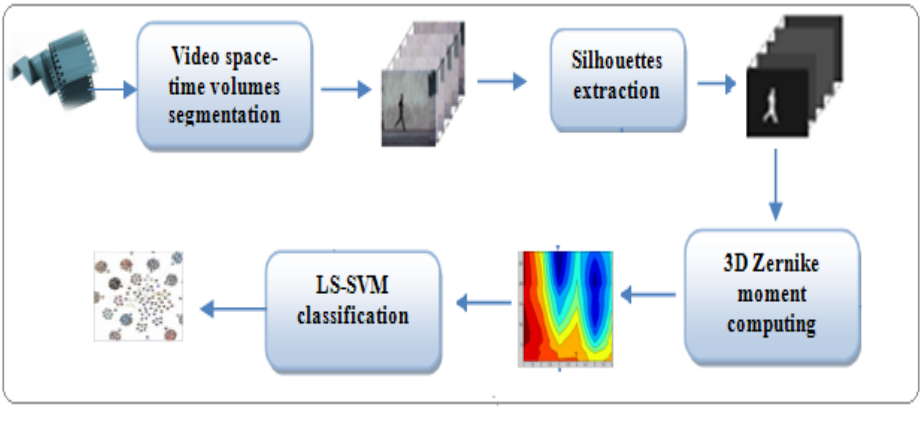


Fig. 1. General Architecture of the proposed approach

3.1 Segmentation and 3D Silhouettes Extraction

Given that video contains body separable actions and that the camera is fixed, the background can be rebuilt easily. In the proposed work, the median value of all pixels in the temporal direction is simply used to obtain the background. The shape of the body is computed by subtracting the background from each frame as shown in Figure 2 where the first row presents video into space-time volumes, the second row presents a static background and the last row explains a silhouette volumes obtained by static background subtracting from image. The last row in Figure 2 present volumes of extracted silhouettes, this obtained silhouettes are stored at the same resolution as original image.

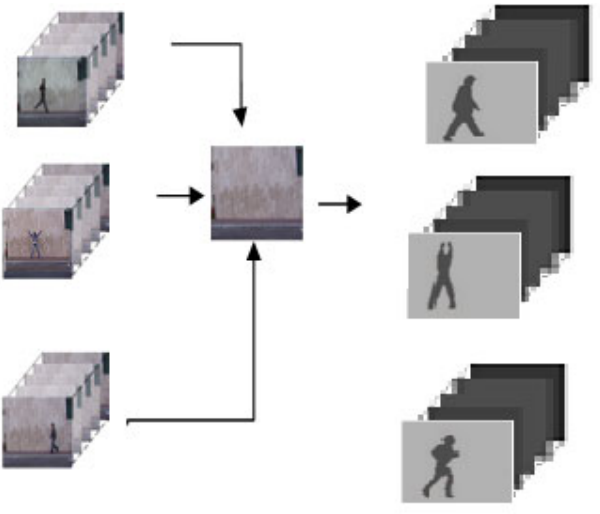


Fig. 2. Background subtraction

3.2 3D Zernike Moments Computing

The main contribution of the proposed approach is to use 3D Zernike moments also known as Zernike velocity moments for video action classification. Indeed, Zernike moments are the best among multiple invariant moments in terms of overall performance. Zernike moments are a class of orthogonal moments and have been considered as effective in terms of image and shape representation. In more, Zernike moments are rotation invariant and can be easily constructed to an arbitrary order. Although higher order moments carry more fine details of an image, they are also more susceptible to noise. Shutter and Nixon in [17] proved that Zernike moments perform well when applied to analyzing walk sequences resulting in a good recognition rate and a compact description. In the proposed approach, 3D Zernike moments are chosen to characterize the actions in video sequences. Therefore, Moments are calculated for each silhouettes series and experimented with different orders of 3D Zernike moments to determine the optimal order which can resolve the

proposed problem. To compute Zernike moments for the obtained silhouette volumes, the are two axis. The first one treat time as the zaxis when applied to images into three-dimensional XYT (x,y and time) block. However, this method confounds the separation of the time and space information, as they are embedded in the data and not specific to the descriptor. Time must be treated separately to space because they are fundamentally different. For our work, we choose to apply the second axis which resolves this problem. In fact, it reformulate the moment descriptor to incorporate time, enabling the separation and/or combination of the time and spatial descriptions. If Zernike velocity moments describe just spatial information (no motion) of a moving rigid shape, then the correlation between images is exploited, and is advantageous, refining the description of the rigid shape as the sequence increases in length. The final Zernike velocity moments of this sequence can be considered as refined (or averaged) Zernike moments of a single image, the descriptions of which are orthogonal. In our work, the shape is moving and deforming (non-rigid), such as a person walking. Then the spatial correlation between consecutive image descriptions is reduced. The Zernike velocity moments are a weighted sum of the Zernike moments over multiple consecutive images. The weighting (velocity) is real-valued and scalar, and the spatial description of each consecutive image in the sequence are orthogonal. The final descriptors of the moving and deforming shape are temporally correlated due to the use of the image sequence. The Zernike velocity moments are expressed as:

$$A_{mn} = \frac{m+1}{\pi} \sum_{i=2}^{\text{images}} \sum_{x} \sum_{y} U(i,\mu,\gamma) S(m,n) P_{ixy} \tag{1}$$

They are bounded by x2 + y2 $\leq$1, while the shape's structure contributes through the orthogonal polynomials:

$$S(m,n) = [V_{mn}(r,\theta)]^* \tag{2}$$

where * denotes the complex conjugate. The Zernike polynomials [3] Vmn(x; y), expressed in polar coordinates are:

$$V_{mn}(r,\theta) = R_{mn}(r) \exp(jn\theta) \tag{3}$$

where (r; θ) are defined over the unit disc and $R_{mn}(r)$ is the orthogonal radial polynomial, defined as:

$$Rmn(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s F(m,n,s,r) \tag{4}$$

Where $$F(m,n,s,r) = \frac{(m-s)!\, r^{m-2s}}{s!\left(\frac{m+|n|}{2}-s\right)!\left(\frac{m-|n|}{2}-s\right)!} \tag{5}$$

Figure 3 presents the series of silhouette in the first line and their associate 3D Zernike moments in the second line.

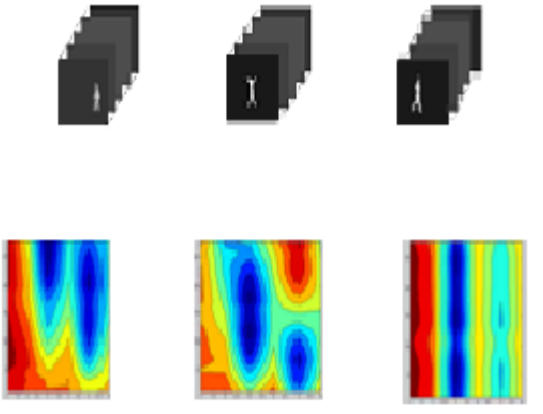


Fig. 3. Series of silhouette and her associate 3D Zernike moments

3.3 LS-SVM Classification

To classify actions in videos, many works are proposed using several attributes such as neural networks [18], GMM [19], etc. The most used classifier is SVM which allows obtaining efficient results to classify actions. It is a powerful classifier used successfully in many pattern recognition problems. For this reason, the Least squares support vector machines LSSVM method is used in the proposed work. It is a variant of standard SVM which simplify the SVM formulation without losing any of its advantages. This method is proposed by Suykens and Vandewalle [20] where the training algorithm solves a convex problem like SVM. It has been shown by a meticulous empirical study that the generalization performance of LS-SVM is comparable to that of SVM [21]. In addition, the training algorithm of LS-SVM is very simplified since a linear problem is resolved instead of a quadratic programming (QP) problem in the SVM case. Given a training set (x; y), i=1,2,,l with input data xi $\in$ R^n and output label data $y_i \in R^n$, and the classifier takes the following form:

$$y(x) = sign\ [\omega^T \varphi(x) + b] \tag{6}$$

Where $\varphi(.)$: Rn $\rightarrow$ Rn is the mapping to the high dimensional and potentially infinite dimensional feature space. In the primal weight space, the optimization problem becomes:

$$\min_{\omega,b,\xi} j(\omega,\xi) = \frac{1}{2}\omega^t\omega + \gamma\frac{1}{2}\sum_{i=1}^{l}\xi_i^2 \tag{7}$$

$$y_i[\omega^t\varphi(x_i)+b] = 1-\xi_i, i = 1,2,...,l \tag{8}$$

Where $\xi_i > 0$ denotes a real constant used to control the punishment degree for misclassification. Because ω becomes infinite dimensional, this primal problem cannot directly be solved. Therefore, let us proceed by constructing the following Lagrangian:

$$L(\omega, b, \xi, \alpha) = J(\omega, \xi) - \sum_{i=1}^{l} \alpha_i (y_i [\omega^t \kappa(x_i) + b] - 1 + \xi_i) \quad (9)$$

Where the values are Lagrange multipliers, which can be positive or negative due to the equality constraints [21].

4 Experimental Results

To evaluate the proposed approach, it was applied on publicly available benchmark dataset proposed by weizmann [22]. This last one is usually used to test action classification models, that originally contained 81 low resolution videos (180* 144). This dataset consists of 93 videos, where 9 people perform 9 different actions: running, bending, waving with one hand, jumping in place, jumping jack, jumping, walking, skip, and waving with two hands. Each video clip contains one subject performing a single action. Illustrative examples for each of these actions are shown in Figure 4. Each video in the dataset is segmented in the space-time volume. Then, the associated silhouettes are extracted. 3D Zernike moments are calculated for each silhouette volumes to describe video. These moments are calculated for square images. For this reason, we have resized the frames series a common size of 140*140 pixels. In the experiments, we select 1/3 of the silhouettes series from each action category to form the training set (40 silhouette sequences) which are used for training LS-SVM, and the rest of series are used for testing LS-SVM. The outputs of LS-SVM process present the corresponding classes of actions.

Fig. 4. Example of actions in the weizmann dataset

To choice the optimal and representative order of Zernike moments which allows to resolve the proposed problem, different orders are tested. The experimentations show that order seven is the best. In fact, it permits to obtain the best values on the diagonal of the confusion matrix. In more, the order higher than seven didn't permit more improvement for classification. The following array presents the confusion matrix for the third order (a1-"walk," a2-"run," a3-"skip," a4-"jack," a5-"jump," a6-"jump in place," a7- "wave with one hand," a8- "wave with two hands," and a9-"bend").

	a1	a2	a3	a4	a5	a6	a7	a8	a9
a1	85	10	2	0	0	0	3	0	0
a2	7	80	0	0	0	0	8	5	0
a3	10	0	83	7	0	0	0	0	0
a4	0	3	0	75	4	3	4	5	6
a5	8	11	3	8	70	0	0	0	0
a6	13	12	3	0	7	65	0	0	0
a7	7	12	0	0	7	0	68	6	0
a8	11	0	9	0	0	0	9	71	0
a9	0	0	0	0	3	0	9	7	81

Array (a)

	a1	a2	a3	a4	a5	a6	a7	a8	a9
a1	100	0	0	0	0	0	0	0	0
a2	2	98	0	0	0	0	0	0	0
a3	0	3	97	0	0	0	0	0	0
a4	0	0	0	100	0	0	0	0	0
a5	10	0	0	0	83	7	0	0	0
a6	0	0	2	0	0	98	0	0	0
a7	0	0	0	0	2	0	96	0	2
a8	0	0	0	0	0	0	4	96	0
a9	0	0	0	0	0	0	1	0	99

Array (b)

	a1	a2	a3	a4	a5	a6	a7	a8	a9
a1	88	11	0	0	0	1	0	0	0
a2	4	93	3	0	0	0	0	0	0
a3	4	8	66	0	11	11	0	0	0
a4	0	0	0	93	0	0	4	3	0
a5	0	0	7	0	91	2	0	0	0
a6	0	0	0	0	9	91	0	0	0
a7	0	0	0	0	0	0	90	10	0
a8	0	0	0	0	0	0	9	91	0
a9	0	0	3	0	3	0	0	0	94

Array (c)

	a1	a2	a3	a4	a5	a6	a7	a8	a9
a1	82	1	3	0	0	14	0	0	0
a2	2	35	51	0	10	2	0	0	0
a3	0	41	44	0	9	0	6	0	0
a4	0	0	0	96	0	1	0	1	2
a5	14	14	26	0	29	0	16	1	0
a6	0	0	0	13	0	85	0	1	1
a7	0	0	0	1	0	7	90	2	0
a8	0	1	0	1	0	0	44	52	2
a9	0	0	0	5	3	0	3	5	87

Array (d)

Matrix (a) show that, using order three for Zernike moments, some actions are misclassified. For example, jumping in place action presents a modest classification ratio (65). Matrix (b) show that the seventh order is the best order. In fact, it allows to obtain the better classification ratio for almost all actions. To improve these efficient results, they are compared with two other works [24][1]. The following arrays (c) and (d) present confusion matrix obtained with. The comparative study with other works prove that the proposed approach improve classification results for the most of actions.

5 Conclusion

This paper proposed a new approach for action classification in video based on 3D Zernike moments. In fact, after segmentation of the input video and a subtraction of the background, 3D Zernike moments are calculated on the 3D shape from silhouettes

found in the space-time volume. Finally LS-SVM is used to classify actions. The proposed method has been evaluated by carrying out experiments on the Weizmann databases. The results show a good classification rate for most of tested actions. In future work, we will extend the approach to the classification of videos containing body which makes different actions at same time and videos containing several bodies in action. In more, we can also treat the occlusion effect on the characterization of action.

References

1. Zelnik-Manor, L., Irani, M.: Event-Based Analysis of Video, Computer Vision and Pattern Recognition, Computer Vision and Pattern Recognition. In: Proceedings of the 2001 IEEE Computer Society Conference, vol. 2, pp. 123–130 (2001)
2. Wu, X., Ngo, C.W., Hauptmann, A.G., Tan, H.K.: "Real-Time Near-Duplicate Elimination for Web Video Search with Content and Context. IEEE Transaction on Multimedia 11, 196–207 (2009)
3. Efros, A., Breg, C., Mori, G., Malik, J.: Recognizing Action at a Distance. In: Computer Vision Proceedings. Ninth IEEE International Conference, vol. 2, pp. 726–733 (2003)
4. Dollar, P., Rabaud, V., Cottrell, G., Belongie, S.: Behavior recognition via sparse spatio-temporal features. In: International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pp. 65–72 (2005)
5. Gavrila, D.: The visual analysis of human movement: A survey. Computer Vision Image Understand. 73, 82–98 (1999)
6. Cedras, C., Shah, M.: Motion-based recognition: A survey. Image Vision Computer 13, 129–155 (1995)
7. Rao, C., Shah, M.: View-invariance in action recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 316–321 (2001)
8. Song, Y., Goncalves, L., Perona, P.: Unsupervised learning of human motion. IEEE Transaction on Pattern Analyses and Machine Intelligence 25, 814–827 (2003)
9. Yacoob, Y., Black, M.: Parameterized modeling and recognition of activities. Computer Vision on Image Understand. 73, 232–247 (1999)
10. Ali, A., Aggarwal, J.: Segmentation and recognition of continuous human activity. In: Procedure of Intelligent Workshop on Detection and Recognition of Events in Video, pp. 28–35 (2001)
11. Bobick, A., Davis, J.: The recognition of human movement using temporal templates. IEEE Transaction on Pattern Analyses and Machine Intelligence 23, 257–267 (2001)
12. Weinland, D., Ronfard, R., Boyer, E.: Motion history volumes for free viewpoint action recognition. Presented at the IEEE Workshop Modeling People and Human Interaction, pp. 87–89 (2005)
13. Gorelick, L.: al Actions as Space-Time Shapes. IEEE Transactions Pattern Analysis and Machine Intelligence 29, 2247–2253 (2007)
14. Guo, K., Ishwar, P., Konrad, J.: Action Recognition in Video by Covariance Matching of Silhouette Tunnels. In: XXII Brazilian Symposium on Computer Graphics and Image Processing, pp. 299–306 (2009)
15. Kellokumpu, V., Pietikainen, M., Heikkila, J.: Human activity recognition using sequences of postures. In: IAPR Conference on Machine Vision Applications, p. 6570–6573 (2005)

16. Sminchisescu, C., Kanaujia, A., Li, Z., Metaxas, D.: Conditional models for contextual human motion recognition. In: Procedure International of Conference on Computer Vision, vol. 2, pp. 1808–1815 (2005)
17. Shutler, J.D., et al.: Statistical gait recognition via temporal moments. In: Image Analysis and Interpretation 4th IEEE Southwest Symposium, pp. 291–295 (2000)
18. Baccouche1, M., Mamalet, F., Wolf, C., Garcia, C., Baskurt, A.: Une approche neuronale pour la classification d'actions de sport par la prise en compte du contenu visuel et du mouvement dominant. In: Compression et Représentation des Signaux Audiovisuels (CORESA), pp. 25–30 (2010)
19. Tian, Y., Liu, Z., Yao, B., Zhang, Z., Huang, T.: Action Detection Using Multiple Spatial-Temporal Interest Point Features. In: IEEE International Conference on Multimedia and Expo (ICME), pp. 340–345 (2010)
20. Suykens, J.A.K., Vandewalle, J.: Least Squares Support Vector Machines. Neural Processing Letters 9(3), 293–300 (2002)
21. Gestel, T.V., Suykens, A.K.: Benchmarking Least Squares Support Vector Machine Classifiers, Technical report of Internal Report 00-37 on ESAT-SISTA 54, 5–32 (2000)
22. Blank, M., Gorelick, L., Shechtman, M., Irani, M., Basri, R.: Actions as space-time shapes. IEEE Transactions on pattern Analysis and Machine Intelligence 29, 2247–2253 (2005)
23. Green, R., Guan, L.: Quantifying and recognizing human movement patterns from monocular video images. IEEE Transaction Circuits on System Video Technologys 14, 179–190 (2004)
24. Dhillon, C.H., Nowozin, P.S., Lampert, S.: Combining appearance and motion for human action classification in videos. In: Computer Vision and Pattern Recognition Workshops, pp. 22 – 29 (2009)

Visual Information Based Argument Categorization for Semantics of Chinese Verb

HaiPeng Liu[1], XiaoJie Wang, and YiXin Zhong

[1] Center for Intelligence Science and Technology,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
pliu1981@gmail.com

Abstract. Recently, language acquisition with aids of multi-modal information have drawn more and more attention. However, semantic grounding of verbs has been less concerned due to their complex semantic representation. This paper proposed a novel way to combine visual information into semantic representation of Chinese verb. While introducing original representation of two constituents, which are verb frame and argument from Frame Semantic, both of them are linked with visual information for verb semantic. And a visual information based categorization for arguments is mainly discussed. For achieving it, a collection of {video, its text description} pairs is first built. After preprocessing on both sides, the correspondence between arguments of verbs and related visual features is constructed basing on SOM groups. A video describing system has also been built to generate sentences for new videos. The evaluation of the describing system shows the effectiveness of our visual semantic representation on Chinese verbs.

Keywords: verb semantic representation, visual grounding, argument categorization, Self-Organizing Map.

1 Introduction

Lexical semantic analysis is an important research topic in Natural Language Processing. In most existing theories and technologies, representations of semantics are based on relations between words or concepts. Briefly, it is to explain one word conceptually by some other words or relations with other words [1-3]. This type of semantics has been widely applied in many fields such as machine translation and question answering systems [4, 5]. However, it can do little in some other tasks, for example, situated human-machine interaction, automatic text description for images and so on. The main reason is that the linguistic words have no relations with perceptive information in this type of semantics. To link language and perception, a new trend is to learn semantics from various sensorimotor information, among which, vision cognition and its relationship with language ability have gained special attention. However, present researches mainly focus on nouns and adjectives comparing with less attention on verbs [6-8]. This is caused by complexities of verb representation [9-11].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 206–213, 2011.

Complexities on visual semantic acquisition of verbs mainly exist on both aspects of vision and language. On vision, scenes described by verbs are dynamic events occurring within space and time. These scenes with scalability on space and time need to be depicted with sequential visual information. On language, performances of verbs in sentences always need the participants of other lexical categories [10]. Verbs have a dominant function to these lexical categories. For example, a video description sentence always has a structure which takes a verb as its head word to dominate a few other nouns.

Till now many works on verb grounding are still at the initial stage. Siskin's LEONARD system tries to use force dynamics as semantic of verbs for event classification [12]. Pangburn's EBLA extend these verbs into protolanguage [13]. Fleischman and Roy embodied the idea of multi-modal information learning into sports video retrieval [14]. Satish has built an Image Schema representation to describe relative motions between two objects [15]. However, these works mainly focus on the linking process between verbs and visual features, ignoring the proper representation on origin verb semantic for a learning system. So that acquired verbs in above works are all transitive verbs with simple argument structure, which leads to simple SVO sentence structure. Verbs are only conjunctions between subject and object. By this way, visual representations of above verb semantics have no substantial difference with nouns and adjectives.

This paper tries to build a semantic representation which can reflect linguistic structure of verbs and which is simultaneously based on visual information. Thus, we start our work from both linguistics and vision-language association. On linguistics, the structure of verb semantic is firstly defined basing on Frame Semantics, which includes two constituents of frame and arguments. On vision-language association, we use Self-Organizing Map to build connections between verb arguments and visual features through a certain scale of video–text paired corpus. Thus we can get a visual semantic representation on verbs.

2 Video-Based Meaning Acquisition Model of Chinese Verbs

Video based Meaning Acquisition of Chinese Verb (ViMac-V) is a supervised model whose structure is shown in Fig. 1. ViMac-V takes video-text paired inputs as training samples to constructed verb semantic representation based on video information.

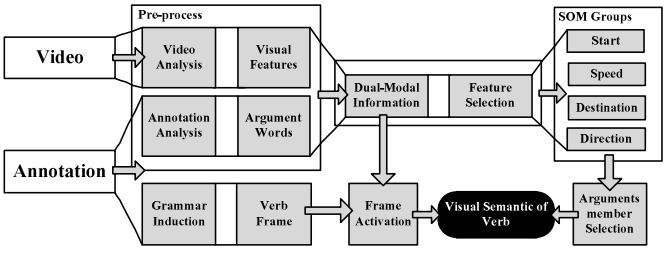


Fig. 1. Structure of ViMac-V System

2.1 Verb Semantic Representation Based on Frame Semantics

Gillette believed that verb has an excellent ability on describing dynamic events in space and time [11]. But this ability cannot be realized without supports from other basic lexical semantics, especially for satellite-framed language such as English or Chinese [16]. Thus representation of semantic on verbs needs participants of other basic lexicons and necessary grammar rules to govern these argument. This linguistic way synthesizing grammar information in semantic construction is named ***Frame Semantics*** [17]. According to Frame Semantics, our verb semantic representation is constituted by two parts:

$$\text{Verb} = \text{Frame} + \text{Arguments}$$

In above representation, frame is used by the verb to dominant other arguments. Then detailed description can be realized through selection on various arguments. With these two constituents, verb's dominant function and its relation with grammar, which distinguish verbs from nouns and adjectives, can be well embodied.

2.2 Pre-process on Video-Text Paired Corpus

2.2.1 Video Feature Analysis

Video feature analysis is to extract space-time visual features related with motion concepts from raw videos. Here our defined visual features form a 12-dimension vector:

$$V_f = [sn, x_i, y_i, x_f, y_f, s, d_x, d_y, sp_x, sp_y, ep_x, ep_y]^T$$

Details of each visual feature are shown in Table 1. In Table 1 the nine-square grid coordinate system is to equally divide the motion area into a 3×3 plots. Then the label of each plot is used to substitute the origin coordinate value. It can be seen as a data reduction method to coordinate values of start and end points, the over-averaged distribution of which will affect later categorization result (Section 4.4).

2.2.2 Annotation Corpus Analysis

For Chinese description, segmentation and POS tagging are first carried, then frames are extracted using bigram model [18]. For arguments abstraction, according to the frequencies and POS information, the tagged words and expressions are sorted into different classes to form the argument W_A. There are total 4 classes of argument lexicons which are $\{W_A^{speed}, W_A^{dir}, W_A^{start}, W_A^{des}\}$ and 5 verb frames of $\{F_0, ..., F_4\}$ extracted from annotation text corpus. In evaluation, output sentences are organized by the frame with categorized argument lexicons, for example:

小球在中间慢慢的向右运动到尽头

A small ball moves right slowly from middle to the end

Table 1. Defined space-time features

Visual feature	Physical Meaning	Descriptions
sn	File index	Indicate the sequence of video files or annotation sentences
x_i	Position	Abscissa value of start point
y_i	Position	Ordinate value of start point
x_f	Position	Abscissa value of end point
y_f	Position	Ordinate value of end point
s	Speed	Speed value of moving subject
d_x	Direction	Difference of abscissa values between start and end point: $d_x = x_i - x_f$
d_y	Direction	Difference of ordinate values between start and end point: $d_y = y_i - y_f$
ns_x	Position	Abscissa value of start point under nine-square grid coordinate
ns_y	Position	Ordinate value of start point under nine-square grid coordinate
ne_x	Position	Abscissa value of end point under nine-square grid coordinate
ne_y	Position	Ordinate value of end point under nine-square grid coordinate

3 Argument Categorization Based on SOM Groups

In ViMac-V, defined verb semantic needs to be linked with visual information. Till now, associations between visual information with both frame and arguments are complete. Due to space limitations, this paper will mainly focus on argument categorization process.

3.1 Modeling and Tranning of SOM Groups

Self-Organizing Map [19] has been widely applied in Feature Extraction and Large-scale Data Mining [20, 21]. In ViMac-V, argument categorization process is through 3 steps: SOM Training, Neuron Clustering and Lexicon Acquisition.

SOM groups for categorization are made up by 4 sub maps, which correspond to 4 arguments. We take speed sub map as an example for detailed explanation. Input signal $x = [x^1, x^2, \ldots, x^n]^T$ can be divided into two parts, such as $[V^{speed}, W_A^{speed}]^T$. Visual part V^{speed} contains visual features extracted from V_f (Section 2.2.1). And linguistic part W_A^{speed} is made up by speed argument lexicons. After x is received by speed sub map, each neuron c_i competes with each other to match x. The nearest neuron c_w to x is activated. Then weight adjustment is following WTM rules to c_w as well as its neighbor neurons. The coverage and strength of adjusted neurons is decided by the neighborhood function N_w, which can be replaced by Gaussian Function in practice. Through competition and weight adjustment, the distribution of diverse speed information sensed on vision can be reflected on the speed sub map.

3.2 Neuron Clustering and Lexicon Acquisition

After map training, distribution of speed sensory on the map plane still needs to be transformed into explicit expressions in order to be linked with argument lexicons. Our clustering method is a top-down agglomerative algorithm, in which the distance measurement adopts ward distance [22]. Then each cluster of neurons represents a lexical category C_i^{speed} in speed argument. C_i^{speed} needs to be associated with linguistic feature W_A^{speed} to acquire its language concept. In previous training step, each adjusted neuron c_i records the linguistic feature W_A^{speed} carried by input signal x. The record with the maximum counting number will be regarded as the language concept of c_i. As to the entire category, its language concept is acquired with also the maximum counting number principle. Each linguistic feature of c_i is counted to choose the maximum one as the lexicon of the entire category C_i^{speed}. Then the categorization for speed argument is completed. After the categorization of all 4 sub maps, their category members can be chosen and dominated by different frames to constitute different verb semantics.

4 Experiments and Evaluation

4.1 Categorization Experiments of SOM Groups

1000 videos of a small ball's straight-line motions are manufactured. It is assumed that all the static concepts describing its shape or color are already acquired in previous version of ViMac-V [11]. Annotated sentences are all opening natural language which reflect full cognitive responses of annotators when they are observing videos. The categorization results of each sub map are shown in Figure2. There are 3 categories in speed argument, 4 in direction and 9 categories both in start and destination categories. In Fig 2(a), the hierarchy of speed argument is very clear. The 3 categories means 快速(fast), 中速(medium) and 慢速(slow) respectively. While to the other arguments, they have to face more categories and diversed linguistic features. So that their acquisition results are more dependent on the effects of feature selection.

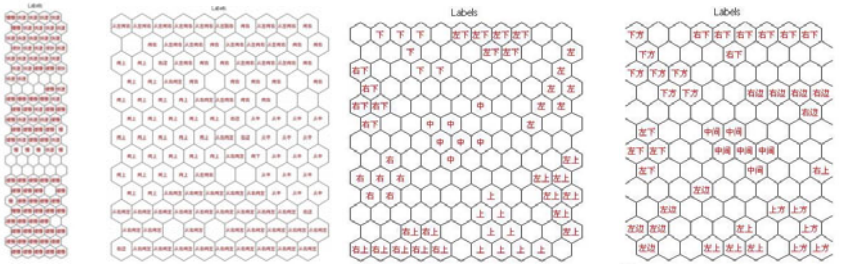


Fig. 2. From left to right: (a) Speed argument; (b) Direction argument; (c) Start argument; (d) Destination argument

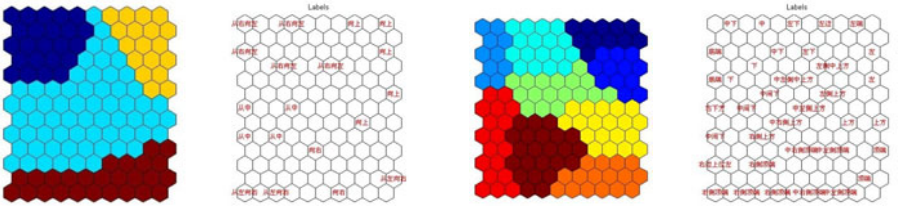


Fig. 3. From left to right: (a) Clustering result with original visual features for direction arguments; (b) Acquisition result with original visual features for direction arguments; (c) Clustering result with original visual features for start arguments; (d) Acquisition result with original features for start arguments.

4.2 Effects of Feature Selection to Categorization

When selecting visual features, besides the physical meaning of arguments, the effects to clustering and lexicons acquisition should also be concerned. Take direction argument as an example. Assuming it simply choose $[d_x, d_y]$ as visual features. The results of categorization are shown in Fig 3(a) and (b). Because the training data densely distributes in two-dimensional visual space V_{d_x,d_y}, most neurons on the map plane are clustered into one class (the middle part with light color in Fig 3(a)). The rest neurons will suffer reduction on activation chances. And these will lead to bad lexicon acquisition results (Fig 3(b)). However, if additional coordinate values of start and destination $[x_i, y_i, x_f, x_f]$ are added into visual features, the two-dimensional visual space will be extended to six-dimensional. And dense distribution of training data will be more spread to help reducing dead zones. The results of categorization will get improved (Fig 2(b)).

As to the arguments of start and destination, training data scatter too evenly in visual spaces V_{x_i,y_i} and V_{x_f,y_f}. It is because the start and end positions of the small ball cover the entire screen uniformly. Meanwhile, different annotators have diversified subjective expressions on describing location concepts. So that it is difficult to get a consensus linguistic representation for location categories (Fig 3(c) and (d)). To solve these problems, we improve feature selection from both vision and language aspects. As to vision, we add nine-square grid features $[sp_x, sp_y]$ and $[ep_x, ep_y]$ besides coordinate values. As to language, nine-square grid labels replace original linguistic expressions. Then lexicon acquisition of location arguments can be a two-layer structure. One uses natural lexicons and the other uses nine-square grid labels (Figure 2(c), (d)). Thus in generation stage, ViMac-V can choose to output natural lexicons or regular nine-square grid labels.

4.3 Evaluation of Description Output

Categorization results can be directly used in ViMac-V together with an activated verb frame to generate text description for videos. Since we focus on argument categorizations, the verb frame is temporarily set to motion frame F_0. It means all four arguments will be activated and the verbs are used with 运动(*move*) uniformly.

The test videos are 20 unlabeled AVI files. Three Chinese native speakers are invited to give their evaluations. Figure 4 shows some videos and their description sentences generated by ViMac-V. Through evaluation, the descriptions generated by ViMac-V to these 20 videos are all confirmed by 3 experimenters. However, the generated sentences with improved feature selection get faster consensus. While to the sentences using original features, there exist certain disputes on describing start and destination. It is difficult to get a fast agreement among 3 experimenters.

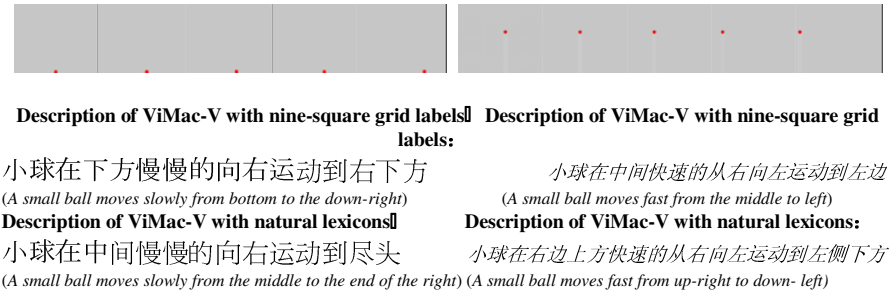


Fig. 4. Examples of output sentences by ViMac-V. English sentences in brackets are just free translation of above Chinese sentences

5 Conclusions

This paper proposes a Chinese verb semantic representation based on both visual and linguistic information. Evaluation proves that SOM can be used for visual categorization of different arguments. The improvements on feature selection can avoid existing subjective diversities in annotation corpus. Consistency between generated sentence of ViMa-V and normal human cognition proves the validity of our visual semantic representation on verbs.

However, evaluation on output sentences is still limited to small-scale corpus. And the dynamic visual scenes are simulated videos manufactured by computer. It lacks the evaluation under real situated circumstances. Also, we want to transplant ViMac-V onto a robot platform. All these works will be further studied in the future.

Acknowledgment

This research was supported by the major research plan of National Nature Science Foundation of China (90920006), and by Research Fund for the Doctoral Program of High Education of China (20090005110005).

References

1. Palmer, M., Kingsbury, P., Gildea, D.: The Proposition Bank: An Annotated Corpus of Semantic Roles. Computational Linguistics 31(1), 71–106 (2005)
2. Sameer, S., Pradhan, E.H.H., Marcus, M.P., et al.: Ontonotes: a Unified Relational Semantic Representation. Int. J. Semantic Computing 1(4), 405–419 (2007)

3. Collin, F., Baker, C.J.F., Lowe, J.B.: The Berkeley FrameNet project. In: Proceedings of the 26 Annual Meeting of the ACL and 7th International Conference on Computational Linguistics, San Francisco, California (1998)
4. Sameer Pradhan, V.K.W.W., Jurafsky, D., Martin, J.H.: Using Semantic Representations in Question Answering. In: ICON 2002, Bombay, India (2002)
5. Karin Kipper Schuler, A.K.: VerbNet overview, extensions, mappings and applications. In: NAACL 2009, Boulder, Colorado (2009)
6. Ballard, C.Y.a.D.H.: On the Integration of Grounding Language and Learning Objects. American Association for Artificial Intelligence (2004)
7. Kubat, R.: Semantic Context Effects on Color Categorization. In: Proceedings of the 31st Annual Cognitive Science Society Meeting, Amsterdam (2009)
8. Zhang, C., Zhang, W., Liu, H., Wang, X.: Visual Information based Meaning Acquisition of Chinese. In: 10th Chinese National Conference on Computational Linguistics, Shan Dong Province, China, pp. 260–266 (2009)
9. Gentner: Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In: Language Development, NJ (1982)
10. Gleitman, L.: The structural sources of verb meanings. Language Acquisition (1990)
11. Jane Gillette, H.G., Gleitman*, L., Lederer, A.: Human simulations of vocabulary learning. Cognition 73, 135–176 (1999)
12. Siskind, J.M.: Grounding the Lexical Semantics of Verbs in Visual Perception using Force Dynamics and Event Logic. Journal of Artificial Intelligence Research 15, 31–90 (2001)
13. Pangburn, et al.: EBLA: a perceptually grounded model of language acquisition. In: Proceedings of the HLT-NAACL, NJ, USA (2003)
14. Fleischman, M., Roy, D.: Unsupervised content-based indexing of sports video. In: International Multimedia Conference, Augsburg, Germany (2007)
15. Mukerjee, G.S.: Acquiring Linguistic Argument Structure from Multimodal Input using Attentive Focus. In: Development and Learning, ICDL 2008, Monterey, CA (2008)
16. Talmy, L.: Toward a Cognitive Semantics. MIT Press, Massachusetts (2000)
17. Fillmore, C.J.: Frame semantics, in Linguistics in the Morning Calm, e. In: The Linguistic Society of Korea, Seoul, pp. 111–137 (1982)
18. Collins, M.: A new statistical parser based on bigram lexical dependencies. In: Proceedings of the 34th Annual Meeting of the Association of Computational Linguistics, Santa Cruz, CA (1996)
19. Kohonen, T.: Automatic formation of topological maps of patterns in a self-organizing system. In: Proc. 2SCIA, Scand. Conf. on Image Analysis, Helsinki, Finland (1981)
20. Jain, J.M.a.A.K.: Artificial Neural Networks for Feature Extraction and Multivariate Data Projection. IEEE Transactions on Neural Networks 6(2), 296–317 (1995)
21. Jorma, T., Laaksonen, J.M.K., Erkki, O.: Class distributions on SOM surfaces for feature extraction and object retrieval. Neural Networks 17, 1121–1133 (2004)
22. Rousset, P., Guinot, C.: Distance between Kohonen Classes Visualization Tool to Use SOM in Data Set Analysis and Representation. In: Mira, J., Prieto, A.G. (eds.) IWANN 2001. LNCS, vol. 2085, p. 119. Springer, Heidelberg (2001)

A Novel Remote Detection Method of Illegal Electricity Usage Based on Smart Resistance

Byambasuren Bat-Erdene[1], Sang-Yep Nam[2], and Dong-Han Kim[1]

[1] Department of Electrical Engineering, Kyung Hee University, Korea
[2] Department of IT, Kookje University, Korea
{baterdene,donghani}@khu.ac.kr, r13337@unitel.co.kr

Abstract. A power delivery system that uses power line communication can communicate with other electrical systems, monitor the quality of electrical energy, and create an economical solution for automatic meter reading. If we use a smart meter with the smart resistance proposed in this paper, we can easily add an illegal electricity usage detection capability to an electrical power delivery system for automatic meter reading. The main current of the circuit can be calculated from the value of the smart resistance and compares it with the measured main current. A mismatch between the measured current and the calculated current implies there is either illegal electricity usage inside the smart meter's circuit. This paper presents the architecture of the system, the algorithm of the detection system, the model of the system, and the results of the simulation.

Keywords: Automatic meter reading (AMR), illegal electricity usage, power line communication, remote detection.

1 Introduction

Information can be transmitted economically through electrical power transmission lines using power line communication (PLC) with no extra wiring [1]-[9]. PLC communication can also be used for many other purposes such as the following: to communicate with other electrical devices, to monitor the power quality, to control the optimal flow of an electrical load, and to implement an automatic meter reading (AMR) system. Remote monitoring of the electricity consumption of every user will be possible if an AMR system can be implemented in a power delivery system [10]-[15].

There is already an emerging possibility of monitoring the electricity consumption of customers through the use of an AMR system based not only on PLC communication but also broadband over power lines (BPL) [16].

An AMR system in a power delivery system can be used to remotely detect any illegal consumption of electricity in the system. Illegal usage of electricity is a major problem for electricity distribution companies. For instance, the Minister of Energy and Natural Resources in Turkey said the illegal usage of electricity constituted 19 percent of Turkey's total electricity consumption [17].

The same problem exists in Iran [18]. And even Mongolia is now seeking an optimal solution for remote detection of illegal electricity usage [19]-[21].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 214–223, 2011.

The online detection of illegal electricity usage is not yet possible through methods that involve statistical data derived from a time domain reflectometer in the main power cable [22]-[25]. In order to detect a fault of power line and location of illegal connections, many companies/countries use detection methods based on a time domain reflectometer in the power cable of consumers. However, these methods are unsuitable for remote online detection of illegal electricity consumption. There is consequently a growing need for an optimal means of remotely detecting illegal electricity consumption.

2 Detection of Illegal Electricity Consumption

2.1 History

In October 2004, H.Cavdar suggested a new AMR-based method of detecting illegal electricity consumption. He assumed that electricity could be used illegally in the following four ways:

1. Use of mechanical objects.
2. Use of a fixed magnet.
3. Use of an external phase before the meter terminals.
4. Switching of energy cables at the meter box.

H. Cavdar's basic method involves the use of two electrical meters: one is the user's meter; the other is placed outside the user's residence. Any variation in the energy measurements of the two energy meters confirms the illegal usage of electricity [17]. H. Cavdar's method can detect any illegal consumption of electricity between the points where the extra energy meter and the main energy meter are installed. However, in many countries, illegal consumption is rife through a street-level connection to the main power line without any energy meter. H. Cavdar's method requires the use of an extra energy meter and it can only be used on a single low voltage grid.

In 2007, A. Pasdar and S. Mirzakuchaki suggested that the illegal usage of electricity could be detected through a remote method based on a smart meter. The smart meter disconnects all users to transmit a low voltage signal of high frequency through the main power line. Their method detects illegal usage of electricity by measuring and comparing the impedance value with that of a normal regime [18]. However, the main drawback of their method is the fact that all users of the power delivery system must be disconnected.

This paper presents a new approach that enables illegal usage of electricity to be detected remotely without having to disconnect users.

2.2 New Method for Finding Illegal Connections

Whenever electricity is used illegally directly from the power line in any of the four ways mentioned above, the new method uses a resistance that called smart resistance for remote real time detection of the illegal usage. We assume that the smart resistance is changed value of resistance by a control system. Figure 1 shows application of the smart resistance.

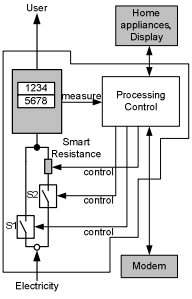


Fig. 1. Block diagram of smart metering with smart resistance

The smart meter in the figure is installed in the AMR system of a substation. The smart meters are controlled by the transmission of information to each other. A smart meter with smart resistance can detect illegal usage of electricity by instantly changing the voltage of the user during a small interval and this small value of voltage is depended by the electric load of the user.

The proposed detection system compares the value of the main energy meter on the substation with the value received from each user's smart meter. It then checks the balance of the values online. An imbalanced result confirms that electricity in the power delivery system is being illegally consumed. The following formula is valid when there is no illegal consumption of electricity:

$$E_{main} = E_1 + E_2 + \ldots + E_n + E_{error} \tag{1}$$

There E_{main} is value of the main energy meter on the substation; E_1; E_2; … E_n are values of each user's smart meter; E_{error} is the energy lost, including the energy lost on the line of the substation. When the AMR system detects illegal usage of electricity, it sends the following command to each user's smart meter: "check illegal consumption of electricity". The circuit of smart metering with smart resistance is shown in Figure 2.

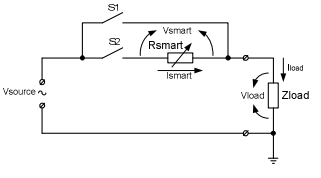


Fig. 2. The circuit of smart metering with smart resistance

When the control command is received, the control system of the smart meter turns *S1* off and *S2* on. In this case, the smart resistance is connected by *S2* and the control system saves data about the value of the smart resistance, the value of the energy meter's voltage and current. The control system then disconnects the smart resistance and turns *S1* on. From the value of the changed smart resistance, we can calculate the current and voltage of the smart resistance. Because the smart resistance connects

with load of electricity in series, the current of the user should be the same as the current of the smart resistance. From this circuit, we can easily find the current of smart resistance following form:

$$I_{smart} = V_{smart} / R_{smart} = (V_{source} - V_{load}) / R_{smart} \quad (2)$$

From the measurement of smart meter and control system, we know next parameters: complex value of load current and voltage - I_{load}; V_{load}; complex value of source voltage - V_{source}; smart resistance - R_{smart}. If the measurement of smart meter is true which means there is no illegal usage, the calculated current I_{smart} and the measured current of smart meter I_{load} are same:

$$I_{smart} = I_{load} \quad (3)$$

If condition (3) is false, we can detect illegal electricity usage in this smart meter.

In the next section, we discuss the simulation of a smart meter with smart resistance and present our results.

3 Simulation Results and Detection Algorithm

3.1 Simulation Results

Using Simulink tools of MATLAB, the detection system can be experiment a normal state (the absence of any illegal electricity usage) and an illegal state (the presence of illegal electricity usage). In the illegal state, we will carry out the simulation that is same following four illegal usages: to use either a fixed magnet or mechanical objects; to switch the energy cables at the meter connecter box and to use external phase before the meter terminals.

The simulation circuit of the detection system under normal state is shown in Figure 3. In order to easily see the result, load resistor is chosen an active resistance and this circuit can measure real voltage of source, real current of source, voltage of load and current of load. We assume that *R shunt* resistance illustrated by the Figure 3 is taken infinite in normal state. Figure 4 shows the simulation results in normal state. We implement simulation following conditions. First, *Breaker 1* is switched on in 0.02 sec. Then, *Breaker 2* is switched off and *Breaker 1* is switched off in 0.06 sec. Finally, both *Breaker 1* and *2* are switched off in 0.1 sec. When smart resistance is connected in normal state, the simulation yields the same result for both the calculated current and the electric meter current, on the other hand an illegal electricity usage does not exist on this smart meter.

In shunt state, we change *R shunt* resistance illustrated by the Figure 3 to zero. When we perform simulation in illegal electricity usage of condition “to switch the energy cables at the meter connector box” and “to use external phase before the meter terminals”, *R shunt* resistance is chosen zero value.

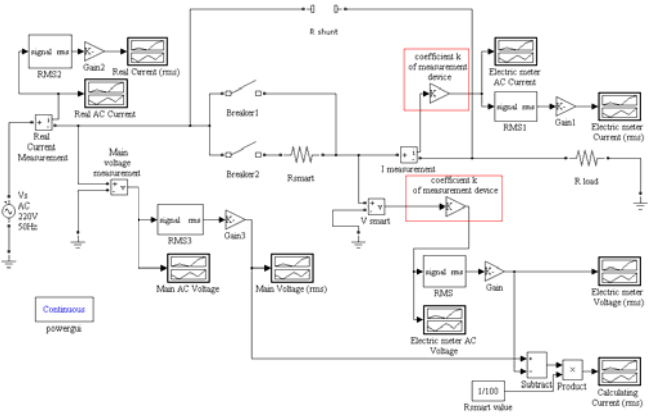


Fig. 3. The simulation circuit of the detection system under normal state

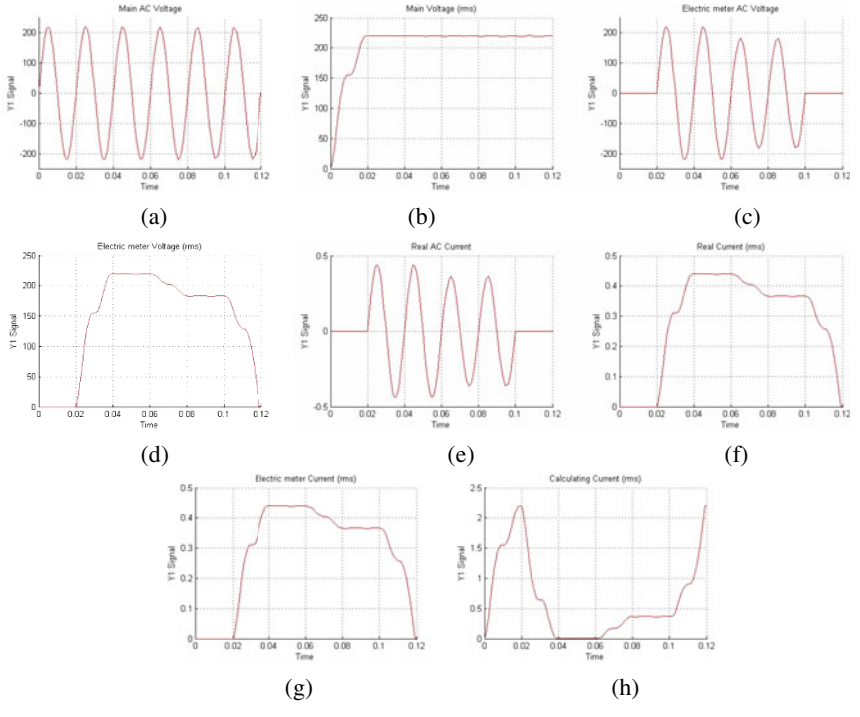


Fig. 4. Simulation results of the normal state: (a) shows the main AC voltage of the source, (b) shows the root mean square value of main voltage of the source, (c) shows the electric meter AC voltage under load resistance, (d) shows the root mean square value of electric meter voltage under load resistance, (e) shows the real AC current of the source, (f) shows the root mean square value of real current of the source, (g) shows the root mean square value of electric meter current under load resistance, (h) shows the root mean square value of calculating current.

In this case, the simulation let call shunt state. In shunt state, the system is subject to another condition where the energy cables are switched at the meter connector box during the external phase before the meter terminals.

Figure 5 shows simulation results of the shunt state. As shown in Figure 5, the calculated current and the current of the energy meter are zero which means no charge for users. But when smart resistance is connected, voltage of the electric meter cannot be changed by any value. Therefore, the smart meter can confirm the presence of illegal electricity usage on this user.

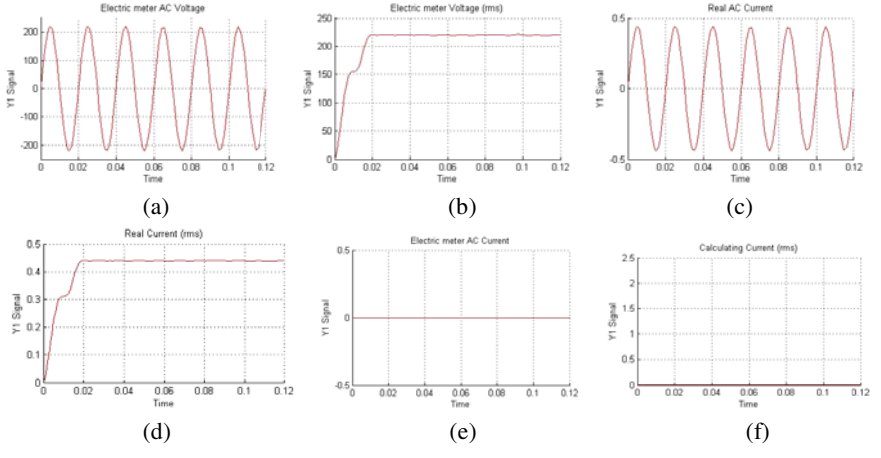


Fig. 5. Simulation results of the shunt state: (a) shows the electric meter AC voltage under load resistance, (b) shows the root mean square value of electric meter voltage under load resistance, (c) shows the real AC current of the source, (d) shows the root mean square value of real current of the source, (e) shows the electric meter AC current under load resistance, (f) shows the root mean square value of calculating current.

Illegal usage type of using fixed magnet or mechanical objects are same issue which the coefficient k of measurement devices for electric meter current and voltage are changed by simulation circuit in Figure 3. When we perform simulation in illegal electricity usage of condition "to use either a fixed magnet or mechanical objects", the coefficient k of measurement devices for electric meter current and voltage are chosen 1/2 and zero. First, the experiment is carried out condition that the coefficient k of measurement devices for electric meter current and voltage are chosen 1/2. Figure 6 shows the simulation result of the coefficients k of measurement devices for electrical meter current and voltage are chosen to 1/2. From results for changing coefficient of measurement devices, the calculated current and the electric meter current are not same when the smart resistance is connected. In this case, we know what an illegal electricity usage is presence. When a fixed magnet was used illegally, we can think alike above experiment which the coefficient k of the measuring device is chosen 1/2. Figure 7 shows the simulation result of the coefficients k of measurement devices for electrical meter current and voltage are chosen to zero.

The experiment which the coefficient *k* of the measuring device is chosen zero is same condition for a mechanical objects were used illegally. The results show that the calculated current of the connected smart resistance differs from the current of the energy meter. This difference means that the smart meter can be used to indicate any illegal electricity usage.

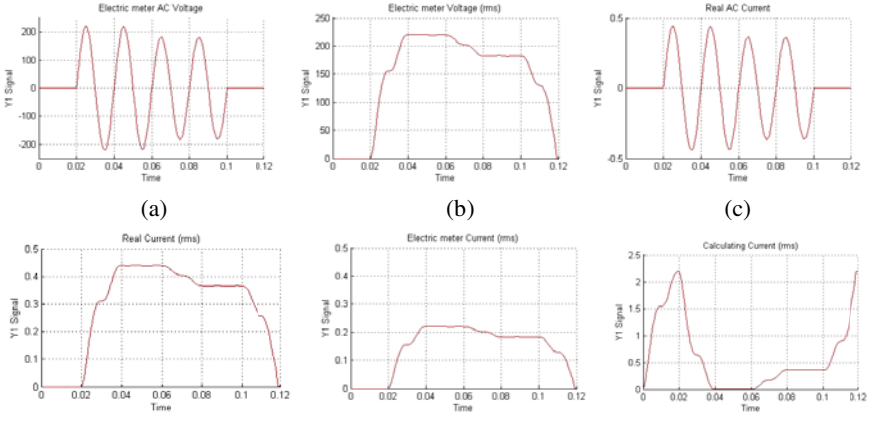


Fig. 6. The simulation result of the coefficient k of measurement devices for electrical meter current and voltage is chosen to 1/2 (*k=1/2*): (a) shows the electric meter AC voltage under load resistance, (b) shows the root mean square value of electric meter voltage under load resistance, (c) shows the real AC current of the source, (d) shows the root mean square value of real current of the source, (e) shows the root mean square value of electric meter current under load resistance, (f) shows the root mean square value of calculating current.

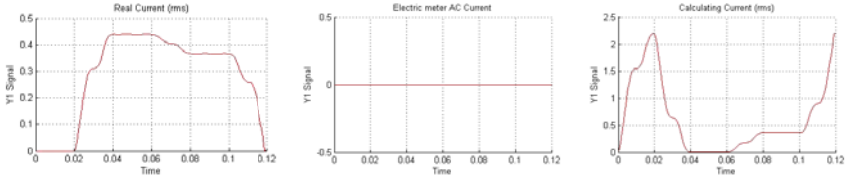


Fig. 7. The simulation result of the coefficient k of measurement devices for electrical meter current and voltage is chosen to zero (*k=0*): (a) shows the root mean square value of real current of the source, (b) shows electric meter AC current under load resistance, (c) shows the root mean square value of calculating current.

From above these methods, we can use the smart meter for remote online detection of any illegal electricity usage.

3.2 Detection Algorithm

We can use the simulation results to create the main algorithm for detecting illegal usage of electricity. The main algorithm is shown in Figure 8. From this algorithm,

we can use a smart meter on a power delivery system to detect and identify the type of illegal electricity usage. First of all, this algorithm checks either balanced or unbalanced main electric meters of substations. If the system detect unbalanced main electric meter of substation, the system starts detection process from this unbalanced substation. Then the detection process is implemented in smart meters of this unbalanced substation by the algorithm that illustrated Figure 8. There are several types of illegal usage. One type is the shunt state (which involves the use of an external phase before the meter terminals and the switching of energy cables at the meter connector box). Another type is a state where the system fails to measure the real current (by the use of mechanical objects or a fixed magnet). First, the detection process checks either electric meter current is zero or not zero and then if electric meter current is zero, we can know the conditions which user of smart meter is not load or shunt state of illegal usage, otherwise state of the system doesn't measure real current ($k=0$). When the smart resistance is connected by control system, if voltage of electric meter does not change or *Umain* main voltage of source is same with *Usmart* voltage of electric meter, we can detect an illegal usage for shunt state. If main voltage of source is not same with voltage of electric meter, the system can calculate ratio of voltage for source and voltage for electric meter, because current of user is zero. In this case, if the calculated voltage equals the voltage of electric meter, it is not illegal usage. Otherwise, it is condition that the system does not measure real current ($k=0$) and this is an illegal usage.

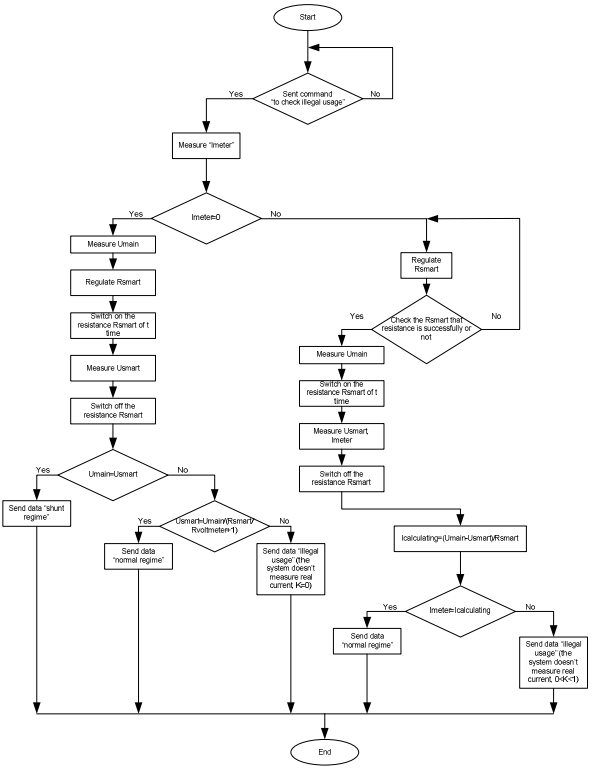


Fig. 8. The main algorithm of the detection system

If current of electric meter is not zero, we should compare current of electric meter and calculated current under the smart resistance. If current of electric meter equals with calculated current under the smart resistance, there is not any illegal usage in this smart meter. However, we can detect an illegal usage in smart meter, if current of electric meter is not same with calculated current.

If the smart meters do not indicate any illegal consumption of electricity, the illegal consumption must obviously exist outside of the system.

4 Conclusion

In this paper, we propose a novel method of using a smart meter with smart resistance to remotely detect illegal electricity usage. We also investigate the feasibility of a remote system of detecting illegal electricity usage.

The proposed system was simulated with the aid of Simulink modeling for MATLAB. The simulation results show that if an AMR system is installed in a power delivery system including a smart meter with smart resistance can detect any illegal electricity consumption on the power delivery system.

The detection algorithm, which was created from the simulation results, can detect any types of illegal electricity usage.

The following aspects of the system are expected to be investigated in future research: an automatic system for controlling the optimal value of the smart resistance; the control range of the smart resistance; experimentation on a real system; and testing on a power delivery system in Mongolia.

Acknowledgments. This research was supported by a grant from Construction Technology Innovation Program (CTIP) funded by Ministry of Land, Transportation and Maritime Affairs (MLTM) of Korean government and the Industrial Strategic Technology Development Program (No. 10035544-2010-01) funded by the Ministry of Knowledge Economy (MKE, Korea). The first author is grateful to Dr. Dong-Han Kim for your contributions to this research.

References

1. Ahola, J.: Applicability of power line communications to data transfer of on line condition monitoring of electrical drivers, Phd. Thesis, Dept. Electrical & Electronic. Eng., Acta Univ. (2003)
2. Zuberi, K.H.: Power line carrier communication systems, M. S. thesis, Dept. Microelectronic and information technology. Eng., Kista Univ. (2003)
3. Park, B.S., Hyun, D.H.: Implementation of AMR System Using Power Line Communication, pp. 18–24. IEEE, Los Alamitos (2002)
4. Lim, T.Y., Chan, T.W.: Experimenting remote kilowatt hour meter reading through low voltage power lines at dense housing estates. IEEE Trans. Power Delivery 17, 708–711 (2002)
5. Hong, D., Lee, J., Choi, J.: Power Quality Monitoring System Using Power Line Communication. In: Information, Communications and Signal Processing, Fifth International Conference, pp. 931–935 (2005)

6. Dostert, K.: Powerline communications. Prentice-Hall, Englewood Cliffs (2001)
7. Tanaka, M.: High frequency noise power spectrum, impedance and transmission loss of power line in Japan on intrabuilding power line communication. IEEE, Los Alamitos (1988)
8. Tanaka, M.: Transmission characteristics of a power line used for data communications at high frequencies. IEEE, Los Alamitos (1989)
9. Sado, W., Kunicki, J.: Personal communication on residential power line-assessment of channel parameters. IEEE, Los Alamitos
10. Newbury, J., Miller, W.: Multiprotocol routing for automatic remote meter reading using power line carrier systems. IEEE Trans.Power Delivery 16, 1–5 (2001)
11. Newbury, J.: Communications field trials for total utility metering. IEEE Trans. Power Delivery 11, 690–695 (1996)
12. Sridharan, K., Schulz, K.N.N.: Outage management through AMR systems using an intelligent data filter. IEEE Trans. Power Delivery 16, 669–675 (2001)
13. Mak, S., Radford, D.: Design considerations for implementation of large-scale automatic meter reading systems. IEEE Trans.Power Delivery 10, 97–103 (1995)
14. Miura, N., Sato, H., Narita, H., Takaki, M.: Automatic meter reading system by power line carrier communications. Proc. Inst. Elect. Eng. C 137, 25–31 (1990)
15. Arrington, R.A.: Load management and automatic meter reading through the use of power line carrier. Electr. Power Syst. Res. 4, 85–104 (1981)
16. Liu, Q., Zhao, B., Wang, Y., Hu, J.: Experience of AMR systems based on BPL in China. In: IEEE International Symposium Power Line Communications and Its Applications, ISPLC 2009, pp. 280–284 (2009)
17. Cavdar, H.: A solution to remote detection of illegal Electricity usage via power line communications. IEEE Trans. Power Delivery 19(4) (October 2004)
18. Pasdar, A., Mirzakuchaki, S.: A Solution to Remote Detecting of Illegal Electricity Usage Based on Smart Metering. In: 2th IEEE International Workshop on Soft Computing Applications, SOFA 2007, August 21-23 (2007)
19. Natsagdorj, C., Nyamtseren, B., Bat-Erdene, B.: Processing measurement information system based on electrical energy meter of distribution electricity and network in Erdenet industry. In: A Compilation Dedicated to 45th Anniversary of Power Engineering School, pp. 83–90 (2005)
20. Bat-Erdene, B., Lhagvademberel, B.: Registration device for registering of commutation device position, and its usage. In: Third International Forum on Strategic Technologies, IFOST 2008, pp. 385–387 (2008)
21. The annual report 2009 of East Distribution Center of Mongolia, East distribution center of Mongolia (2010)
22. Nagi, J., Mohammad, A.M., Yap, K.S., Tiong, S.K.: Non-Technical Loss Analysis for Detection of Electricity Theft using Support Vector Machines. In: 2nd IEEE International Conference on Power and Energy, pp. 907–912 (December 2008)
23. Nagi, J., Yap, K.S., Tiong, S.K.: Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines. IEEE Transactions on Power Delivery 25(2) (April 2010)
24. Pok, H.L., Yap, K.S., Abidin, I.Z., Hashim, A.H., Hussien, Z.F., Mohamad, A.M.: Abnormalities and fraud electric meter detection using hybrid support vector machine and modified genetic algorithm. In: 19th International Conference on Electricity Distribution, Vienna, May 21-24 (2007)
25. Kruno, T., Marinko, S., Darko, P.: Reduction of non-technical losses based on time domain reflectometer (TDR) Principles and function. In: 18th International Conference on Electricity Distribution, Turin, June 6-9 (2005)

Applied Method for One Source Multi Use (OSMU) in the Broadcasting Communication Convergence Environment

Hyunmi Jang, Sengphil Hong*, Kyongjin Kim, and Jae-Jung Kim

School of IT, Sungshin Women's University,
Seongbuk-gu, Seoul, Korea
{hmjang,kyongjin,philhong}@sungshin.ac.kr, jajukim@hotmail.com

Abstract. Users increasingly use portable devices such as smart phone, tablet PC, and hence increase the advancements of technology for electronic transactions. With the advancements of technology, the development of OSMU (One Source Multi Use) system has shown rapid growth, but they have had limited use in anywhere. This paper suggested ICDS (Interactive Contents Delivery System) architecture for effectively transmits and manages the one source multi use service method by using the diverse digital devices in the broadcast communication convergence environments.

Keywords: One Source Multi Use, Interactive Contents, Heterogeneous Devices, Broadcast Communications Convergence.

1 Introduction

Korea has the Internet access environment at the world-class level through infrastructure-oriented growth policies, it was based on the advent of mobile Internet era. With increasing use of smartphones and other mobile devices, the Internet usage has moved from the wireline to the wired and wireless environments[4].

With a paradigm shift from fixed data to mobile data[3], devices change have provided opportunities of new entertainment service including N-Screen service introduction that providing content across a variety of platforms such as the TV, PC and mobile devices to share and consume information. Wireless communications infra structure, however, is significantly more complex than the wire-based infrastructure, and resources are limited only by the total network capacity. We need to prepare that change is coming, mobile convergence environments that all things are connected freely available through network including extension of the Wi-Fi infrastructure, IPv6 transition, etc.

In this paper is organized as follows. Chapter 2 examines the status of mobile service in Korea. We discuss the issues of mobile convergence environments in chapter 3. Later, in chapter 4, we suggest the system for large-scale contents delivery that conducts 3 mechanisms. Then, we describe overall algorithm and prototype

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 224–229, 2011.

implementation in chapter 5. In chapter 6 we suggest the expected benefits. Lastly, chapter 7 concludes the papers and suggests the future research work.

2 Status of Mobile Service in Korea

According to the smartphone and smartpad propagation forecasts in Korea[1] by Korea Communications Commission, the rate of total mobile subscribers that takes about 12% overall is for the number of smartphone subscribers at the end of November. The number of domestic smartphone subscribers is expected to surpass 40 million populations by 2015.

Mobile traffic is expected to widen in the future, as well as mobile devices. Morgan Stanley[2] recently announced that mobile will be bigger than desktop Internet in 5 years. As increasing traffic of SNS community and mobile multimedia contents, domestic wireless data traffic is expected to increase about 40 times more in 2015 compared to 2010[1].

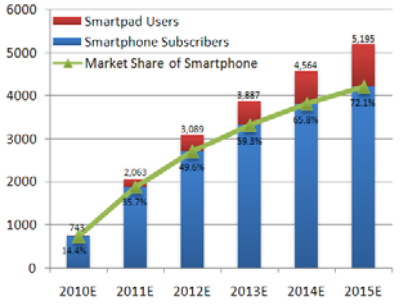


Fig. 1. The smartphone and smartpad propagation forecasts in Korea, 2010-2015E

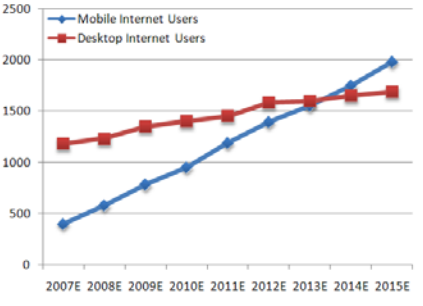


Fig. 2. Global mobile vs. desktop internet user projection, 2007-2015E

3 Issues in Mobile Convergence Environments

The mobile devices have become synonymous with a part of daily life for many people, allowing users to make calls, read e-mail, and get the job done. Although this has resulted in increased use of mobile device, many devices fail to meet the usability requirements of users.

There are problems that degrade the profitability of the mobile: portable devices that cannot support all utilizations because of limited resources such as low bandwidth, power consumption, etc; multimedia contents that may not be suitable for limited heterogeneous platforms; difficulties in high-speed transfer of large contents in wireless environments.

- Heterogeneous platform with limited resources
- Low-performance in wireless network

Thus, the issues faced by the IT society need to be resolved in order to support efficient mobile convergence service. In the following chapter, we suggest the system of ICDS (Interactive Contents Delivery System) to solve these several issues.

4 Interactive Contents Delivery System

We are propose to ICDS(Interactive Contents Delivery System) which is provided to transmission of interactive contents for diverse digital devices (N-Screen) when the one content is purchased through the user's diverse digital devices. ICDS is composed of 3 Mechanism: 1) UDAM (User & Device Authentication Mechanism) 2) ICPM (Interactive Contents Processing Mechanism) 3) LCAM (Large-scale Contents Accelerator Mechanism). (Fig.3) shows the represented how ICDS works between user's diverse digital devices.

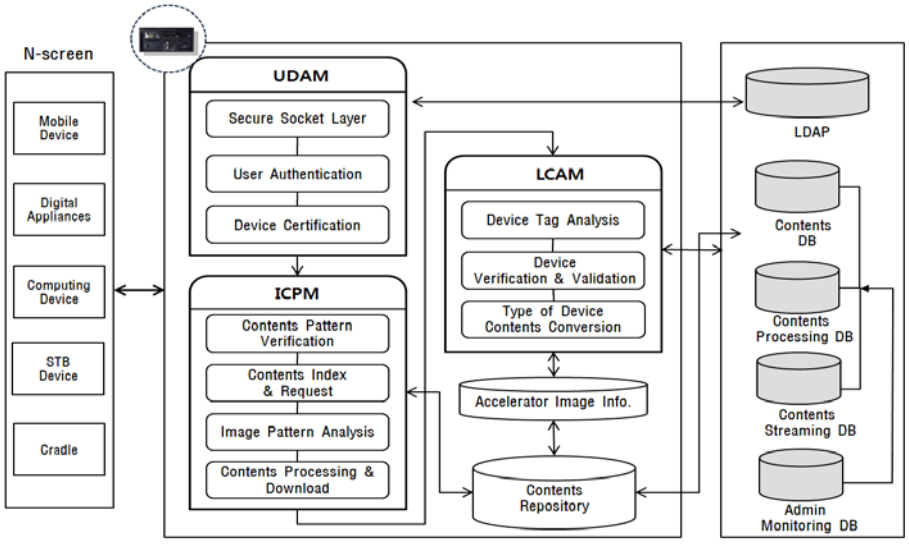


Fig. 3. The architecture of ICDS (Interactive Contents Delivery System)

User & Device Authentication Mechanism (UDAM) maintains the confidentiality of the subjects concerned with ICDS. Through the user's diverse digital devices, the certification given by the certification authority provides the verification and execution process of the user's identity and the final certification of the diverse digital device by CA. When the certification of the device is complete, it is encoded through the session. **Interactive Contents Processing Mechanism (ICPM)** is the given interactive contents service using image capture processing method, when authenticated user's diverse digital devices request for the contents. The contents or images selected by the users are extracted from the HTTP streaming servers as image (e.g., screen capture images) and then, the image on the device and the image to be request are analyzed and compared to see how extensive the changes are. If any change is detected, the relevant image is extracted and saved in the predetermined image file formats (e.g., jpg, png, bmp). Furthermore, when it is necessary to have the contents which have been already downloaded once, it is possible to send several image file formats to various heterogeneous devices. **Large-scale Contents Accelerator Mechanism (LCAM)** can quickly help transmit the requested contents

from user, after analyzing the diverse digital device's tag attribution (e.g., device type, device browser type, device mainboard model). The image file formats that can make the images or contents transferred fastest are preconfigured for each type of devices in Accelerator Image Info. So, the image information from the changed image is requested from the ICPM. At that time, the processing is done using XML & MPEG4 and to provide web based the interactive contents delivery service in a fast and scalable way.

5 ICDS Algorithm and Implementation

5.1 Algorithm

The ICDS (Interactive Contents Delivery System) can provide the interactive contents services swiftly by extracting the images in real time, which is required for each type of the diverse digital devices (N-Screen). In this section, we present an algorithm to provide these mechanisms (see Algorithm 1).

Algorithm Interactive Contents Delivery Management

```
1: // User & Device Authentication Mechanism
2:  if Login(user.id, user.pwd, user.cert) = true then
3:      pki.auth ← Check-Auth(user.id)
4:      cert ← CheckCert(pki.auth)
5:      return serialNum ← CA(CheckUserInfo(pki.auth))
6:          if Login (user.id, user.cert, serialNum, timestamp) = true then
7:              accessKey ← CheckCert (mpki.auth)
8:          else return false
9:   else SendDenyMsg(err, 'Access Deny')
10:       return false
11:  end if
12: // Interactive Contents Processing Mechanism
13:  if accessKey ≠ null then
14:      list_contents ← SearchContentsList(my_contents, accessKey)
15:      if list_contents = null then
16:          SendWarningMsg(err, 'Saved Contents Not Exist')
17:      else
18:          DBSave(list_contents)
19:          if UpdateContentsList(list_contents) = true then
20:              SendWebContentsSVR(my_contents, url)
21:              url ← GetHttpRequest(WebContentSVR)
22:              content_loc ← new ActiveContent(url)
23:              contents_loc.open('GET', url)
24:              blank_file ← CreateFile()
25:              new_img ← ContentsRepositoryAccess(content_loc)
26:              if blank_file = null then
27:                  if new_img is valid then img ← OverallCapture(new_img) end if
28:              else old_img_size ← SizeofMemory(old_img)
29:                  new_img_size ← SizeofMemory(new_img)
30:                      if old_img_size ≠ new_img_size then img ← UpdateCapture(new_img) end if
31:              end if
32:          end if
33:    end if
34:  end if
35: // Large-scale Contents Accelerator Mechanism
36:      device_character ← CheckDevice(Device_attr)
37:                          /*Device_attr ← Device Type, DeviceTypeofBrowser */
38:      img_kind ← SearchAccleratorInfo(device_character)
```

```
39:     switch device_character.type
40:         case MOBILE:   // SmartPhone Device
41:             mobile_kind ← device_character.model            /* SKT|LGT|KTF */
42:             device_kind ← device_character.device_browser    /* aulnatem|melteIson|....|polaris */
43:             if mobile_kind is valid and device_kind is valid then
44:                 MPEG4_img ← MobileEncording(img_kind, mobile_kind, device_kind)
45:                 SendToMobile('HTTP_MOBILE_RESPONSE', MPEG4_img)
46:             else SendMsg(err, 'unknown Device')
47:             end if
48:         case STB:   // STB Device
49:             stb_kind ← device_character.model
50:             if stb_kind is valid then
51:                 MPEG4_img ← STBEncoding(img_kind, device_character)
52:                 SendToSTB('HTTP_STB_RESPONS', MPEG4_img)
53:             else SendMsg(err, 'unknown Device')
54:             end if
56:         case PC:   // Web Device
57:             pc_kind ← device_character.model
58:             if pc_character is valid then
59:                 MPEG4_img ← PCEncoding(img_kind, device_character)
60:                 SendPC('HTTP_RESPONSE', MPEG4_img)
61:             else SendMsg(err, 'unknown Device')
62:             end if
63:     end switch
64: end if
```

5.2 Implementation

The following case shows the Seoul Metropolitan Facilities Management Corporation application of 'IPTV Traffic Information System' in Seoul. In the case, the real-time traffic images information service is provided through diverse digital devices (N-screen). (Fig. 4) shows the represented CCTV Road Traffic Screen and CCTV DB Table.

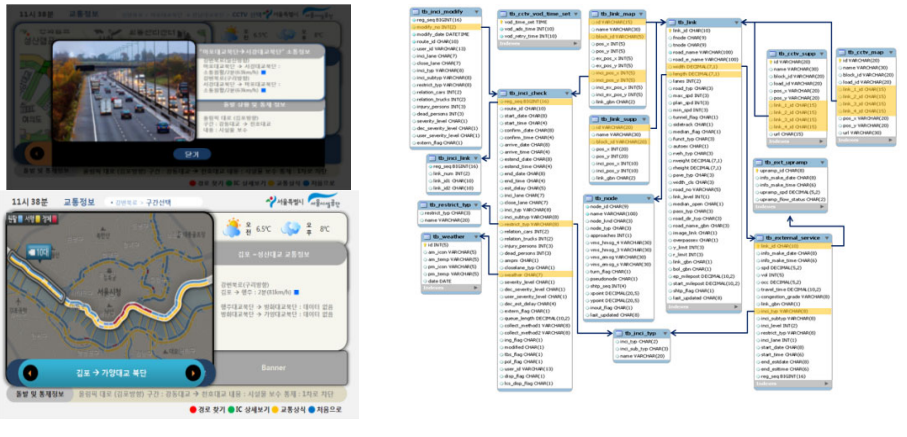


Fig. 4. CCTV Road Traffic Screen & CCTV DB Table

6 Expected Benefits

ICDS can offer following 2 categories functional efficiency.

- Fusible approach for One Contents Multi User (OSMU)

Since the user can maintain the contents which have been already downloaded once from the suggested ICDS and send them to various heterogeneous devices, it is possible to reduce the burden caused by the use of high-priced data.

- Interoperability of Real-time Interactive Contents

The user's heterogeneous devices only handle simple graphics when there is a request for the large-scale contents. Therefore, it is possible to use the real-time interactive contents service rapidly. Also, since such a structure forms the open platform based on the Internet technology without causing any change of the communication infrastructure, it is possible to receive the services which are compatible with every standard.

7 Conclusion and Future Research

This research suggested ICDS (Interactive Contents Delivery System) as an interactive contents delivery method of using image capture processing method. The purpose of the ICDS is to provide a solution for problems in which the files cannot be played due to reasons such as the performance differences between heterogeneous devices, supported data size and the resolution. That is, by suggesting an image processing method to allow the image files to be played in a number of device, more effective interactive services can be provided as a result of this study. The ICDS can be used in an actual project for Seoul and maybe later on could be expanded to a nationwide usage. In the future, the proposed ICDS will be the basis for a system design to be proposed, which is more reliable and safer when the payment is made between devices of different types.

Acknowledgments. This research was supported by the Seoul R&BD Program (JP090988).

References

1. Korea Communications Commission: Smart Mobile Security Plan (2010)
2. Meeker, M., Devitt, S., Wu, L.: Internet Trends. The Mobile Internet Report (2010)
3. World Economic Forum: Scaling Opportunity: Information and Communications Technology for Social Inclusion (2010)
4. Korean Creative Content Agency: Trend of Smartphone Technology and Industry. In depth report of Culture Technology, no. 3 (2010)

Finger Gesture Estimation for Mobile Device User Interface Using a Rear-Facing Camera

Jun-Ho An, Jin-Hong Min, and Kwang-Seok Hong

School of Information and Communication Engineering, Sungkyunkwan University,
300 Cheoncheon-Dong, Jangan-Gu, Suwon, 440-746, South Korea
{amadasv,kaysi}@skku.edu, kshong@skku.ac.kr

Abstract. Mobile user interfaces is an increasingly important topic because of the rapid growth of smart phones and tablet PCs. Thus, we propose a new type of mobile user interface, focusing on mobile devices equipped with a rear-facing camera. In the proposed system, users can interact with mobile software applications using finger gestures with the hand which holds a mobile device, which are detected by a rear-facing camera. Skin-color segmentation, morphological operation, blob detection and skeletonization are used in finger pointer tracking and finger gesture estimation. Mouse events, directional commands and mobile applications can be controlled by the proposed finger gesture-based mobile user interface.

Keywords: Mobile User Interface, Human Computer Interaction, Fingertip Tracking, Gesture Recognition.

1 Introduction

The rapid growth of mobile devices such as a smart phones and tablet PCs has introduced multimedia content and web services to people's daily lives. Moreover, users want to increase their user interaction efficiency when using mobile devices and applications. To support this demand, the importance of user interfaces (UI) has increased. Mobile user interfaces represent the interaction between a user and a mobile device. A user's ability to control and act efficiently are intimately related to the user interface.

There has been research on new types of input interface systems. Accordingly, the latest mobile devices can support a variety of UIs from keypads to multitouch screens and sensors. For example, accelerometers and gyro sensors could be added to mobile devices to support new types of UIs. However, as the number of sensors for Uis increase, they will become difficult to integrate into existing small form-factor mobile devices at the hardware level [1]. Therefore, vision-based mobile user interfaces can serve as an important way of using camera-equipped mobile devices since no new hardware must be introduced.

Vision-based gesture recognition is a significant research area in the field of Human Computer Interaction (HCI). A gesture-based UI has an advantage in that the direct use of a user's hands and fingers is a natural means for humans to communicate

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 230–237, 2011.

with devices [2]. Furthermore, users can interact more intuitively with mobile devices than traditional UIs. In other papers, the use of glove-based devices or two-handed interaction is extensively researched [3]. However, even though these methods achieve high accuracy and speed in measuring hand gestures, these are not suitable for mobile devices because of restricted hand motions.

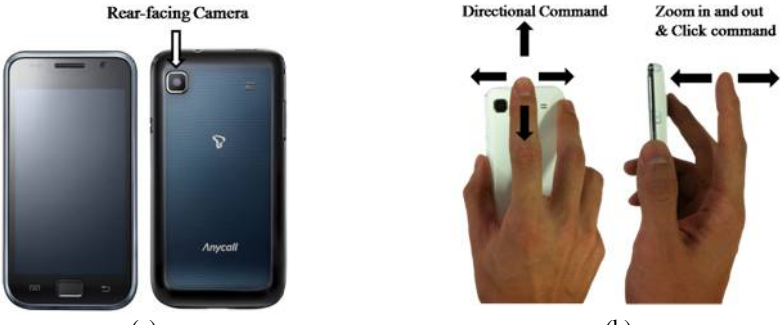


Fig. 1. Proposed mobile user interface :
(a) Mobile device and rear-facing camera (b) Basic control principle

This paper presents a mobile user interface based on finger gesture estimation with a rear-facing camera. In the proposed system, users can execute a required task on mobile devices with finger gestures from the fingers of the hand that is holding the device. The system estimates the movement and gestures of the finger to operate mobile applications using the finger gesture images through the rear-facing camera. The basic idea of the proposed system relies on the finger gestures recognized from captured images of the mobile device's camera as depicted in Fig. 1(a) and the basic control principle in illustrated in Fig. 1(b).

2 Fingertip Estimation

Fig. 2 shows the implementation details of the fingertip estimation system that can accurately estimate a fingertip from a single image captured by a rear-facing camera. The system consists of a pre-processing module which uses skin-color segmentation, morphological operations, blob detection and skeletonization. The estimated coordinates and area of the fingertip is then further refined and fed into the finger gesture recognition module which determines the commands of mobile applications.

In computer vision research, color information is important and widely used in image processing and analysis. There are a variety of color spaces, for example RGB, normalized RGB, LUV, LAB, XYZ, YUV for color coding, YIQ, HSV, HIS, and GLHS from computer graphics [4]. The most common color space is RGB. However, it is not suitable for constructing an accurate skin color model since a high correlation exists between the three components, R, G and B. Thus we use YCbCr space for skin color segmentation [5]. There are non-linear relations between chrominance components (Cb, Cr) and a luminance component (Y). The chrominance components

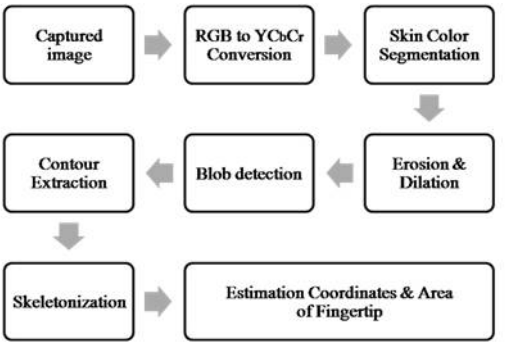
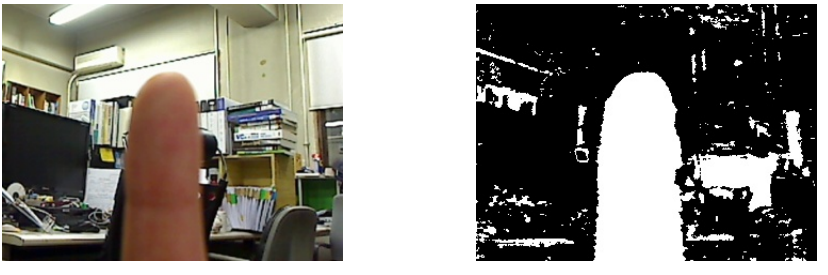


Fig. 2. Proposed fingertip estimation system

are almost independent of the luminance component in the space. Hence, lots of skin color researches operate only on the Cb and Cr plane. Each pixel is classified as either skin or non-skin and they are converted into a new binary image as shown in Fig. 3 with our threshold value analysis, which is defined as the following equation:

$$SkinColor(x,y) = \begin{cases} 1 & if\ (77 \le C_b \le 127) \bigcap (133 \le C_r \le 178) \\ 0 & otherwise \end{cases} \quad (1)$$



(a) (b)

Fig. 3. (a) Original Image (b) Skin Color Segmentation

By conducting skin color segmentation, a binary image is obtained with the skin region masked in white and the non-skin regions masked in black. The two basic morphological operations [6], which are erosion and dilation, are used in the system. Both operations are applied with a 3 by 3 structuring element to remove small background objects and to separate the finger blob. However, non-skin colored objects, which are larger than the 3 by 3 structuring element, still exist. To remove large objects except the finger, we labeled each blob. In the blob detection, connected component labeling is used to divide and label blobs [7].

According to the peculiarity of the proposed system, we keep the blob which contains the lowest row of the image and eliminate other blobs. The contour of the blob is extracted by the edge linking algorithm, which links edge pixels if they are

within a small neighborhood and have similar magnitude or direction [8]. Using the extracted contour, a skeletonization is performed. The aim of the skeletonization is to extract a region-based shape feature representing the general form of an object [9]. Generally, in most existing skeletonization algorithms, when a pixel is determined to be a boundary pixel, it is either deleted directly from the image or flagged and not deleted until the entire image has been scanned. However, the common skeletonization algorithms require a large computation time and are not suitable for mobile devices. Thus, in the proposed system, the method finding the center pixels of the finger blob can be substituted for a skeletonization algorithm as in Fig. 4. As a result, the coordinates and area of the fingertip pointer with the thin line are applied to the finger gesture command recognition.

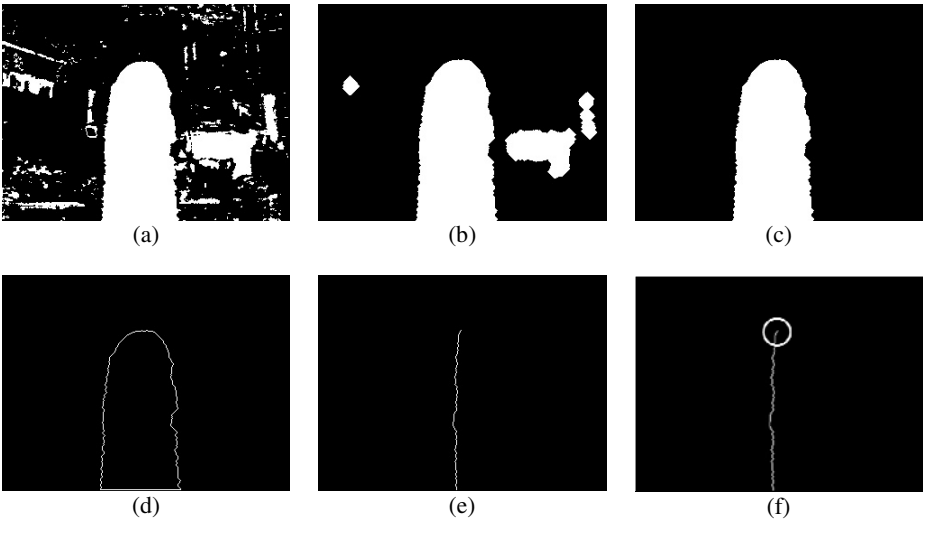


Fig. 4. (a) A binary image (b) Erosion and dilation (c) Blob detection (d) Contour Extraction (e) Skeletonization (f) Extracted finger pointer

3 Fingertip Gesture Commands

In the proposed system, the general directional commands and mouse events can be operated as the same function as other interfaces. First, the directional commands (up, down, left and right) are defined by chessboard distance using the coordinates and movement of the finger pointer. The moved distance, direction and the instant speed of the finger pointer between two frames determine the directional commands.

If the size of the input image I is $N \times M$ with $I[i, j] \in 0 \le i \le N-1, 0 \le j \le M-1$, the current pixel position of the finger pointer locates at (i_2, j_2) and the previous pixel position located at (i_1, j_1), the chess board distance can be defined as:

$$d_{chess} = \max(|i_2 - i_1|, |j_2 - j_1|) \tag{2}$$

Each of the directional commands is shown in Fig. 5. The variable σ can be varied according to the frame rate of the system properly. In the proposed system $\sigma = 4$ is selected. Each of the directional commands is shown in Fig. 5. The variable σ can be selected by the frame rate of the system properly.

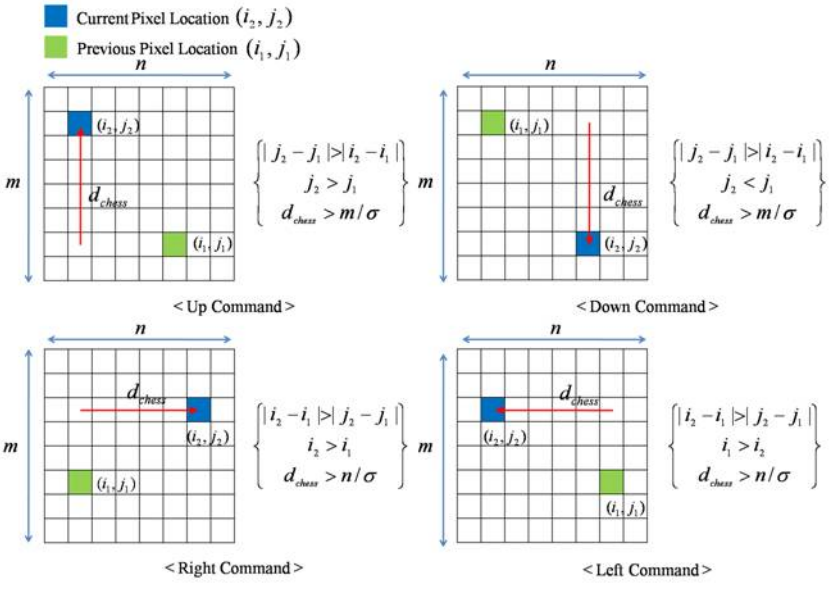


Fig. 5. Directional commands by finger gesture recognition

Commands for clicks, zooming in, and zooming out are generated by the area of the fingertip. The instantaneous rate of change of the fingertip area is applied to the click command by finger gestures, and is defined as the fingertip area of the current frame over the fingertip area of the previous frame. When a user moves his finger back and forth toward the camera similar to the gesture of clicking a mouse, a click command is triggered. The threshold value is determined by the experimental result as in Fig. 6.

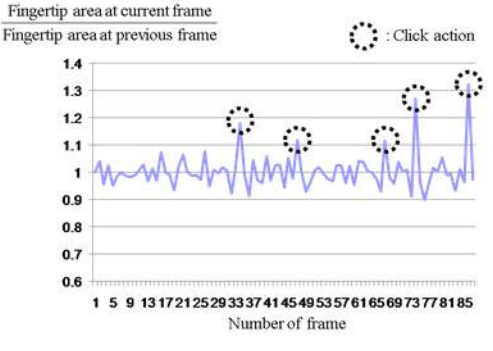
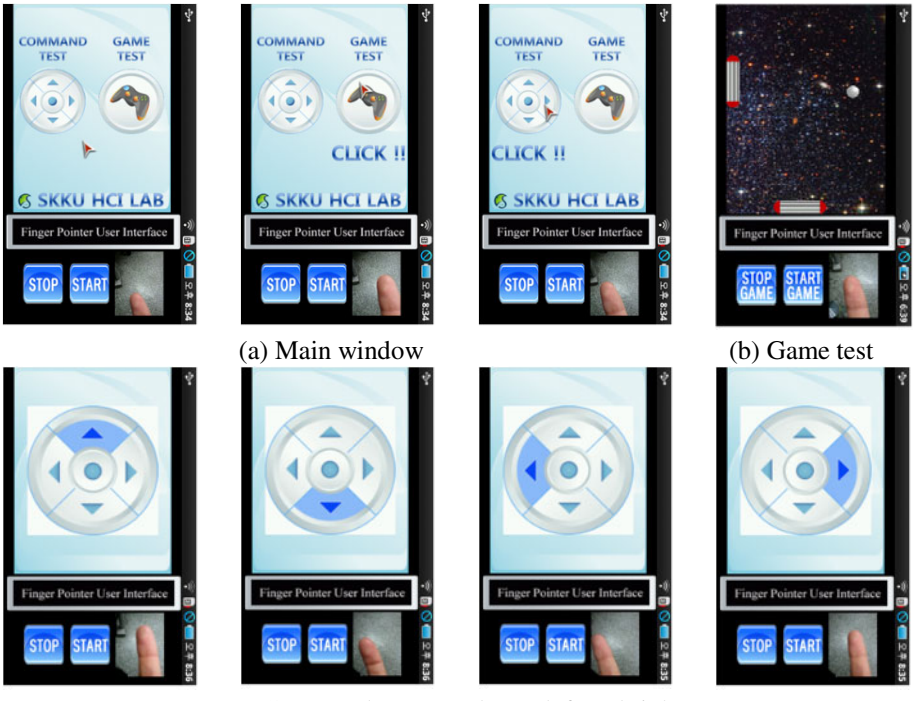


Fig. 6. Example of a click command

4 System Configuration and Experiment

The proposed system is implemented on the Samsung Galaxy S smart phone, which is provided with a 1GHz processor, 512 MB of memory, and a 5MP rear-facing camera. The operating system is Android with Java support. Since Android applications based on Java are slower than applications written in native C/C++ languages, we use the Android NDK and the JNI interface [10][11]. The performance of the Android application can be improved by the Android NDK because Android applications written in Java are only slower than the applications using a native C library. Furthermore, our system frequently requires memory access which needs complex calculations due to real time image processing. Thus, the code related to fingertip estimation and finger gesture command recognition algorithms was written in native C and other code related to the GUI was written in Java with the Android SDK.

Fig. 7 shows the developed system which contains the main window with cursor movement, directional command windows and game applications. A cursor moves corresponding to finger movements and a user can select a game test program and a command test program using the defined finger click action. In the game test, a user can move a small paddle on the left and bottom easily and hit a ball using finger movement.



(a) Main window (b) Game test

(c) Command test (up, down, left and right)

Fig. 7. Finger gesture command test

The accuracy assessment of the fingertip extraction algorithm was performed by the developed program. We used the camera at a frame rate of 10fps and a resolution of 320 x 240 pixels. Five participants in the fingertip extraction experiment moved the cursor using the developed program for one minute. About 3000 images, 600 images per person, were used in accuracy assessment. The overall accuracy was 98.2% under uniform lighting conditions. The experiment of finger gesture commands was performed. Each finger gesture input command, click, up, down, left, and right, was tested 50 times. The overall recognition rate was 92% as shown in Table 2.

Table 1. Fingertip extraction

Participant number	The number of fingertip extraction	Extraction rate
1	592/600	98.6%
2	585/600	97.5%
3	594/600	99%
4	579/600	96.5%
5	598/600	99.6%
	2948/3000	98.2%

Table 2. Finger gesture command recognition

Commands	The number of fingertip gesture recognition	Recognition rate
Click	41/50	81%
Up	43/50	86%
Down	49/50	98%
Left	49/50	98%
Right	48/50	96%
	230/250	92%

5 Conclusion

A mobile user interface based on gesture estimation using a rear-facing camera was proposed in this paper. The reason why we focused on a mobile UI system using a rear-facing camera is that cameras are now widely available on mobile devices and can be used without introducing new sensors. Some of the previous studies on mobile UI systems used a front-facing camera [12]. However, those systems could make mobile users feel inconvenient due to two-handed usage. On the other hand, the proposed system can support a user interface using only one finger with the hand holding a mobile device. For future work, the method of a more natural finger gesture command, robust skin-color segmentation and fingertip extraction to light and complex backgrounds should be studied.

Acknowledgments. This research was supported by MKE, Korea under ITRC NIPA-2011-(C1090-1121-0008) (NTIS-2010-(1415109527)) and PRCP through NRF of Korea, funded by MEST (2010-0020210).

References

1. Koh, E., Won, J., Bae, C.: On-premesie skin color modeling method for vision-based hand tracking. In: IEEE 13th International Symposium Consumer Electronics, ISCE 2009, May 25-28 (1981)
2. Wu, Y., Huang, T.S.: Vision-based gesture recognition: A review. In: Braffort, A., Gibet, S., Teil, D., Gherbi, R., Richardson, J. (eds.) GW 1999. LNCS (LNAI), vol. 1739, p. 103. Springer, Heidelberg (2000)
3. Sasaki, H., Kuroda, T., Manabe, Y., Chihara, K.: Augmented reality based input interface for wearable computers. In: Heudin, J.-C. (ed.) VW 2000. LNCS (LNAI), vol. 1834, pp. 294–302. Springer, Heidelberg (2000)
4. Wu, Y.-W., Ai, X.-Y.: Face Detection in Color Images Using AdaBoost Algorithm Based on Skin color Information. In: Knowledge Discovery and Data Mining 2008, pp. 339–342 (2008)
5. Vezhnevets, V., Sazonov, V., Andreeva, A.: A survey on pixel based skin color detection techniques. In: GraphiCon, Moscow, Russia (September 2003)
6. Shapiro, L., Stockman, G.: Computer Vision, pp. 63–73. Prentice Hall, Englewood Cliffs (2002)
7. Dillecourt, M.B., Samet, H., Tamminen, M.: A general approach to connected-component labeling for arbitrary image representations. Journal of the ACM 39(2), 253–280 (1992)
8. Umbaugh, S.E.: Computer Imaging: Digital Image Analysis and Processing (2005)
9. Abeysinghe, S., Baker, M., Chiu, W., Ju, T.: Segmentation-free skeletonization of grayscale volumes for shape understanding. In: SMI, pp. 63–71 (June 2008)
10. Android NDK | Android Developers, `http://developer.android.com/sdk/ndk/index.html`
11. Java Native Interface, Wikipedia, `http://en.wikipedia.org/Java_Native_Interface`
12. Hirobe, Y., Niikura, T., Watanabe, Y., Komuro, T., Ishikawa, M.: Vision-based Input Interface for Mobile Devices with High speed Fingertip Tracking. In: 22nd ACM Symposium on UIST, pp. 7–8 (2009)

An Adaptive Parameter Estimation Method for Wireless Localization Using RSSI Measurements

Jinhyung Park[1], Hyunhun Cho[1], Seunghae Kim[1], Dowoo Park[2], Ahsol Kim[2], and Joongoo Park[2]

[1] Dept. of Computing & Networking Resources,
Korea Institute of Science and Technology Information,
335 Gwahangno, Yuseong, Daejeon, Korea
{ntoskr,hhcho,shkim}@kisti.re.kr
[2] Graduate School of Electrical Engineering and Computer Science,
Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu, Korea
{dwpark,ahsol,jgpark}@knu.ac.kr

Abstract. In location-based service (LBS), it is important to determine the location of a mobile user more, exactly and efficiently. In case of a wireless network, such as IEEE 802.11, a method using a received signal strength indicator (RSSI) for each access point is advantageous because it does not require additional infrastructure and it is suitable for the simplification. However, this technique of localization is required to investigate the propagation environment to determine the location of mobile user in advance. We propose a method in which the parameters of the propagation environment are determined using only RSSI measurements obtained during localization.

Keywords: signal strength, WLAN, RSSI, log model, distance function.

1 Introduction

Due to the popularization of the internet and the development of mobile communication technology, users want to obtain information and service anywhere and at any time. Location-based service (LBS) satisfies these needs and provides each user with differentiated service. The basis of LBS is to determine the location of each mobile unit. It is easier to predict the location of a mobile user using a received signal strength indicator RSSI based on Wireless-LAN indoors [2]. However, when using the location estimation method based on a RSSI, it is necessary to investigate the indoor propagation environments prior to localization [2]. If this method is adopted in a large space, this requirement of prior investigation is likely to become obstruction hindrance to localization because of the large amount of labor needed. To make up for weakness of the RSSI method, we propose a method that does not require prior investigation.

In this study, the proposed method was evaluated experimentally, and we demonstrated that the localization can be successful even when the preliminary parameter measurements are omitted.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 238–244, 2011.

2 Attenuation Log Model for Ranging

The attenuation log model was used in this study. The attenuation log model assumes that the ideal propagation condition is a line-of-sight path between the transmitter and the receiver. This model does not consider any effect of multipath fading and other path loss. In an indoor space, this model is not precise. However, if there are many sensor nodes in the space, it is still appropriate to determine the distance.

Here, it is assumed that the received signal power P follows a lognormal distribution; thus, the random variable P $[dBm]$ = $10log_{10}$ P is Gaussian. In this model, the relationship between RSSI and distance is described as

$$\begin{aligned} &P[dBm] \sim N(\overline{P}[dBm], \sigma^2) \\ &\overline{P}[dBm] = P_0[dBm] - 20\log_{10}(\frac{4\pi d}{\lambda}) \end{aligned} \tag{1}$$

$$\begin{aligned} &\overline{P}[dBm] = \textit{the mean power in decibel milliwatts} \\ &\sigma^2 = \textit{the variance of the shadowing} \\ \textit{where,}\quad &P_0[dBm] = \textit{offset} \\ &\lambda = \frac{c}{f} = \frac{3\times10^8[m/s]}{2.4[GHz]} \end{aligned}$$

RSSI is a method of measuring distance between nodes using the Received Signal Strength Indicator(RSSI). In general, the RSSI attenuates in proportion to the distance between nodes. If the offset P_0 is known when the node measures the RSSI, the distance between nodes can be determined.

Now it is assumed that there are three fixed nodes that the positions are prior known as shown in Fig.1, and names node *a*, node *b*, and node *c*. When the signal is sent from the fixed each node to the target node unknown the position, the target node measures the received signal strength from each node, respectively. When the received signal strength obtained from each fixed node is named P_a, P_b, and P_c respectively, the value of them is converted into the distance between target node and each fixed node, and determined as d_a, d_b, and d_c. To estimate the position of the target node, we draw three circles centered on the each fixed node. Their radiuses of the circles are distances between each fixed node and the target node determined ahead. Next, we determine the position of the intersection, the position of the target node can be estimated by this intersection.

In this method, however, to estimate the position of the target node, it is necessary to known the value of offset P_0 prior, and unfortunately, this value has the feature of taking a different value in each measured environment.

The following is an example to explain the relationship between offset P_0 in different environments and distances, where the P_0 of the two measured environments are P_{0_1}, P_{0_2} respectively as in Fig. 1. Even if the values of the RSSI obtained by a target node are same, the estimated distances between the target node and the fixed node become different value according to environment. Therefore, we'll have to use suitable value of the offset P_0 in each environment to estimate the accurate position of the target node.

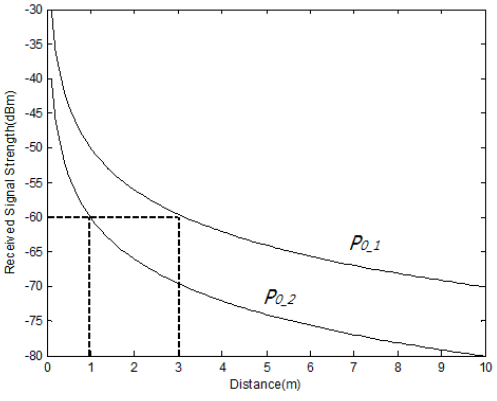


Fig. 1. Ranging on offset P_0

In the past, to know the suitable value of the offset P_0 in each measured environment, prior measurement was conducted and obtained a large amount of sample data was averaged, and the propagation characteristic was determined. However, to conduct this prior measurement on the measured environment of all is difficult because of taking time and troublesome. In the next section, as a solution of the problem a proposal method the can be determined the value of the offset P_0 without investigating of the measured environment prior is explained.

3 Adaptive Parameter Estimation for Ranging

In case of localization procedure using RSSI measurements, the position information of mobile unit can be estimated or calculated without prior information. In this section, we propose an enhanced ranging algorithm with optimized parameters adapting to the radio propagation environments. This algorithm can be implemented with low computation burden.

3.1 The Estimation Model of the Offset Constant

First, the N Access Point (AP) nodes are considered in the testbed. A Mobile Unit (MU) at position θ (x, y, z) receives a signal from the APs and measures the received signal strength indicator. The RSSI measurement obtained by AP_i is given by,

$$P_{ri}[dBm] = P_0[dBm] - 20\log_{10}(\frac{4\pi d_i}{\lambda}) + X_\sigma$$
$$d_i = \sqrt{(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2} \tag{2}$$

where, X_σ is the Gaussian noise with a zero mean and a variance of σ^2, and d_i is the distance between AP_i and MU.

Here, N values of (1) were added to the both sides of (2). Then, the offset P_0 is calculated as follows:

$$P_0 = \frac{20(N \log_{10} \frac{4\pi}{\lambda} + D) + \sum_{i=1}^{N} P_{ri}}{N} [dBm]$$
$$D(x, y, z) = \sum_{i=1}^{N} \log_{10} d_i \tag{3}$$

$D(x, y, z)$ is a distance function between APs represents the summation of the logarithm of the distance from each AP to MU [1]. The received power P_{ri} of AP_i is a value measured during localization, and N is the total number of APs. Therefore, the offset P_0 is inferred from the distance function between APs, $D(x, y, z)$. However, during localization, the coordinates of MU are unknown, and thus, $D(x, y, z)$ cannot be determined. Consequently, the value of the distance function is approximate.

3.2 Approximation of the Distance Function $D(x, y, z)$

The approximation of the distance function between APs consists of three steps as follows:

Step 1: Set the height of the MU.

The coordinates of MU cannot be determined; however, the height of MU and each AP can be set relative to the height of the ground. As a result, the variable z in $D(x, y, z)$ is set, and the function becomes $D(x, y)$. This function is the distance function of the testbed, as shown in Fig. 2.

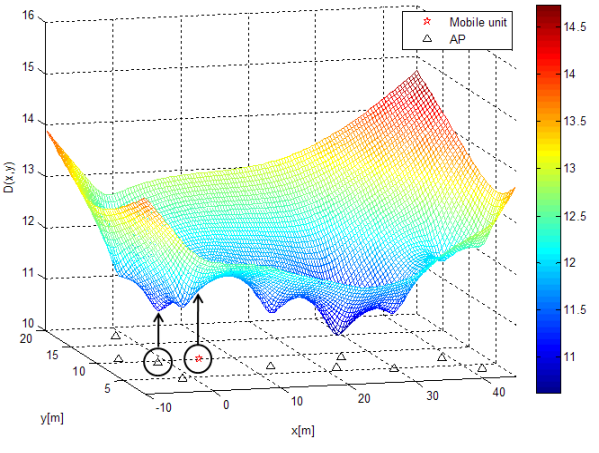


Fig. 2. Distribution of $D(x, y)$ on the testbed

Step 2: Search for the AP that received the maximum power

AP_{imax}, the AP for which the RSSI measurement obtained during localization is the highest, is identified. In general, when the distance between nodes is shorter, the noise of the measurements is smaller. It is inferred that the real position of the MU is near AP_{imax}.

Step 3: Approximate the coordinates (x_{imax}, y_{imax})

In contrast to the identification of the position of the AP_{imax}, which requires two steps, the identification of the real location of the MU requires a variable number of steps. However, it is understood that there is not much difference in each value of the distance function D. In addition, the value of D does not change drastically near AP_{imax}, even if the MU is placed more than 10m from the AP. Therefore, $D(x, y)$ can be approximated by $D(x_{imax}, y_{imax})$, which was obtained in the identification of AP_{imax}.

According to the above-mentioned steps, the unknown function D can be approximated, and the values of the offset P_0 can be determined using (3). However, the measurements P_i [dBm] of the received power given by (2) are influenced by disturbances, for example, multipath fading, reflection and scattering. Moreover, because of the approximation of the distance function, a different result of the calculation is obtained every time when communicating for localization. Therefore, we updated the estimated value in the localization..

3.3 Updating Offset

When the P_0 value obtained from the measurements from the N-th trial is assumed to be α'_n, the actual P_0 used for localization is given by,

$$\alpha_n \leftarrow r_0\alpha'_n + r_1\alpha'_{n-1} + r_2\alpha'_{n-2} + \cdots + r_m\alpha'_{n-m} \tag{4}$$

where r_k ($k = 0, 1, \ldots, m$) are weighting functions, and the summation is set to be *1* in this iterative approximation. In this way, by using not only the data from the latest m th trial, the influence of the multipath fading or shadowing can be inhibited even if the last signal to be received was degraded in this way [1].

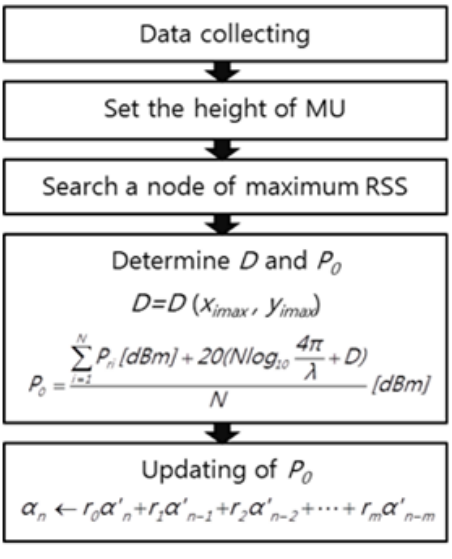


Fig. 3. Flowchart for P_0 estimation

Furthermore, in the measured environment, in which the propagation characteristic changes every time, it is possible to account for the change in the environment adequately by erasing the past data and using some values from the latest data. In this paper, the weighting function $r_k(k=0,1,...,m)$ are $1/m$ and $m=3$. A flowchart explaining the procedure from the data collection to the estimation of P_0 is shown in Fig. 3.

4 Experimental Results

Fig 4 shows our experiment testbed of 50 m × 20 m. 10 APs were placed in this testbed at a height of 1.2 m. The MU moved, generating 12 test points that were localized using the RSSI. Localization was performed in the same way 300 times. The DARL algorithm was used in the localization [2]. The results comparing each case are shown in Fig. 5.

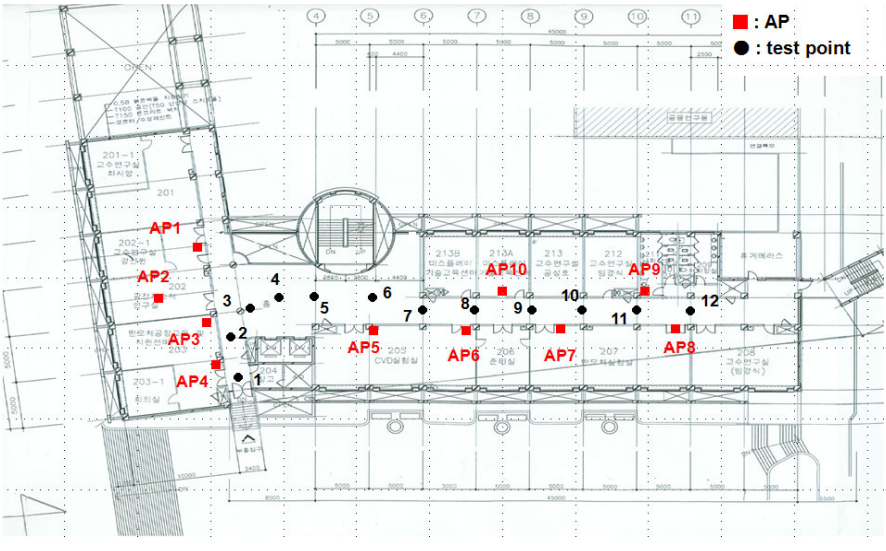


Fig. 4. Map of experimental testbed

Table 1. Specification of the arrangement of APs

Sensor Field	50 m × 20 m
Number of APs	10
z-coordinate of AP(h_{AP})	1.2 m
z-coordinate of MU(h_{MU})	1.4m

As shown in Fig. 5, mean distance error of total test point is reduced about 2 m, comparing method by prior P_0 measurement. Especially, distance errors of test point number 4 and 5 are about 4 m reduction. Consequently, it is thought that distance information from RSSI is improved by recalibration P_0 during localization in poor propagation environment.

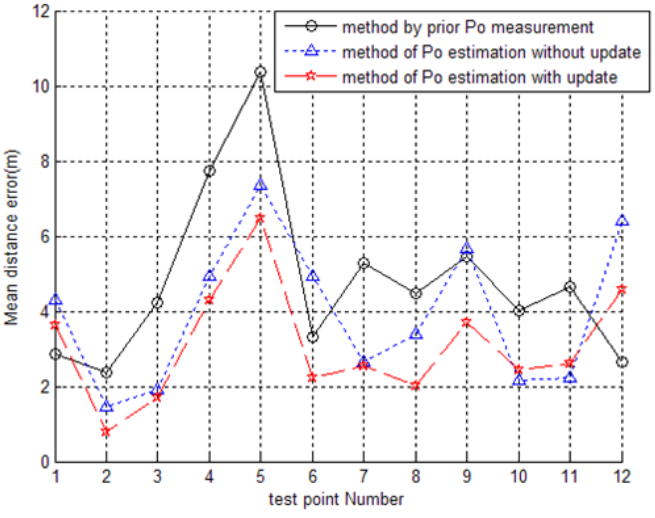


Fig. 5. The mean distance error of each case by experiment

5 Conclusions

We developed a method to estimate the coordinates of MU using parameter estimation of the propagation environment without investigating these parameters prior to localization. The mean distance errors from three methods of localization were evaluated. The results confirmed that the mean distance error of the proposed method was approximately 7% less than that of the P_0 estimation method without updating and 36% less than that of the method including prior P_0 measurement. Therefore, proposed algorithm is an effective way to estimate the location of the MU.

References

1. Tateishi, K., Ikegami, T.: Estimation Method of Attenuation Offset During Localization in RSSI. IS Communication and Information Technologies, 482–487 (2008)
2. Lim, C.B., Kang, S.H., Cho, H.H., Park, S.W., Park, J.G.: An Enhanced Indoor Localization Algorithm Based on IEEE 802.11 WLAN Using RSSI and Multiple Parameters. In: IEEE ICSNC, p. 238 (2010)
3. Tateishi, K., Sugiyama, Y., Watanabe, K., Ikegami, T.: An examination no Location Estimation to Consider Statistics of RSSI. In: Proc. ICICE General Conf., A-21-3, p. 394 (2007)
4. Malekpour, A., Ling, T.C., Lim, W.C.: Location Determination Using Radio Frequency RSSI and Deterministic Algorithm. In: IEEE CNSRC (2008)
5. Patwari, N., O'Dea, R.J., Wang, Y.: Relative location in wireless networks. In: Proc. IEEE VTC, vol. 2, pp. 1149–1153 (2001)
6. Saha, S., Chaudhuri, K., Sanghi, D., Bhagwat, P.: Location determination of a mobile device using IEEE 802.11b access point signals. In: Proc. IEEE Wireless Communications and Networking Conf., pp. 1987–1992 (2003)
7. Nafarieh, A., Ilow, J.: A Testbed for Localizing Wireless LAN Devices Using Received Signal Strength. In: IEEE CNSRC (2008)

Sidelobe Suppression Methods of Cubic-Phase Linear Chirp in Two-Ray Multipath Model for UWB Ranging

Jinhyung Park[1], Hyunhun Cho[1], Seunghae Kim[1],
Changbeom Lim[2], and Joongoo Park[2]

[1] Dept.of Computing & Networking Resources,
Korea Institute of Science and Technology Information,
335 Gwahangno, Yuseong, Daejeon, Korea
[2] Graduate School of Electrical Engineering and Computer Science, Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, Korea
{ntoskr,hhcho,shkim}@kisti.re.kr, {cblim,jgpark}@knu.ac.kr

Abstract. The UWB communication system occupies a very large bandwidth, 7.5 GHz. This enables high precision ranging that provides correct position information. In this paper, we will present a method to reduce side-lobes in receiver. Reducing side-lobes can enhance the accuracy of ranging system. Ranging system with high ratio of main-lobe to 1st side-lobe produces lower calculation error in ranging. Proposed cubic-phase chirp signal method can reduce the power of 1st side-lobe, that is, realize higher accurate ranging.

Keywords: Chirp, UWB, LBS, Ranging, Positioning.

1 Introduction

Recently the wireless communication technologies are widely applied to LBS (Location Based Service). The UWB system has high potential in ranging applications with the advantage of extremely wide bandwidth, 7.5 GHz (from 3.1 to 10.6 GHz) [1]. UWB signal bandwidth should be more than 500 MHz or the fractional bandwidth should be larger than 0.2 with the PSD (Power Spectral Density) [2]. The chirp signal has a few desirable advantages such as low-power consumption and low peak-to-average power ratio (PAPR) compared to the conventional Gaussian pulses [3]. The Gaussian pulses have very high PAPR so that the signal generation in hardware processing is difficult but relatively the chirp signal can be easily generated by the surface acoustic wave (SAW) filter [4].

In this paper, we introduce the cubic-phase linear chirp signal. It has high ratio of main-lobe to the 1st side-lobe ratio in the receiver. And it realizes the high accuracy in range system. But higher ratio of main-lobe- to the 1st side-lobe is needed in the multipath environment. So side-lobe suppression methods with chirp signals will be presented. The chirp signals ranging is performed in the simplified multipath model, Two-ray multipath Model. Computer simulations will evaluate for ranging capability of the chirp signals in several environment. Finally research summary and concluding remarks will be stated.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 245–255, 2011.

2 Linear Chirp

2.1 Chirp Theory

A chirp signal can be represented by the band-pass signal.

$$C_n(t) = \alpha(t)\cos(2\pi f_c t + \pi u t^2 + \phi_0),\ |t| < \frac{T_s}{2} \tag{1}$$

where $\alpha(t)$ is the envelope of the chirp signal, f_c is the carrier frequency, u is the chirp rate, Ts is the chirp duration and φ_0 is the frequency offset. The chirp bandwidth is defined as B=|u| Ts and the time-bandwidth product is TB. A chirp with positive u is an up-chirp, otherwise a down-chirp. Equation (1) implies that the frequency also varies as well as the time and the phase is quadratic to the time.

2.2 Cubic-Phase Chirp Signal

The cubic-phase chirp signal [6] is introduced in formula as below.

$$C_n(t) = \alpha(t)\cos(2\pi f_c t + 0.67\frac{\pi u t^3}{T_s} + \phi_0),\ |t| < \frac{T_s}{2} \tag{2}$$

It was originated by the inspiration with the increase rate of the natural numbers to the exponent as from 1, 4, 9, 16, . . to 1, 8, 27, 64, . . In particular the parameter Ts has an importance to make the state of cubic-phase linear (CPL) chirp analyzable. The significant difference from the quadratic-phase chirp is the increase of the power of t so that the increase rate of t gets faster and it leads to the better correlation ratio between main-lobe and the 1st side-lobe. Though the instantaneous frequency becomes quadratic for which it is the derivative of the phase φ_0, it maintains the almost linearity within chirp duration Ts but outside it changes quadratically. And the coefficient 0.67 is multiplied to have same bandwidth parameter condition with quadratic-phase.

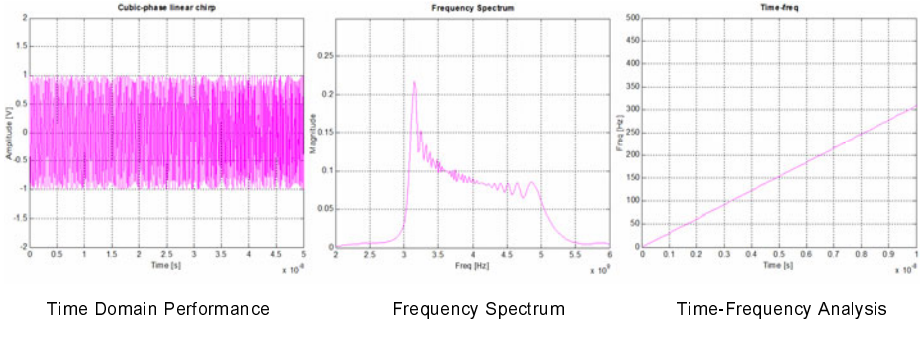


Fig. 1. Cubic-Phase Linear Chirp

Fig 1 shows simulations of the cubic-phase chirp. The time-domain performance is almost same with the quadratic-phase case but the frequency sweeping rate is a little faster. The frequency spectrum became unbalanced at the beginning for up-chirp. The time-frequency analysis reveals the linearity within time duration Ts.

The pulse compression process is same as the case of quadratic-phase chirp but the output equation of matched filter becomes different since the phase degree has been increased. The impulse response of matched filter is expressed.

$$h_m(t) = C^*_{CPL}(-t) = \alpha(t)\cos(2\pi f_c t - 0.67\frac{\pi a t^2}{T_s}),\ |t| < \frac{T_s}{2} \tag{3}$$

And the pulse compression is performed by passing CPL chirp through matched filter. But it could not be neatly completed as *sinc* function.

$$\phi_m(t) = h_m(t) * C_{CPL}(t) = \int_{-\infty}^{+\infty} h_m(\tau)\cdot C_{CPL}(t-\tau)d\tau \tag{4}$$

The correlation performance has an significant relation with ranging accuracy. Higher main-lobe and lower side-lobe levels are well-correlated. And one more important factor for the correlated signal is the ratio of main-lobe to the 1st side-lobe. If the ratio is small that the 1st side-lobe approaches to the main-lobe, it causes wrong peak detection. Especially when signal is propagated through the noisy channel including receiver noise, it gets worse. And the chirp signal whose the ratio of main-lobe to the 1st side-lobe ratio is lower can perform higher-precision ranging.

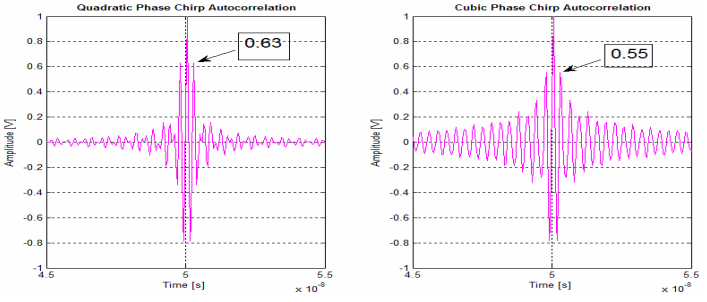


Fig. 2. Auto-Correlation Performance: Quadratic and Cubic-Phase Chirp

Fig 2 provides the auto-correlations of quadratic-phase and CPL chirps. The CPL chirp has lower 1st side-lobe level compared to the quadratic-phase. But it has increased side-lobe ripples. The ratio of main-lobe to 1st side-lobe at 2 GHz bandwidth is 1.803 for CPL chirp and 1.583 for quadratic-phase chirp. And the overall side-lobe levels are analyzed by correlation variance.

3 Side-Lobe Suppression

In the previous section, the cubic-phase chirp signal shows higher ratio in autocorrelation than quadratic-phase chirp. The high peak of side-lobe causes wrong peak detection in receiver. Especially it is getting worse when the signal is propagated through the noisy channel. The ratio of main-lobe to 1st side-lobe is important in the ranging system. So the cubic-phase chirp signal realizes higher accuracy than quadratic-phase chirp signal in a ranging system.

The cubic-phase chirp signal has high ratio but it needs to suppress its side-lobe. There are two methods to suppress the side-lobe ripples. One is the Internal hamming weighting and another is the Amplitude tapering. The Internal hamming weighting realizes in matched filter and the Tukey window function is added in the transmitted signal.

3.1 Matched Filter Frequency Response

In a practical SAW linear-FM pulse compression system, side-lobes below -40dB are generally not controlled by the weighting function but rather by amplitude and phase error of the delay lines used for chirp generation and compression, and by unavoidable spurious signal. Therefore Hamming is most commonly used in such systems. This function is closely modified truncated Talyor function. Talyor function allows a compromise between side-lobe reduction on one hand and pulse widening and mismatch loss widening of the main-lobe and the mismatch loss inverse with decreasing side-lobe design level.

For the case of using an unweighted DDL and an external frequency domain Hamming weighting filter in the receiver, the resulting compressed pulse *r(t)* can be calculated from

$$r(t) = F^{-1}[C(f)H(f)] \quad :output \;\; signal \tag{5}$$

$$C(f) = F[c(t)] \quad :transmitted \;\; signal \tag{6}$$

$$H(f) = F[c(t)\omega_H(t)^*] \quad :frequency \;\; response \tag{7}$$

$$\omega_H(t) = 0.08 + 0.92\cos^2 \pi \frac{t}{T}, \;\; |t| \le \frac{T}{2} \tag{8}$$

with F and F^{-1} denoting Fourier transform and its inverse respectively. $\omega_H(t)$ is the Hamming weighting function in time domain and $H(f)$ is the matched filter frequency response.

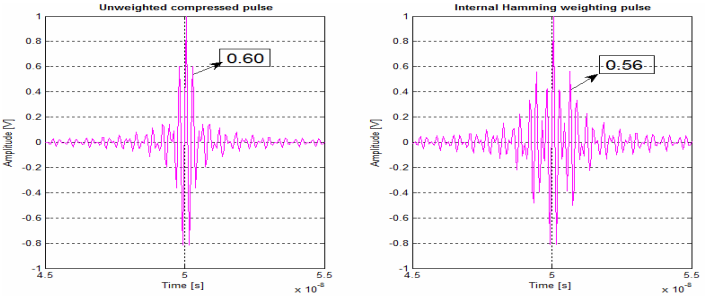


Fig. 3. Autocorrelation in Quadratic-Phase Chirp (TB=100)

Fig. 3 provides the autocorrelation of quadratic-phase chirp using the Hamming weighting in matched filter. It shows that the side-lobe is suppressed when the Hamming weighting function is used.

The Hamming weighting was also adopted in cubic-phase chirp autocorrelation in Fig 4. But the 1st side-lobe lever is lower in cubic-phase than in quadratic-phase chirp autocorrelation. The ratio of main-lobe to 1st side-lobe is 2.38 in the cubic-phase chirp and 1.8 in the quadratic-phase chirp. The cubic-phase chirp shows the lower side-lobe peak level comparing with the quadratic-phase chirp.

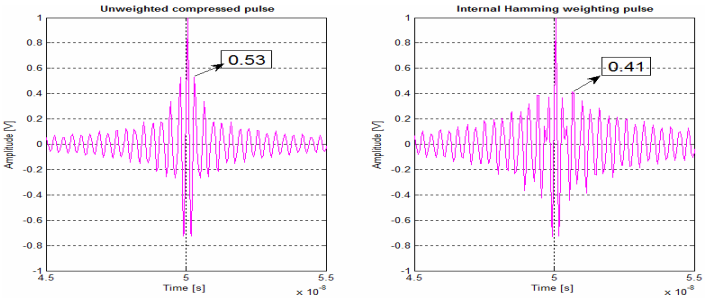


Fig. 4. Autocorrelation in Cubic-Phase Chirp (TB=100)

3.2 Amplitude Tapering

The amplitude ripples of the spectrum can be attributed to the sharp rise-time of the time envelope because a linear FM signal with infinite duration has no ripples. Data correlating pulse rise-time and spectrum ripple were reported previously. Based on this analysis, a modified chirp with amplitude tapering of the transmitted signal has been generated. The attainable ripple reduction is a function of the signal's bandwidth, the choice of the tapering function, and its duration.

The amplitude spectrum can be smoothed by tapering the rectangular chirp envelope. If cosine tapers of length αT are added to the FM pulse, the transmitted signal is described by

$$c(t) = f_T(t)\cos\left\{ j2\pi(f_c + \frac{B}{2T}t^2)\right\},\ |t| \le (0.5+\alpha)T \tag{9}$$

where

$$f_T(t) = \begin{cases} 1 & ,|t| \le T/2 \\ 0.5\left\{1+\cos(\pi\dfrac{|t|-T/2}{\alpha T}\right\}, & T/2 \le |t| \le (0.5+\alpha)T \\ 0 & ,elsewhere \end{cases} \tag{10}$$

is the cosine-tapered window or Tukey window function [5]. For $\alpha=0.1$, there is no ripple observed in the centre of the frequency response. The peak side-lobe level of the correlation pulse is lowered and, more importantly, the far side-lobes fall of rapidly [7].

The amplitude ripples of the spectrum can be attributed to the sharp rise-time of the time envelope because a linear FM signal with infinited duration has no ripples.

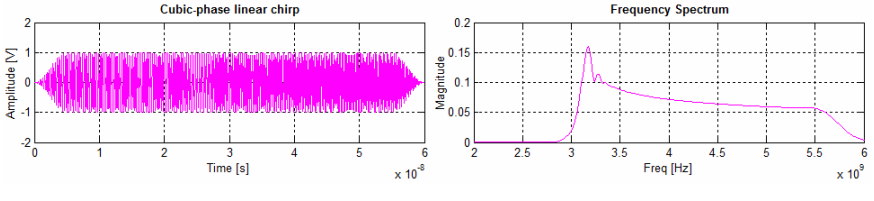


Fig. 5. Cubic-Phase Chirp with Amplitude tapering

Time-bandwidth product TB is 100, T_s time duration is 50ns and Bandwidth is 2GHz. $\alpha T(\alpha=0.1)$ is added in the transmit signal time duration and it is shown in Fig 5. The spectrum amplitude of cubic-phase chirp is tapered.

The amplitude spectrum ripples are reduced with the amplitude tapering method. The ratio of main-lobe to 1st side-lobe in autocorrelation is shown in the Fig 6. The ratio of quadratic-phase chirp in autocorrelation in Fig 4 is same as the ratio in Fig 6. However, the cubic-phase chirp ratio is higher than with only internal hamming weighting used. The ratio of cubic-phase chirp is 2.94 and it of the quadratic-phase chirp is 1.78.

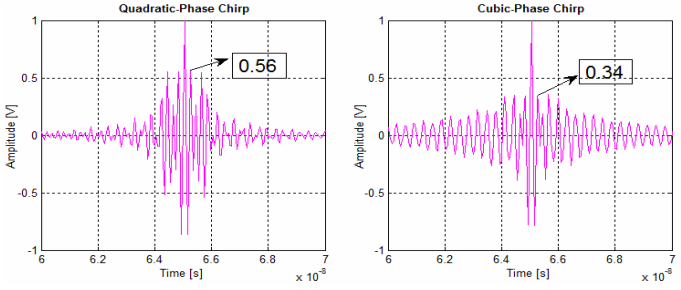


Fig. 6. Autocorrelation Quadratic and Cubic-Phase Chirp

Table 1. Mainlobe-1st.sidelobe Ratio

Side-lobe reduction Method	Ration of Main-lobe to 1st Side-lobe	
	Quadratic-Phase	Cubic-Phase
Unweighted	1.59	1.80
Internal Hamming	1.78	2.40
Cosine-tapering and Hamming	1.78	2.94

Table 1 shows the ratio of main-lobe to 1st side-lobe in autocorrelation. Both the ratio of signal are increased. But the increasing rate of cubic-phase chirp ratio is higher than quadratic-phase.

4 Chirp Signal in Two-Ray Multipath

4.1 Reflection Coefficient

To analyze the multipath problem, the two-ray model is used for simplicity which is depicted in Fig 7. A transmitted signal reaches the receiver through a line of sight path and a reflected path which arrives at the receiver with reduced power and time delay of τ. d is the distance between the transmitter and the receiver, h is the height and θ is the incident angle.

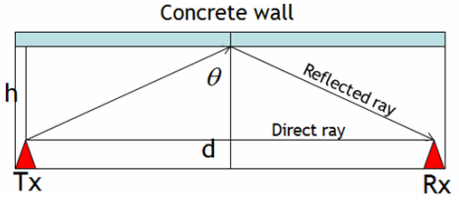


Fig. 7. Two-ray Multipath Model

The transfer function representing the path characteristics may be expressed in impulse response as the equation.

$$h(t) = \delta(t) + r(t)\exp(j\theta\pi/180)\cdot\delta(t-\tau) \tag{11}$$

where $r(t)$ and θ represent the reflection coefficient and phase of the second path. τ is the time delay of the second path relative to the LOS path [8]. $r(t)$ is expressed in terms of its vertical and horizontal polarization coefficient, respectively. The reflection coefficient in Time-Domain (TD) is expressed as

$$f_T(t) = \begin{cases} -\left[K\delta(\tau_r) + \dfrac{4k}{1-k^2}\dfrac{e^{-aT_d}}{T_d}0\sum(-1)^{n+1}nK^bI_n(\alpha\tau_r)\right], hp \\ \left[K\delta(\tau_r) + \dfrac{4k}{1-k^2}\dfrac{e^{-aT_d}}{T_d}0\sum(-1)^{n+1}nK^bI_n(\alpha\tau_r)\right], vp \end{cases} \tag{12}$$

$$where \quad K = \frac{1-K}{1+K}, \begin{cases} K = (\xi\varepsilon_r)^{-1} and \quad \alpha = \dfrac{\xi}{2}, hp \\ K = \xi \qquad and \quad \alpha = \dfrac{\nu}{2}, \dfrac{\cos^2\phi}{\varepsilon_r} \langle\langle 1, vp \end{cases}$$

$$\xi = \frac{\sqrt{\varepsilon_r - \cos^2\phi}}{\varepsilon_r \sin\phi}, \quad \zeta = \frac{\nu}{1-\dfrac{\cos^2\phi}{\varepsilon_r}}, \nu = \frac{120\pi\sigma\lambda 0}{\varepsilon_r}$$

where ε_r is the normalized relative dielectric constant of the reflecting surface, and σ is the conductivity of the reflected surface [10]. I_n is the modified Bessel function and it has tremendous complexity. The approximate expression for $r(t)$ is given by

$$r(t) = \pm\left[K\delta(t) + \frac{4K}{1-K^4}\frac{e^{-\alpha t}}{t}\sum_{n=1}^{N}(-1)^{n+1}nK^nI_n(\alpha t)\right] \tag{13}$$

4.2 Quadratic vs. Cubic-Phase Chirp in Two-Ray Multipath

The reflected surface is generally the dry concrete. The parameters, ε_r and σ in this environment is 5 and 0.7. And the proposed novel cubic-phase chirp has the 50ns time duration. The parameter αt in (13) would be over 100 and the second term in (13) is approximately zero. Then the reflection coefficient function is rewritten by

$$r(t) = \pm K\delta(t) \qquad K = \frac{1+K}{1+K}\begin{cases} K = \dfrac{\sqrt{\varepsilon_r - \cos^2\theta}}{\varepsilon_r \sin\theta}, vp \\ K = \dfrac{\varepsilon_r \sin\theta}{\sqrt{\varepsilon_r - \cos^2\theta}}, hp \end{cases} \tag{14}$$

The distance between the transmitter and the receiver is 10m and the height is 3m. And the reflected surface is the dry concrete wall. Fig 8 provides the autocorrelation of quadratic and cubic-phase chirp signal in Two-ray multipath model. The signals are with two methods (amplitude tapering and internal hamming weighting) of side-lobe reduction suggested in previous section.

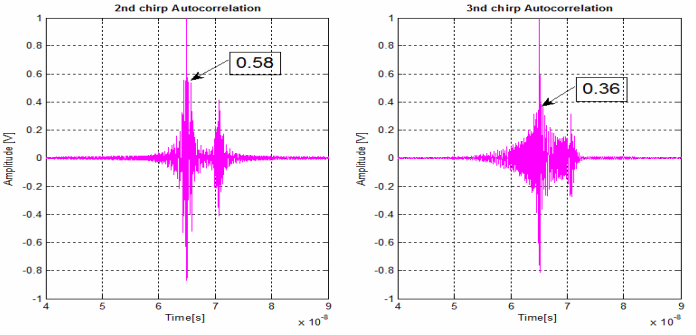


Fig. 8. Autocorrelation in Two-ray Model

The ratio of main-lobe to 1st side-lobe in both signals are increased. But the ratio of the cubic-phase signal is higher than the quadratic-phase signal. The cubic-phase signal realizes higher accuracy on ranging system in multipath model.

5 Simulation Analysis

5.1 Signal Properties

For empirical communication environment the propagation loss exists along to the path. It increases proportionally according to the distance between the transmitter(TX) and the receiver(RX). In this paper, the log-normal path loss model is adopted.

$$\overline{PL(d)_{dB}} = PL(d_0)_{dB} + 10\eta \log_{10}(\frac{d}{d_0}) + S(d);\ d > d_0 = 1m \tag{15}$$

$$\overline{PL_{FS}(\lambda, d)} = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{16}$$

where PL_0, the intercept point, is the path loss at $d=1m$, $10\eta log_{10}(d/d_0)$ is the median path loss referenced to 1m and η is the path loss exponent whose value is 2 in free space and it depends on the structure, λ is the wavelength that can be obtained by $\lambda=c/f=c/f_s$ and S is the log-normal shadow fading in dB [9]. The shadow fading term S has an RMS (Root Mean Square) value of the standard deviation σ in dB. And the simplified path loss model is presented in reference [11] for LOS (Line of Sight) and non-LOS.

$$LOS \quad : \quad P_{RX} = e^{-1.9d} P_{TX} \tag{17}$$

$$non-LOS: \quad P_{RX} = e^{-3.4d} P_{TX} \tag{18}$$

where e stands for 'exponential' and d is the transmitter-receiver distance in meters. The received signal power in UWB will decay by e. In this chapter, the simplified path loss model in LOS is adopted in simulation.

The correlation process is for signal compression and detection. There are two kinds of correlations as auto-correlation and cross-correlation that performs for itself

and between different signals, respectively. The auto-correlation function is stated in (19) and the cross-correlation can be obtained just by replacing *s(t)* to other signal.

$$\phi_{cor}(t) = \int_{-\infty}^{+\infty} s(\tau) \cdot s(t-\tau) d\tau \tag{19}$$

5.2 DS-UWB Ranging Simulation

The chirp signal ranging simulations evaluates with the cubic-phase and the quadratic-phase chirp at 2 GHz bandwidth and the time duration T_s is 50ns . The ranging in the SNR -15 ~ 0 dB shows quite high accuracy at both the cubic-phase and quadratic-phase chirp. So the ranging simulation in the rough environment (SNR = -50 ~ -15dB) was performed. The ranging simulation evaluates by average of 500 times repetition under the SNR -50 ~ -15 dB in AWGN channel. The parameters of chirp DS-UWB ranging simulation are presented in Table 2.

Table 2. Parameters of Chirp DS-UWB Ranging Simulation

SNR	-50 ~ -15 dB
Available UWB bandwidth	2 GHz
Sampling frequency	20 GHz
Carrier frequency	3.1 GHz
Spreading code length	7
Chirp duration	50ns
Channel Model	AWGN
Data modulation	BPSK
repetition	500 timed

Ranging performance when the amplitude tapering and internal hamming weighting was used in both signals in AWGN channel environment. The distance between Tx and Rx is 30m and 50m.

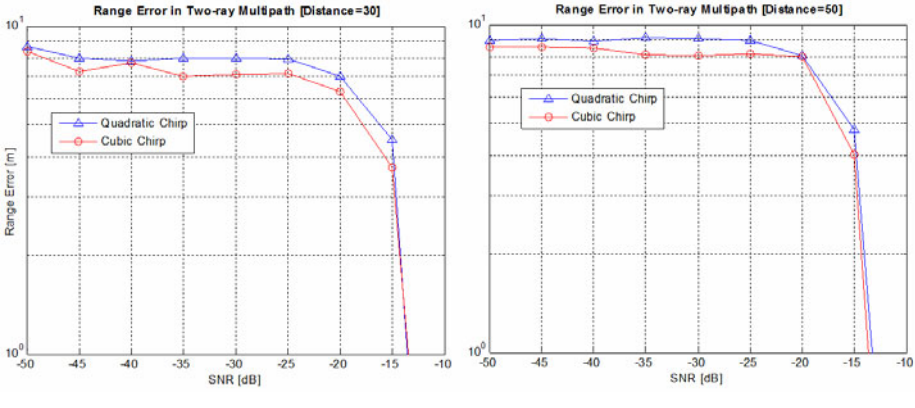


Fig. 9. Chirp Ranging (SNR = -50 ~ -10dB)

The results of simulation show the improved ranging accuracy in cubic-phase chirp signal throughout the SNR -50~-10 dB in the 10m distance between Tx and Rx. Fig 9 shows the chirp ranging under Two-ray multipath model with path loss. And the chirp ranging in Fig 9 is performed under 30m and 50m distance. Each simulation evaluates the ranging simulation at LOS with path loss. But the cubic-phase chirp has the lower ranging error than the quadratic-phase chirp at all environment. The cubic-phase chirp has higher accuracy in ranging system.

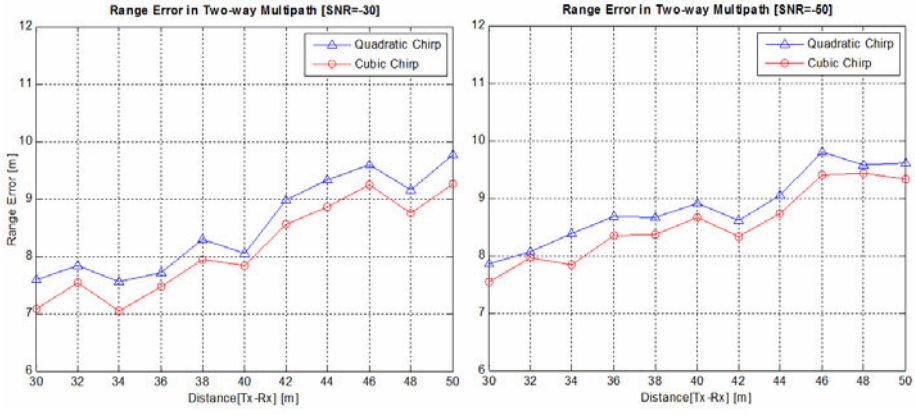


Fig. 10. Chirp Ranging (Distance = 30 ~ 50m)

The results of simulation in Fig 10 show the improved ranging accuracy in cubic-phase chirp signal throughout the distance 30~50m. Fig 10 shows the chirp ranging under two-ray multipath model. simulation evaluates the ranging simulation at LOS with path loss. Fig 10 show the ranging error when SNR is -30dB and -50dB. But the cubic-phase chirp has the lower ranging error than the quadratic-phase chirp at all environment.

6 Conclusion

The quadratic and cubic-phase chirp signal was introduced in this paper and research the potential of its use for UWB ranging. The cubic-phase chirp signal exhibits better correlation properties in the mainlobe-1st.sidelobe ratio than quadratic-phase chirp signal. But the cubic-phase chirp signal has side-lobe ripples. So the methods to suppress the side-lobe ripples are suggested. In the amplitude tapering Tukey window function is adopted in transmitted signal in time domain. And the internal Hamming weighting function is used in frequency response. Both methods are adopted in ranging system with quadratic and cubic-phase chirp signal. The Cubic-phase chirp signal shows the significant high accuracy in ranging system. The enhanced correlation properties in the mainlobe-1st.sidelobe ratio is higher than the quadratic-phase chirp signal. Next the multipath model is considered and Two-ray Model is adopted in a ranging system. Two-ray Model is the simplified multipath model. It is

composed with the direct path and the reflected path. The reflected path has the reflection coefficient. The cubic-phase chirp signal has also higher accuracy in two-ray model. Finally the computer simulation in AWGN channel confirmed the improved ranging accuracy for cubic-phase chirp signal. The cubic-phase chirp signal revealed enhanced ranging capability with the enhanced correlation and low PAPR advantage.

References

1. Di Benedetto, M.-G., Giancola, G.: Understanding Ultra Wide Band Radio Fundamentals. Prentice Hall, Upper Saddle River (2004)
2. Federal Communications Commission, Revision of part 15 of the Commission's rules regarding ultra-wideband transmission system, First order and report Tech. Rep. FCC 02-48 (2002)
3. Cook, C.E., Bernfeld, M.: Radar Signals. Artech House, Norwood (1993)
4. Springer, A., Gugler, W., Huemer, M., Reindl, L., Ruppel, C.C.W., Weigel, R.: Spread Spectrum Communications Using Signals. In: Proc. IEEE/AFCEA EUROCOMM Conference, pp. 166–170 (May 2000)
5. Harris, F.J.: On the use of windows for harmonic analysis with the discrete Fourier transform. Proc. IEEE, 51–83 (1978)
6. Shin, Y.-i., Park, J.-g.: An Optimal High-Order-Phase Chirp for DS-UWB Ranging
7. Kowatsch, M., Stocker, H.R.: Effect of Fresnel ripples on sidelobe suppression in low time-bandwidth product linear FM pulse compression. Proc. IEEE 129, 41–44 (1982)
8. Kim, E.S., Kim, J.I., Kang, I.-S., Park, C.G., Lee, J.G.: Simulation results of Ranging Performance in Two-ray Multipath Model. In: International Conference on Control, Automation and System, pp. 734–737 (October 2008)
9. Di Benedetto, M.-G., Kaiser, T., Molish, A.F., Oppermann, I.: UWB Communication System: A Comprehensive Overview. Hindawi Publishing Corporation (2006)
10. Yao, Y.: A Time-domain ray model for Predicting Indoor Wideband Radio Channel Propagation Characteristics. In: IEEE International Conference, Universal Personal Communications Record, vol. 2, pp. 848–852 (1997)
11. Mei, H.: Modeling and Performance Evaluation of a BPPM UWB System, MD thesis, Delft University of Technology, Delft Netherlands (2003)

Device Collaboration System Based on Context-Aware Architecture in Smart Environment

Sook-Youn Kwon, Chan-Yung Choi, and Jae-Hyun Lim

Computer Science Department, Kongju National University,
Budeadong 275, Cheonan, Chungnam, Korea
{sookyoun,chlcks0,defacto}@kongju.ac.kr

Abstract. In recent times, research is being actively carried out to enable the seamless use of a single content using many different varieties of devices in smart environments, anytime, anywhere. However, present researches are limited to receiving provider-centered services from various cell phones with the same operating system and therefore this thesis suggests, within this model, a device collaboration system based on context-aware architecture to provide a user-centered service. To allow content sharing between different devices, this system collects not only internal/external contextual information but the locations of all devices within the smart environment, resources and service information. Then, through suggested context interpretation structure it predicts the user's intention and service, and selects the best device right for using the content to be run and therefore provides a personalized user-centered service. To verify the suggested system structure, smart environment and scenario were built and whether or not the content sharing service functioned through the best device depending on context-awareness was tested.

Keywords: Context-Aware, Content Sharing, Device Collaboration, Smart Environment.

1 Introduction

In recent smart environments, various devices owned by individuals have increased in number and users are increasingly requesting to seamless use the contents they own using any kind of phone anytime and anywhere[5][11][14]. In other words, personally selecting the device that is appropriate within the vicinity of the user and the environment and thereby carrying out the service wanted regardless of place and time, or according to context information such as the type of the content to be played, the user's location or time and preference, intelligently infer according to the most appropriate device system and receive the user-centered services[1][2][6]. To provide identical content-using service within different devices in this kind of smart environment, research in communication technology within devices, context-awareness technology and collaboration technology between devices and more are being carried out[4][9][10][13]. Service is provided targeting various cell phones with identical operating systems installed in them or a provider-centered service structure where a specialized service that is already registered in the server and is provided and

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 256–263, 2011.

providing service due to user request, are the characteristics of present researches [9][13]. Unlike these, in this thesis the system is designed to allow content sharing regardless of platform, and the context of the user and the device are analyzed through context-awareness structure; through these it suggests a context-aware base device cooperative system that infers the optimum device depending on the user's location.

The structure of these is as follows. In page 2, as related research Stanford University's iRoom and DLNA and IBM's Celadon project were looked into and in page 3 the structure of device cooperative system based on suggested context interpretation structure is explained. In page 4 test bed following the scenario is restored and confirmed through testing the possibility of the suggested system, and in page 5 the conclusion and the future project is described.

2 Related Works

2.1 Standford University's iRoom

Stanford University's iRoom(interactive room)was researched as a part of Interactive Workspaces Project(i-Workspaces) of 1999. This project started with the aim of maximizing the efficiency of conferences, discussions, researches and more, using screen-sharing by high quality display devices within a Ubiquitous computing environment and was produced based on iROS(Interactive Room Operating), an open source base[3]. By everybody operating on each person's computers, not only can individual operations be carried out at the same time but resources and information can be shared through sharing of display devices.

2.2 IBM's Celadon

IBM's Celadon which has been pushed ahead since 2004 is a context-aware based intelligent ubiquitous cooperative system project[9][13]. It supports cooperative management within certain regions to allow the diverse types of wireless cell phones to communicate wirelessly with devices close by and even other variety of devices and equipment with platforms that are different to the operating systems to be interlocked to provide a personalized service which is its aim. However, this project only targets services provided to the user that is already registered in Celadon Zone. Such methods cannot be viewed as a personalized service that considered the individual user's context. Therefore this research aims to provide a more personalized service by predicting intelligently the user's intention and type of service and the best device for running the contents.

3 Proposed Device Collaboration System

This research proposes a device collaboration system based on context-aware architecture for content sharing between heterogeneous devices. Device collaboration system blueprint is made up of Device Collaboration Server and Device Manager, as shown in Fig. 1.

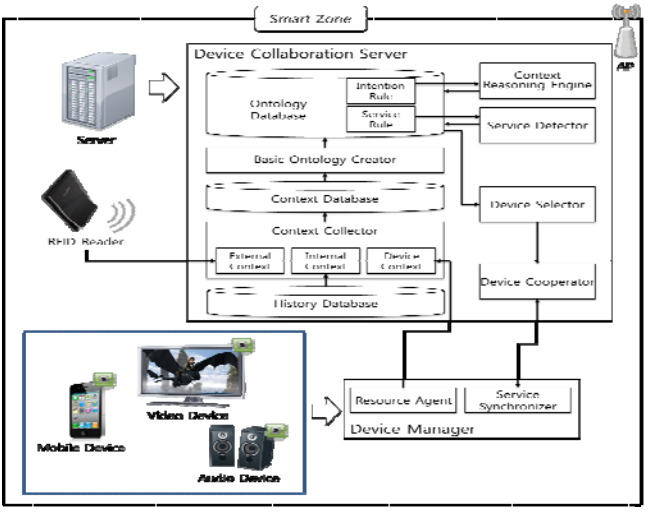


Fig. 1. Device Collaboration System Architecture

3.1 Context Collector

3.1.1 Internal Context

Internal context information means accumulated data that has been entered into the History Database before such as the user's identity, preferences and service background information. Table 1 is user's service background information stored in History Database.

Table 1. User's Background Information

User	Audio Device	Display Device
User_A	001cbf5d3b52	00e09120852f
User_A	001cbf5d3b52	001cbf5d3b52
User_B	001cbf5d3b52	00242160d545
User_B	cc08e07937d8	00242160d545

3.1.2 External Context

Outer context information collected in Context Collector is used to give the best service according to the location of the user. Firstly RFID Tag is attached to user's mobile device and other similar devices and Tag information is collected using Reader. Then the collected data is sent to Context Collector every certain period of time and analyzed to be used as outer context information when predicting device's location. Device's location information within Smart Environment is collected through the RFID system in order of read count, protocol id, antenna id and id, and is stored in the Context Database.

3.1.3 Device Context

Device Context within Context Collector is collected with all the device resource and service information within Smart Environment using SIGAR (System Information Gatherer and Reporter). SIGAR API is a Java based library used to collect same types of resource data from various operating systems and the collected resource and service information is as shown in Fig. 2.

```
CPU Model        : Core(TM)2 Duo CPU     T7500  @ 2.20GHz
CPU Vendor       : Intel
CPU TotalCPUs    : 2
CPU Mhz          : 2194
OSDescription    : Microsoft Windows Vista
OSArch           : x86
IP               : 192.168.0.4
MonitorWidth     : 1280
MonitorHeight    : 800
MemoryTotal      : 2047
MemoryUsed       : 1157
MemoryFree       : 889
Proc ID : 4      , Proc Name : System
Proc ID : 280    , Proc Name : smss
Proc ID : 372    , Proc Name : csrss
Proc ID : 432    , Proc Name : csrss
Proc ID : 440    , Proc Name : wininit
Proc ID : 492    , Proc Name : services
Proc ID : 516    , Proc Name : winlogon
```

Fig. 2. Device's Resource and Service Information

3.2 Ontology

For content sharing service in heterogeneous devices, a lot of information including user's internal/external and device location, resource and service is required[12]. In this thesis, live-time data adding, editing and deletion are made easy and by the relationship of the data is defined using rules of ontology that can reduce the complexity to intelligently predict the user's intention, service and the device [7].

3.2.1 Basic Ontology Creator

Basic Ontology Creator is a basic ontology structure generator that reads and interprets internal/external context information and device information in Context Database in applications to allow. Firstly, the basic ontology structure created by the Basic Ontology Creator is entered into the memory as a file that consists of abstract classes and properties. Then the context information stored in Context Database is converted into a format that is appropriate for the basic ontology structure that is registered n the memory and the registered basic ontology structure is generated as ontology base model. Here, data adds and defines class, individual, property and relation to be compatible with the collected data using Jena.

3.3 Context Reasoning Engine

Context Reasoning Engine is a context predicting engine that predicts user's intention by applying intention rule to the ontology base model. For understanding user's intention, Intention rule to be applied to the ontology base model is generated through the four context factors' values which are the object, time, location and action[8].

```
@prefix pre: <http://witlab.kongju.ac.kr/DeviceCollaboration.owl#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix owl: <http://www.w3.org/2002/07/owl#>.

[ObjectRule: (?X rdf:type pre:Device) -> (pre:Object pre:hasObject ?X) ]
[LocationRule: (?X rdf:type pre:Device), (?X pre:isLocation ?Y),
               (pre:Object pre:hasObject ?OX), equal(?X, ?OX) ->
               (?OX pre:isLocation ?Y) ]
[TimeRule: (?X rdf:type pre:Device), (?X pre:isTime ?Y),
               (pre:Object pre:hasObject ?OX), equal(?X, ?OX) ->
               (?OX pre:isTime ?Y) ]
[ActionRule: (?X rdf:type pre:Device), (?X pre:isActiveService ?Y),
               (pre:Object pre:hasObject ?OX), equal(?X, ?OX) ->
               (?OX pre:isAction ?Y)]

[IntentionRule_wantBigDisplay1: (pre:Object pre:hasObject ?X),
               (?X pre:isAction ?Y), equal(?Y, "movieplayer") ->
               (pre:Intention pre:hasIntention pre:wantBigDisplay) ]
[IntentionRule_wantBigDisplay2: (pre:Object pre:hasObject ?X),
               (?X pre:isAction ?Y), equal(?Y, "pictureplayer") ->
               (pre:Intention pre:hasIntention pre:wantBigDisplay) ]
[IntentionRule_wantGoodSpeaker1: (pre:Object pre:hasObject ?X),
               (?X pre:isAction ?Y), equal(?Y, "musicplayer") ->
               (pre:Intention pre:hasIntention pre:wantGoodSpeaker) ]
[IntentionRule_wantGoodSpeaker2: (pre:Object pre:hasObject ?X),
               (?X pre:isAction ?Y), equal(?Y, "movieplayer") ->
               (pre:Intention pre:hasIntention pre:wantGoodSpeaker) ]
```

Fig. 3. Intention Rule

This ontology base model and intention rule summoned from Ontology Database and it is made to predict the user's intention through Context Reasoning Engine and the intention rule can be defined like in Fig. 3.

3.4 Service Detector

Service Detector applies the Service Rule to the ontology model with the user's intention predicted, and predicts the service needed by the user. The service rule is applicable to all the services that can be used within the smart environment. The executable service list is defined as the base of the service information collected from device, and the service appropriate for the intention is defined through this type of rule.

3.5 Device Detector

Device Selector searches the device list for what is appropriate for carrying out the predicted service from Service Detector. Here SPARQL inquiry is carried out to request to search and the result is delivered to Device Cooperator and the delivered result

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX : <http://witlab.kongju.ac.kr/DeviceCollaboration.owl#>
SELECT ?Device
WHERE {{ ?Device rdf:type :Device . ?Device :hasMemoryFree ?MFree .
         ?Device :hasDisplayInch ?DI .
         ?Device :hasLocation ?Lo .
         FILTER ( ?MFree >= 200 ) . FILTER ( ?DI >= 15 ) .
         FILTER ( ?Lo, "A" )
}
```

Fig. 4. Example of SPARQL Inquiry Statement

represents the device that provides the content sharing service and the provided device or the device list and selection criteria uses the cumulative user's past service information or preference information, and capacity information such as the device's contrast or audio. Fig. 4 is an example of the SPARQL inquiry for display device search.

3.6 Device Cooperator

Device Cooperator aims for content synchronization and control of cooperation between device that provides service and device that receives service. Synchronization of services and contents between devices is based on Time Synchronization and it sends service type, content and Time Stamp information to each device where the service will be run in and it resends the according reply signal to the server and so enables confirmation of connection, information delivery and synchronization.

4 Experiment

4.1 Test Bed

For the verification of the system structure proposed in this thesis, smart environment and the test bed were created for the user's mobile device, device around the user and the execution of data within the environment to carry out the experiment.

The device cooperative server was created using JAVA SE 1.6 with Windows Server 2008 R2 as a base. All of the data were managed by installing MySQL Server 5.1.41 and Jena 2.6.3 was used for creation, editing and prediction of context interpreting structure's ontology. For user mobile device iPhone4 which contains iOS4.1 operating system was used and the experiment was carried out by creating a wireless environment.

4.2 Device Cooperation Experiment and Result

In this research scenario was constructed for the verification of the suggested system structure and like in Fig. Monitoring screen consists of Device List, Process List and

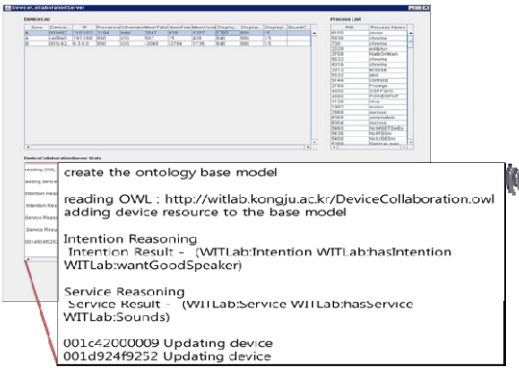


Fig. 5. Monitoring Screen

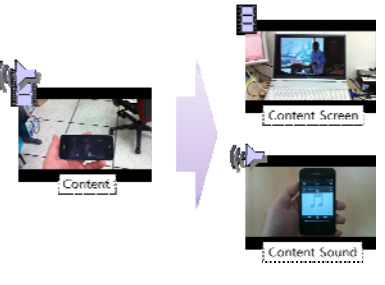


Fig. 6. Test Result

Device Collaboration Server State and device list shows all device location and resource information within the smart environment. Process List generates currently operating service list for each service and Device Collaboration Server State shows the created ontology model's result, predicted user's intention and service and device information. Fig. 6 is the tested result screen based on the scenario and service is provided by User A viewing the video entering the smart environment and sound is played in the user's device according to the user's accumulated information and the video is serviced by being transferred to a different, more advanced display device that is nearby.

5 Conclusion

In this research, user's internal/external context information and device's context information within smart environment is analyzed to intelligently predict the user's intention and service and searches for the more appropriate device and a user-centered personalized service is provided through a device collaboration system based on context-aware architecture for content sharing. Through the suggested system structure, it could be verified that service was provided where content that is being played is freely transferred to another device depending on the context information of user and device. In the future various research such as video streaming service and other effective file management method due to contents will be needed to increase the performance of the device cooperating system.

Acknowledgements

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0028289)

This work was supported by the Grant of the Korean Ministry of Education, Science and Technology (The Regional Core Research Program / Zero Energy Green Village Technology Center)

References

1. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a Better Understanding of Context and Context-Awareness. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, pp. 304–307. Springer, Heidelberg (1999)
2. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5, 4–7 (2001)
3. Brad, J., Armando, F., Terry, W.: The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms. IEEE Pervasive Computing Magazine Special Issue on Systems 2, 67–74 (2002)
4. Naohiko, K., Kenta, M., Kazunori, T., Hideyuki, T.: Smart Device Collaboration for Ubiquitous Computing Environment. In: Proceedings of the Workshop on Multi-Device Interface for Ubiquitous Peripheral Interaction at the 5th International Conference on Ubiquitous Computing (2003)

5. Youngblood, G.M., Heierman, E.O., Holder, L.B., Cook, D.J.: Automation intelligence for the smart environment. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence, pp. 1513–1514 (2005)
6. Gu, T., Pung, H.K., Zhang, D.Q.: A Service-Oriented Middleware for Building Context-Aware Services. Journal of Network and Computer Applications 28, 1–18 (2005)
7. Whee, S.: A Theoretical Study of Using Methods for OWL Vocabulary and Syntactics to Ontology Automatic Construction in Korean. Journal of Korean Library and Information Science Society 37(2), 191–216 (2006)
8. Jae-Hyun, L.: Design and Implementation of Context-Awareness Interpretation Structure for Intelligent Service in Korean. Journal of Korean Society for Internet Information 7(5), 123–134 (2006)
9. Lee, M.C., Jang, H.K., Paik, Y.S., Jin, S.E., Lee, S.: Ubiquitous Device Collaboration Infrastructure: Celadon. In: The Fourth IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems, and the Second International Workshop on Collaborative Computing, Integration, and Assurance (2006)
10. Lee, H.Y., Kim, J.W.: An Approach for Content Sharing among UPnP Devices in Different Home Networks. IEEE Consumer Electronics 53(4), 1419–1426 (2007)
11. Cook., D.J., Das, S.K.: How smart are our environments? An updated look at the state of the art. Pervasive and Mobile Computing 3, 53–73 (2007)
12. Chan-Yung, C., Jae-Hyun, L.: Design of Device Resource Management System for Service Mobility in Home Network Environment in Korean. Journal of Korean Society for Internet Information 11(2), 113–114 (2010)
13. McFaddin, S., et al.: Celadon: Delivering Business Services to Mobile Users in Public Spaces. IBM Research Report, RC24381 (2007)
14. Mun, H.R., Kim, S.H., Chung, B.H.: Context Sharing Technology Trend Analysis in Korean. Technical Report, ETRI (2010)

Developing English Learning Contents for Mobile Smart Devices*

Seongwon Park[1,3], Kwangeak Kim[2], and Bong Gyou Lee[3,**]

[1] Research Institute for Information Strategy Consulting, Seoul, Korea
[2] LiveABC Co.,Ltd, Seoul, Korea
[3] Graduate School of Information, Yonsei University, Seoul, Korea
{seongwon,bglee}@yonsei.ac.kr, kkwang21@liveabc.co.kr

Abstract. The study developed three items in total. First, the application which enables learning contents run on Android OS and Windows Mobile OS was developed. Second, contents which will be applied on the application were manufactured. The contents developed in this study are for learning English presentations. The contents are classified into two parts; Part 1 is for basic English presentations, and Part 2 is for advanced English presentations. The third development is Learning Management System(LMS) that is driven by connected with PC. Contents studied on mobile devices are recorded by LMS, so that users can learn continuously not only on mobile devices but also on PC. We analyzed the questionnaire responses with respect to UI satisfaction and satisfaction of the learning experience. The UI satisfaction results showed that 85% of the participants were satisfied at an ordinary or higher level with our system. And The satisfaction of the learning experience results showed that 95% of the participants were satisfied at the ordinary or higher level with our system.

Keywords: Mobile Learning(m-learning), Smart Learning, English Learning.

1 Introduction

Low quality of the early mobile devices, with insufficient computational power, small displays, and difficulty in accessing the wireless internet connection, constituted a severe barrier to the deployment of rich educational materials[1]. Present mobile communication terminals show unparalleled improvement in performance as well as functions compared to early voice call terminals. As mobile communication terminals have high performance and multi-functionality, the number and size of applications embedded in terminals also are surging naturally. Interests of terminal manufacturers, consumers, and communication companies coincide with each other, the market of smart phones naturally grows rapidly [2]. Fig.1 shows the expected volume of sales classified by smart phone OS and platforms, predicted by ROA group. As shown in Fig. 1, currently Google Android, Windows Mobile and iPhone OS enjoy the greatest

* This research was supported by a research grant of Seoul R&BD Program (# PA090747).
** Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 264–271, 2011.

popularity [3]. The study aims to develop mobile contents for English learning based on Android OS and Windows Mobile OS. In recent years, mobile technologies have dramatically evolved, promoting interest in mobile learning systems. Mobile learning is a newly emerging learning style and offers learners the opportunity to learn at anytime and anywhere. And using digital materials for mobile learning is being popularized by education sector. Digital educational materials include text, picture, or video that are made by computer software [4]. Mobile learning(m-Learning), the next generation of the computer-aided and multimedia-based learning, is based on mobile phone. It can be said that the great majority of the population have mobile phones and carry it with them most of the time. Because of this fact, the independency of time is achieved and space for learning is secured nearly a hundred percent. As a consequence, the main advantage of mobile learning is that you can learn wherever and whenever you want to. You can use idle periods for learning, for example: The time while you are travelling, while you are waiting for the bus, or while you are waiting at the restaurant. Nearly every unused and wasted time can now be used for efficient and effective learning. Therefore mobile learning will be an important instrument for learning, because it will help us to use our time more efficiently [5]. In this research, We developed a mobile English learning system based on Android OS and Windows Mobile OS.

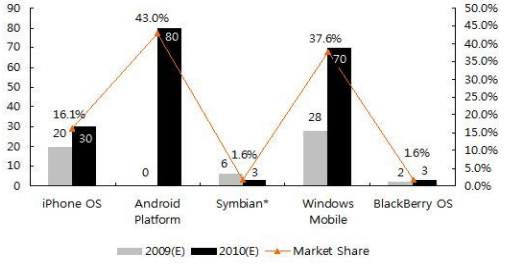


Fig. 1. Market share of smartphone

In Section 2, we review the previous work in mobile learning. Section 3 addresses mobile learning system we developed. Section 4 concludes with findings, limitations and future research.

2 Related Work

2.1 Mobile Learning

M-learning means that users can study by utilizing mobile devices including mobile phones and PMP at any places any time [6]. Chen et al.(2002) argue that the characteristics of m-Learning are accessibility meaning users can access learning materials anywhere anytime, immediacy meaning users can acquire learning materials immediately under any conditions, and situating of instructional activities meaning meaning users can apply learning on daily lives [7]. Roschelle(2003) insists that the technology of m-learning allow users participate in various learning activities including

searching information, access to online information contents, and discussions [8]. Ole & Judith(2003) provide PDA to students and investigate results in order to study the usefulness of mobile technology for medical education and clinical training. Research outcomes prove that students use PDA for communication related with learning [9]. Pauliina & Harrri (2003) state that m-learning brings a new paradigm for learning models with its characteristics including portability as well as mobility [10].

2.2 English Learning by Mobile Devices

Harrison(1999) said that learners can use m-learning as an educational tool to establish self-directed learning, and Internet acts as a catalyst for learners to more actively participate in learning by searching and acquiring required information materials [11]. Myers(2000) told that when making Chinese people learn English with mobile devices which provide voice recognition and interpretation programs, learning with mobile devices was proved to have positive influence on the improvement in learning English [12]. Houser(2005) let college students in Japan study English on mobile devices after sending English learning materials to them, and results showed that learning English using mobile devices appeared to be more effective for students than learning with PC. Also, evaluation results of students for learning English using mobile devices also were positive [13]. Kiernan & Alzawa(2004) studied the usefulness of mobile phones as a tool for studying English, and claimed that it has merits compared to previous learning methods [14].

3 Mobile English Learning System

The study developed three items in total. First, the application including contents transfer system which enables English learning contents run on Google Android OS and Windows Mobile OS was developed. Second, contents which will be applied on the application were manufactured. The third development is Learning Management System(LMS) that is driven by connected with PC. Fig. 2 shows the mobile service for English learning.

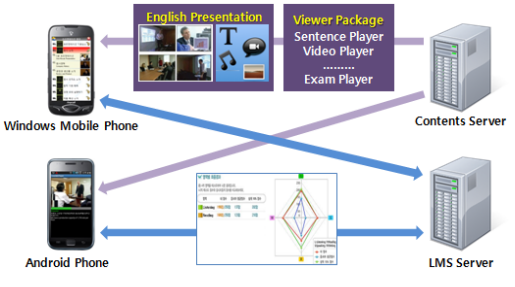


Fig. 2. A structure of mobile English learning system

3.1 Mobile Learning Application Running on Smartphones

Since the development of platforms for smart phones have different development languages and data structures by OS, respective programs should be developed. In case of W from Microsoft, an authorization code should be checked first before reading data for security. The authorization code is created when installing applications by using private ID of a phone, and is used to identify whether the product is installed in proper manners. Accordingly, those who installed the illegally counterfeited player are prohibited from using. Also, in order to use a built-in codec of Windows media player to play videos, .mp3 is available for sound sources while .wmv is available for video files. Sound sources have one mp3 for each unit of sentences, because it is almost impossible to move precisely to the playing stage of each sentence among all sound sources and there is difficulty for repeating sentences as well. XML was applied since data format does not utilize outside database, and because formatting XML is determined by data types and XML files are decided by kinds of languages, related XML cab ne used regarding the language of URI. The service developed in the study based on Android and Windows Mobile also has same operation and functions.

The left picture of Fig. 3 shows the scene of English education contents while playing on Google Android OS while the picture on the right side is for the scene on Windows Mobile OS. To run an application for Google Android OS, Galaxy S from Samsung Electronics was used while Omnia 2 from Samsung Electronics was used for Windows Mobile OS.

Fig. 3. The application running on Google Android and Windows Mobile

In this system, multimedia is provided to learn including texts, sounds, images, and movies. In view of smart phone, user interface is designed suitably for touch. The player developed in this study provides various functions as shown from Fig. 4: first, moving to the previous or next screen: second, selecting one language from three options including ‘Korean and English’, ‘Only English’, and ‘Only Korean’: third, learning by sentences, learning the whole chapter (by cases), and quiz: fourth, playback (the whole or one part): fifth, stop and pause: and sixth, a dictionary.

Fig. 4. Menus in the player

3.2 Contents for English Presentations

The contents developed in this study are for learning English communication. In specific, it is to study English presentations. The contents are classified into two parts; Part 1 is for basic English presentations and Part 2 is for advanced English presentations. Each part is made up with nine units, and each unit is composed differently by topics. The number of whole chapter for both parts is fifth-one. The contents include various presentations, for example there are "Company Introductions", "Launching a New Product", and "delivering an Annual Budget". Table 1 shows contents by parts, by units, and by chapters.

Table 1. A list of presentation contents

Part 1	Title
Unit 1	Making a Presentation in English
	1. Beginning the Presentation
	2. Introducing the Products
	3. Wrapping Up the Presentation
Unit 2	Company Introductions
	4. One-Minute Presentation
	5. Company History
	6. Factory and Product
Unit 3	Organization Chart
	7. One-Minute Presentation
	8. The Management
	9. Key Players
Unit 4	Winning Over New Clients
	10. The ODM Company - A Dry Run
	11. Customized Service - Getting Your Hands Dirty
Unit 5	Boasting on the Procedures
	12. Production Procedures - You Can't Go Wrong
	13. The Basics - Nowhere to Go but Up!
Unit 6	Talking Your Product Up
	14. One-Minute Presentation
	15. Microwave Magic
	16. Pick a Fish Tank
	17. Sleek Cell Phone - Product Benefits & Features
	18. Must-have for Travel - Product Benefits & Features
Unit 7	Launching a New Product
	19. The Selling Point - The Fundamentals
	20. Innovative Marketing - Inspiring he Market
Unit 8	Holding a Press Conference
	21. The Press Conference - The Product Launch Begins
	22. Q&A Session
Unit 9	A Successful Sales Call
	23. One-Minute Presentation
	24. Cold Call - The Sales Call
	25. The Sales Call - The Pitch
	26. Closing the Deal

Part 2	Title
Unit 1	How to Use Graph in Presentations
	27. One-Minute Presentation
	28. Line Graph
	29. Bar Graph & Pie Chart
Unit 2	Charting a Presentation
	30. One-Minute Presentation
	31. You Dropped the Ball
	32. Recovering the Field
Unit 3	Media Advertising Theme
	33. Ole Advertising Theme
	34. Print Advertising
Unit 4	Promotion Plan : Play it Cool
	35. One-Minute Presentation
	36. Tuning In
	37. Setting Your Sights
Unit 5	Presenting a New Product
	38. Product Positioning - Man are Different
	39. Channels of Distribution - Rounding Down
Unit 6	Preparing for a Trade Show
	40. One-Minute Presentation
	41. Fill Us In
	42. Final Push
Unit 7	Summing Up Sales Figures
	43. One-Minute Presentation
	44. Sitting Pretty
	45. Getting a Close Up
Unit 8	Delivering an Annual Budget
	46. Meeting Notice - Annual Budgets Meeting
	47. The Budgets- Sales in the Red
	48. General Affairs Presents
Unit 9	Mapping out the Future
	49. One-Minute Presentation
	50. A Blueprint for Success
	51. Moving Up in a Downward Market

3.3 Learning Management System

Contents studied on mobile devices are recorded by learning management system (LMS), so that users can learn continuously not only on mobile devices but also on PC. To enjoy the service on mobile devices, users should save contens in mobile devices, and it can be done in two ways for this system. Firstly, users access the contents server on PC and download files to save in mobile devices. Secondly, users directly access contents server via WiFi, and save contents. LMS saves contents studied on mobile devices, and save them on learning management server as well. Contents saved in LMS include the list of contents, scores of quizes by units, and learning time. The learning records saved in LMS are analyzed, evaluated, and managed by learning management system, so mobile uers can manage their own learning records on PC and review them.

4 System Evaluation of m-Live

We had the participants use the presentation English learning system and the m-Live system developed in this study running on the smart phones based on the Android OS and Windows mobile OS. Afterwards, we conducted a simple survey using a questionnaire regarding the degree of satisfaction for the user interface (UI) recognized by the users and for the learning experience via the contents. The participants in the evaluation included 60 undergraduate and graduate students from two 4-year universities in Seoul. 41 participants were in their twenties, 14 participants in their thirties, and five in their forties.

We analyzed the questionnaire responses with respect to the UI satisfaction and the results showed that 10% of the total participants responded that it was very satisfactory, 35% satisfactory, 40% ordinary, 12% dissatisfactory, and 3% very dissatisfactory, indicating that 85% of the participants were satisfied at an ordinary or higher level with the m-Live system. Table 2 shows the analysis results for the UI satisfaction questionnaire.

Table 2. Analysis results for the m-Live UI satisfaction questionnaire

Level of Satisfaction	No. of Respondents	Percentage
Very satisfactory	6	10
Satisfactory	21	35
Ordinary	24	40
Dissatisfactory	7	12
Very dissatisfactory	2	3

We analyzed the questionnaire responses with respect to satisfaction of the learning experience and the results showed that 10% of the total participants responded that it was very satisfactory, 55% satisfactory, 30% ordinary, and 5% dissatisfactory, indicating that 95% of the participants were satisfied at the ordinary or higher level with the m-Live system. The degree of satisfaction with the learning system was higher than that with the UI. Table 3 shows the analysis results for the learning satisfaction questionnaire.

Table 3. Analysis results for the m-Live learning satisfaction questionnaire

Level of Satisfaction	No. of Respondents	Percentage
Very satisfactory	6	10
Satisfactory	33	55
Ordinary	18	30
Dissatisfactory	3	5
Very dissatisfactory	0	0

5 Conclusions

In this study, we developed mobile content and a mobile learning system for learning of English presentations based on the Google Android OS and Windows Mobile OS which have continuously gained in the world market share.

This study provides the following benefits: first, the foreign language learning content including dictionary, letters, sound, and movies was developed even though there are not sufficient language learning smart phone applications currently, and the smart phone market is expected to continuously grow; second, a vast English conversation content for presentations was planned and produced, as there is almost no content for presentations although it is now in schools and companies; third, we showed the current status of research on learning and English learning using mobile instruments for researchers and providers who want to develop content for learning and English learning using mobile instruments; fourth, we exemplified and suggested the possibility of developing English learning mobile content for the many content developers and content providers who want to deal with mobile contents at mobile market places (App Store, Android Market, etc.) in the future; Finally, the user evaluation showed us what the user was satisfied with or not, suggesting further follow-up work or study for those who want to produce or study mobile contents for learning.

However, this study focused on the production of mobile content that did not exist before, and the effect was not analyzed in depth and empirically. Specifically, the degree of learning satisfaction for the system developed in this study was investigated by means of a simple questionnaire and by interviewing 60 undergraduate and graduate students at two 4-year universities in Seoul, but the degree of learning satisfaction or leaning effectiveness was not analyzed based on strict methodology with the various users in many locations. Therefore, a future empirical study should be done using stricter methods to determine the degree of learning satisfaction and learning effectiveness when using smart phone-based learning content and learning systems. Additionally, further study needs to add more useful functions supported by the conventional PC-based learning contents such as distinguishing between lectures a user watched previously from those a user has not watched yet, or the ability to control the speed of the playback for the learners with better or poorer English ability.

References

1. Nah, F.F.-H., Siau, K., Sheng, H.: The value of mobile applications: a utility company study. Communications of the ACM 48(2), 85–90 (2005)
2. Jang, Y., Kim, C.: The Evolution of Smarthphone Market & The Effect by Android. Journal of KIISE, 48–56 (2010)
3. ROA Group Korea, http://www.roagroup.co.kr
4. Liu, G., Jiao, Z.: The Design of Mobile Learning System for Teachers' Further Education. In: Proc. of 2010 Second International Workshop on Education Technology and Computer Science, vol. 2, pp. 730–732 (2010)
5. Holzinger, A., Nischelwitzer, A., Meisenberger, M.: Mobile Phones as a Challenge for m-Learning: Examples for Mobile Interactive Learning Objects (MILOs). In: Proc. of the Third IEEE International Conference on Pervasive Computing and Communications Workshops (2005)
6. Clyde, L.A.: m-learning. Teacher Librarian 32(1), 45–46 (2004)
7. Chen, Y.S., et al.: A Mobile Scaffolding- Aid-Based Bird-Watching Learning System. In: Proc. of IEEE International Workshop on Wireless and Mobile Technologies in Education (2002)
8. Roschelle, J.: Keynote paper: unlocking the learning value of wireless mobile devices. Journal of Computer Assisted Learning 19(3), 260–272 (2003)
9. Ole, S., Judith, G.: Personal Digital Assistants in medical education and practice. Journal of Computer Assisted Learning 19(3), 320–329 (2003)
10. Pauliina, Harri: Mobile Learning in Teacher Training. Journal of Computer Assited Learning 19(3), 330–335 (2003)
11. Harrison, N.: How to Design Self-Directed and Distance Learning: A Guide for Creators of Web-based Training, Computer-based Training, and Self-Study Materials. McGraw Hill, New York (1999)
12. Myers, M.J.: Voice Recognition Software and a Hand-Held Translation Machine for Second-Language Learning. Computer Assisted Language Learning 13(1), 29–41 (2000)
13. Thornton, Houser: Using mobile phones in English Education in Japan. Journal of Computer Assisted Learning 21(3), 217–228 (2005)
14. Kiernan, Aizawa: Cell Phones in task based learning: Are cell phones useful language learning tools? RECALL 16(1), 71–84 (2004)

Pre-test Analysis for First Experiences of Korean E-Voting Services

Ki Youn Kim[1], Dae Jung Kim[2], and Bong Gyou Lee[1]

[1] Graduate School of Information, Yonsei University, 134 Shinchon-Dong, Seodaemun-Gu, Seoul, 120-749, South Korea
{seansky,bglee}@yonsei.ac.kr
[2] The Grand National Party, 14-31 Yeouido-dong, Yeongdeungpo-gu, Seoul, 150-748, South Korea
blueha@naver.com

Abstract. This study is a pre-test research conducted before introducing an e-voting system by Korean e-Government. The purpose of this empirical study is to establish research hypotheses based on UTAUT model of Venkatesh et al. and to prove causal relationships among defined factors influencing on user acceptance for an e-voting system. In Korea, though actual user groups were extremely limited, we collected data from 116 persons who had experienced trial operation for this e-voting system and derived both theoretical and practical propositions and implications. As the result of this study, including the new variable of 'ubiquity' in consideration of the specific characteristics of e-voting, overall variables defined in this research model are examined partially by means of empirical analysis process to identify the statistical significance. The outputs of this research may provide theoretical basis and political implication necessary for developing countries' planning to introduce e-voting in the future

Keywords: E-Voting, E-Government, IT Policy, UTAUT, User acceptance, Structural equation model, Ubiquity.

1 Introduction

Users with high capability of utilizing information technology showed more positive attitude for the utilization of e-government. As the realization of e-government is being materialized step by step, many reports claim that it is important to evaluate the efficacy and value of e-government in the position of user [1], [2], [3], [4]. As information technology has been developing rapidly, academic discussion for the previous researches on the information system of public sector such as electronic voting(e-voting) and e-government has been actively proceeding [5], [6], [7].

Recently the discussion for introducing e-voting for e-government implementation has been expanded in Korea, a leading country of information technology. Research questions of this study are to explore determinants for user acceptance in an e-voting system. This research is a case study for the e-voting in Korean e-government and is a

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 272–279, 2011.

pre-test for the user group who had actually experienced trial operation before introducing an e-voting system. Though those actual users were extremely limited because e-voting has not been common yet, it is both theoretically and practically meaningful to in that this research collected data from experienced voters' samples and to derive implications.

The scopes of this research are that conceptualize definitions of core factors influencing on user acceptance over e-voting process and identify the research hypothesis and model suitable for the topic of e-voting based on Unified Theory of Acceptance and Use of Technology(UTAUT) model of Venkatesh et al. [8]. We operationally defined element latent variables such as performance expectancy, effort expectancy, social influence, facilitating condition, behavioral intention and use behavior from the UTAUT base model. Additionally, several new variables considering of the specific features in e-voting system and process are added; cost expectancy, security and ubiquity.

The data for analysis was collected from the sample group constituted with the actual users, who had experienced e-voting and extremely limited in numbers. 116 samples group participated in questionnaires were the members in Congress and government who recognized for extensive knowledge of e-voting systems. The type of e-voting system that was selected in this study is touch screen type e-voting. Contribution of this study is to provide theoretical essential model and policy implications to the countries planning to introduce e-voting such as Korean case in the future.

2 Theoretical Background

2.1 E-Voting Services in E-Government

Many information technologies and systems for realizing e-government have been recently introduced. Thus, The viewpoint of researches has been gradually shifted from controller-oriented to user-oriented topic in accordance with the development of e-voting environment [9]. Though there are so many researches in the field of e-voting technologies, prior researches or case studies focused on the behavior of demanders' behavior are very rarely found because the numbers of country actually utilizes an e-voting system now are limited.

An e-voting system means 'the voting method based on ubiquitous infrastructure is available to vote anytime, anyplace by digitalizing the stages of election preparation, voting and ballot counting. Generally, the type of e-voting system is a broad concept including the voting utilizing digital method such as touch screen voting, internet voting, telephone voting and it has many merits like cost saving, turnout increase, invalid vote protection compared to the existing paper ballot.

This study narrowed the definition for the concept of e-voting to an e-voting system utilizing touch screen. A method of a touch screen voting system is as similar as the operation method of ATM (automated teller machine) in a bank. Electors can

select a candidate shown in the screen by touching with a finger. The procedure of voting is as follows. Voters confirm their information from the integrated database of the electoral roll and issue the identified voting card. When a voting card is inserted into the voting machine installed on a balloting table, the screen for voting appears automatically and then electors can select one of candidates shown at the screen for voting. Lastly, after they confirms the result of voting from the print window and receives the discharged voting card, then they input it into the pickup box the voting card. User can vote in any place where internet is connected after receiving an electronic certificate.

2.2 Behavioral Approach of Voting Participation

There were various prior researches with topics for the voting behavior. The factors affecting voting participation are various private characteristics of users such as the sense of civic duty, political ideology, education, income, age and sex and public characteristics such as the type, size, method and period of election. The population-sociological model theory is the classical theory exploring the factors to determine the intention of voting participation. Milbrath and Goel adopted this theoretical model and proved that the higher sociological position and age is, the higher normal participation ratio in politics such as voting activities is [10].

Park verified that the older showed the higher political [11]. Younger group is not proper to have a social pros and cons such as the housing ownership or property rights, so they take a passive about voting participation. Though the explanatory power of demographic and sociological models is excellence, but these have theoretical limitation. This fact implies that the population-sociological variables are not the only absolute values to measure the influence on the intention of voting behavior. On the contrary, the party preferred by electors, political system, and psychological evaluation to candidates are far more important variables.

3 Research Model and Hypotheses

3.1 Research Hypotheses

This chapter defines the latent variables having influence on the dependent variables and examines research hypotheses based on UTAUT base model to verify the causal relationships among defined variables. Key variables are consisted of seven independent variables(performance expectancy, effort expectancy, social influence, facilitating condition, cost expectancy, safety, ubiquity), one mediated variable directly connecting between independent variables and dependent variable(behavioral intention), and one dependent variable(usage behavior). Additionally, we introduced three new variables; cost expectancy, safety and ubiquity to adopt unique features of e-voting. Research hypotheses proposed causal relations among variables.

Table 1. Research hypotheses

Hypotheses	Research hypotheses
H1	Performance expectancy will influence a positive effect (+)over behavioral intention
H2	Effort expectancy will influence a positive effect(+) over behavioral intention
H3	Social influence will influence a positive effect(+) over behavioral intention
H4	Facilitating condition will influence a positive effect(+) over usage behavior
H5	Cost expectancy will influence a positive effect(+) over the behavior of use
H6	Safety will influence a positive effect(+) over behavioral intention
H7	Ubiquity will influence a positive effect(+) over behavioral intention
H8	Behavioral intention will influence a positive effect(+) over usage behavior

3.2 Definition of Key Variables

Venkatesh et al. defined 'performance expectancy' as 'the degree of individual belief that user's utilization of a specific system is helpful to improve the performance of work'. This study redefined performance expectancy as 'the degree of individual belief that the utilization of e-voting can help improving for the matter related to voting activity and ballot counting. UTAUT model defined 'effort expectancy' as 'the degree of ease related to system utilization'. We redefined 'effort expectancy as 'the degree of individual belief that the method of use is easy when user utilizes e-voting'. Social influence suggested by Venkatesh et al. meant 'the level of recognition for the belief that other people, who are valuable to me, should utilize new systems with me'. We redefined 'social influence' as 'the degree of belief that it is desirable for user and others such as neighbors to utilize e-voting.

'Facilitating condition' means 'the degree of individual belief to believe that the systemic, technical infrastructure exists to support the utilization of system'. Taylor and Todd defined 'recognized behavior control' as 'recognition for internal and external constraints against behavior' and reported that this variable was a direct antecedent variable of behavioral intention [12]. In this study, facilitating condition means 'the level of recognizing the necessity for basic fundamental technology or service should be support when user utilizes e-voting'.

We newly introduced 'cost expectancy' variable. It means 'the degree of belief to think that user's utilization of e-voting has an economical utility'. This study defined 'cost expectancy' as 'the degree of user's belief that the utilization of e-voting has an economic utility'. Compared to hand-operated ballot counting, society is deeply concern about the problem of 'safety' such as security and privacy existed in e-voting systems. Igbaria et al. defined 'safety' as 'recognition for the safety level of suppliers or demanders who visit to use digital products' [13]. We redefined 'safety' as 'user's recognition for the safety degree of e-voting providers when user utilizes e-voting '. Ubiquity is a useful and unique characteristic of e-voting and it makes available to participate in the ballot regardless of spatial constraints over entire voting process after the procedure of confirming ID. This study redefined 'ubiquity' as 'the degree of recognition that users feel it possible to utilize e-voting at once, whenever, wherever they want to.

4 Research Design and Methodology

4.1 Data Collecting and Samples

This research is a case study for e-voting systems that is going to be introduced in Korea. Because e-voting has not been now utilized regularly, actual users who had experienced it were extremely limited. We conducted an empirical pre-test based on 116 actual users who had participated in the trial operation. This sample group was consisted of staffs or official member of Congress and governments who had recognized and experienced for e-voting process. And, required data for analysis was collected by online/offline survey and telephone interview. Additionally, three new variables were introduced in this study; safety, ubiquity and cost expectation. With the exception of demographic variables, all items were evaluated by Likert's seventh-point scale ranged from 'not at all'(first- point) to 'strongly agree'(seventh- point).

The sample group of this study was actual experienced users who had recognized accurately e-voting systems' process and method. As e-voting is not introduced nationwide yet in Korea, we surveyed by telephone, email, and directly visit to the aids for congressman and government officers, and the staff for parties and the Secretariat of National Assembly who had an experience of trial vote or an indirect experience. The sample size was 108 persons constituted of 28 persons with an indirect experience and 80 experienced persons. Showing the gender distribution, male(56.5%) was slightly more than female(43.5%), but there was no great difference. Age group was constituted with 30s (42.6%), 20s(32.4%), 40s(18.5%) and 50s(6.5%). In the distribution according to the experience for e-voting, experienced persons were 74.1% and non-experienced were 25.9%.

5 Results

5.1 Prior Analysis

Factor analysis minimized various sub-factors to supra-factors with homogeneity. In other words, it is an analysis method decreasing the numbers of variables. The most universal way of extracting factors are Principle Component Analysis and Common Factor Analysis(CFA). The standard limit of a factor loading that measures factors' correlations established >+/-50 meaning 'very significant' in the conservative limit. Eigen value, the explanatory power for each factors' variance was limited based on the standard >1. Results of the factor analysis are shown like Table 2. Effort expectancy (EE) showed Eigen value(4.24) and distribution(20.19%), social influence(SI) at 3.19/15.21%, facilitating condition(FC) at 2.87/13.70% and cost expectancy at 2.814/13.38%. In general, independent variables were condensed to seven factors and total explanatory power was 85.17%.

Table 2. Factor analysis of independent variables

Var.	1	2	3	4	5	6	7	Total	Vari-ance	Cumu l-ative	α
PE	.305	.104	.196	-.032	.163	.276	**.662**	1.116	5.315	85.174	.922
	.179	.157	.077	.062	.196	.185	**.572**				
	.309	.175	.142	-.171	.019	.294	**.568**				
EE	**.871**	.187	.062	.035	.202	.131	.133	4.240	20.190	20.190	.918
	.861	.192	.033	-.070	.070	.047	.150				
	.757	.327	.160	-.050	.208	.029	.284				
SI	.107	**.711**	.156	-.071	.109	.162	-.073	3.195	15.214	35.403	.871
	.117	**.675**	.088	-.149	.077	.067	-.213				
	.108	**.529**	.154	.016	.057	.238	-.230				
FC	.510	.423	**.761**	-.041	.193	.458	.189	2.878	13.705	49.108	.916
	.263	.195	**.624**	-.006	.122	.160	.239				
	.356	.193	**.580**	-.099	.114	.144	.040				
CE	.340	.113	.147	**.650**	.303	-.067	.106	2.810	13.381	62.490	.749
	.053	.145	.157	**.613**	.082	.063	.058				
	.051	.119	.153	**.531**	.059	.003	.014				
SF	.201	.121	.197	.085	**.745**	.083	.048	2.291	10.910	73.400	.703
	.020	-.039	.002	.147	**.737**	.053	-.110				
	-.009	-.015	.019	.142	**.615**	-.039	-.101				
UQ	-.142	-.090	.039	.140	-.060	**.723**	.081	1.356	6.459	79.859	.912
	.245	.112	.045	-.023	.177	**.649**	-.011				
	-.032	.243	.014	-.223	.177	**.556**	.252				

5.2 Model Fit and Hypotheses Test

CFA analysis was conducted for each factor after reliability analysis. Bagozzi and Yi explained convergent validity is significant when the standard limit of the factor loading is >0.6 and t-value is >1.96 [14]. In this research, as the convergent validity's value showed higher than an ideal standard limit, it might be verified. Summing up the results of the model goodness of fit, they were $\chi 2$(813.840), p-value(.000), RMSR(.036), GFI(.932), AGFI(.917) and NFI(.903).

These hypotheses established based on structural equation model. In significance level p<.001, the variables whose a significant causal link was verified among variables were as follows. As the ubiquity variable was influential over behavioral intention was the path factor coefficient, E=0.37, C.R=5.58. Thus, this variable was proved to have a remarkably significant positive relation with dependent variable. Ubiquity additionally introduced in this study was based on the specific features of e-voting. Performance expectancy was statistically significantly E=0.33, C.R=4.38. That is, the higher users' performance expectancy, the higher behavioral intention was. Social influence was E=0.62, C.R=7.77 and was positively verified to have a strong influence. Results of path analysis among variables are shown at Table 9. Conclusively, facilitating condition, E=0.52, C.R=3.06, was proved to have a positive relation with dependent variable in p <.01 of significance level. Ubiquity, E=.251, C.R=1.914, was proved to have a positive relation with not only parameter but also dependent variable in p<.05 of significance level. But hypotheses about other new variables including safety were partially supported.

Table 3. Path analysis and hypotheses testing

	Path	Estimate	S.E.	C.R.	P	Result
H1	BI ← PE	0.335	0.076	4.383	***	Supported
H2	BI ← EE	-/0.079	0.07	-1.133	0.257	Rejected
H3	BI ← SI	0.625	0.08	7.771	***	Supported
H4	UB ← FC	0.521	0.170	3.060	***	Supported
H5	UB ← CE	-0.098	0.089	-1.100	0.271	Rejected
H6	BI ← SF	0.048	0.044	1.096	0.273	Rejected
H7	BI ← UQ	0.370	0.067	5.558	***	Supported
H7!	UB ← UQ	0.251	0.131	1.914	*	Supported
H8	UB ← BI	-0.19	0.217	-0.878	0.380	Rejected

*p<0.05, **p<0.01, ***p<0.001.

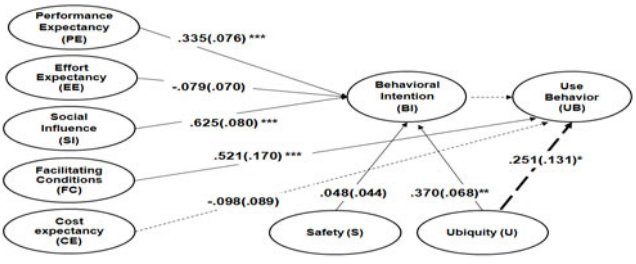


Fig. 1. Confirmatory hypotheses testing

6 Conclusions

This study is an empirical case study to explore determinants influencing on user acceptance for e-voting system. Main contents of this research are to conceptualize important variables through literature reviews. Specially, we added up three new variables; cost expectation, safety, and ubiquity. As the result, the fact was verified that exist a partial possible causal link between independent variables, mediated variable and dependent variable.

The theoretical and practical contribution and implications of this study summed up as follows. Firstly, it is meaningful that this research is a pre-test performed before introducing an e-voting system in Korea. Nevertheless, we selected actual users as the sample group who had experienced the trial operation of the initial e-voting. As the e-voting system is not introduced yet in Korea, it is difficult that collect actual data from the experienced users. So most preceding studies with the similar topic for new information technologies or services such as e-voting would frequently adopt the potential user group. Secondly, looking into various cumulative research models for user acceptance until now, a lot of researchers adopted based on a specific model out of representative models such as TRA, TPB, TAM, C-TAM- TBP, MPCU, IDT, SCT and so on. But this paper adopted UTAUT model as a fundamental model which proposed integrated eight representative models for user acceptance. To investigate possible casual links between variables, we established the adjusted structural

equation model based on Venkatesh et al.'s fundamental model and empirically examined by survey.

The limitation of this study is that it is possible to derive slight different answers from the ordinary user group because our sample constituted mainly with staffs or members of congress government office in the public sector. In future study after introducing e-voting system, it will be meaningful that we would compare with the results from whole people.

Acknowledgement

"This research was supported by the KCC(Korea Communications Commission), Korea, under the CPRC(Communications Policy Research Center) support program supervised by the KCA(Korea Communications Agency)" (KCA-2011-0902-1).

References

1. Robin, G., Shaun, G., Simon, H.: Do They Want It? Do They Use It? The 'Demand-Side' of e-Government in Australia and New Zealand. Government Information Quarterly (2010)
2. Loo, W.H., Paul, H.P., Chong, S.C.: User Acceptance of Malaysian Government Multipurpose Smartcard Applications. Government Information Quarterly 26, 358–367 (2009)
3. Juan, G., Frank, W., Antonio, J.: Factors Affecting the Use of e-Government in the Telecommunications Industry of Spain. Technovation 27, 595–604 (2007)
4. Schedler, K., Lukas, S.: Customer Orientation in Electronic Government: Motives and Effects. Government Information Quarterly 24, 291–311 (2007)
5. Dijk, J., Peters, O., Wolfgang, E.: Explaining the Acceptance and Use of Government Internet Services: A Multivariate Analysis of 2006 Survey Data in the Netherlands. Government Information Quarterly 25, 379–399 (2008)
6. Lau, T.Y., Aboulhoson, M., Lin, C., Atkin, D.: Adoption of e-government in three Latin American countries: Argentina, Brazil and Mexico. Telecommunications Policy 32, 88–100 (2008)
7. Paul, T.J.: Deliberative Democracy and the Conceptual Foundations of Electronic Government. Government Information Quarterly 22, 702–719 (2005)
8. Venkatesh, V., Morris, M., Davis, G., Davis, F.: User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 27(3), 425–478 (2003)
9. Pieter, V., Gino, V.: User-centered E-Government in Practice: A Comprehensive Model for Measuring User Satisfaction. Government Information Quarterly 26, 487–497 (2009)
10. Milbrath, L.W., Goel, M.L.: Political Participation: How and Why People Get Involved in Politics. Rand McNally College Publishing Company, Chicago (1977)
11. Park, C.W.: Characteristics of the Korean National Assembly: A Critical Review. Parliamentary Research 1(1), 14–38 (1995)
12. Taylor, S., Todd, P.: Assessing IT Usage: The Role of Prior Experience. MIS Quarterly 19, 561–570 (1995)
13. Igbaria, M., Parasuraman, S., Baraoudi, J.J.: A Motivational Model of Microcomputer Usage. Journal of Management Information Systems 13(1), 127–143 (1996)
14. Bagozzi, Yi, Y.J.: Multitrait-multimethod Matrices in Consumer Research. Journal of Consumer Research 17(4), 426–439 (1991)

Analyzing User Satisfaction Factors for Instant Messenger-Based Mobile SNS*

Seongwon Park[1,3], Duckshin Oh[2], and Bong Gyou Lee[3,**]

[1] Research Institute for Information Strategy Consulting, Seoul, Korea
[2] Department of Management Information System, Sahmyook University, Seoul, Korea
[3] Graduate School of Information, Yonsei University, Seoul, Korea
{seongwon,bglee}@yonsei.ac.kr, ohds@syu.ac.kr

Abstract. Users of smart phone-based instant messenger such as KAKAO TALK, MSN messenger are rapidly increasing. However, research on user satisfaction for smartphone instant messengers has not been done at all. This study analyzed factors affecting user satisfaction by conducting a survey on 220 users of mobile messengers in smartphones. The survey results showed that self-disclosure, flow, and social presence significantly affected user satisfaction. This study will benefit researchers and contribute to industries that possess interest in factors affecting user satisfaction regarding the use of smartphone applications and mobile messengers.

Keywords: Mobile Instant Messenger, Mobile Instant Messaging, Mobile SNS.

1 Introduction

The smart phone market succeeded in achieving a high growth rate even amid a depression of the cell phone market during the 2009 global economic crisis. While the smart phone market achieved a growth rate of 43% in 2010. The market is expected to achieve a growth rate of 40% in 2013 [1]. Users of smart phone-based Social Network Service (SNS) are rapidly increasing in number along with the development of the smart phone market, and it is predicted that the number of users will continuously increase. SNS refers to the service that facilitates formation of relationships between online users that possess similar interests and provides various activities, such as managing personal connections and sharing information and contents for such relationships [2]. SNS can be largely classified into web-based social network services and Instant Messenger-based social network services [3]. This study investigated user satisfaction of SNS services by doing research on Instant Messenger users among various social network services.

As of October 2010, the number of mobile instant messenger users in Korea was 15,000,000 NateOn messenger users, 4,000,000 KAKAO TALK users, 2,700,000 MSN messenger users, 850,000 Skype users, and 100,000 Yahoo messenger users.

* This research was supported by a research grant of Seoul R&BD Program (# PA090747).
** Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 280–287, 2011.

The use of such mobile messengers is expected to continuously increase in the future [4][5]. In particular, KAKAO TALK acquired more than 10,000,000 Korean members by March 2011 just a year after its release [6]. Fig. 1 shows the ranking for free social networking software in the AppStore, in order of popularity. According to the figure, the instant messenger KAKAO TALK took 1st place in the social networking category with NateOn ranked 4th, Skype 8th, and MSN Messenger 9th thereby verifying the great popularity of instant messengers among social network services.

Fig. 1. Free software ranking for social network service in AppStore

However, in spite of the large number of smart phone SNS users, hardly any research has been done on factors affecting user satisfaction regarding smart phone SNS. Furthermore, research on user satisfaction for smart phone instant messengers has not been done at all. Such facts place a high significance on studying and analyzing factors affecting user satisfaction of smart phone SNS and instant messengers, which have recently gained great popularity. In order to achieve this, this study has conducted a survey on smart phone instant messenger users in order to research user satisfaction and investigate factors affecting user satisfaction.

2 Literature Review and Research Hypotheses

2.1 Mobile Messenger

When internet service first started facilitating inter-communication and cooperation between individuals, messenger was first introduced in 1996 by a company named Mirabilis. Mirabilis launched a service named ICQ (“I Seek You”), which enabled friends to send online messages with one another through their PC [7]. Since Messenger possesses the same meaning as Instant Messenger, the two terms are frequently used together. Messenger refers to the service for actualizing various interactive communications within cyber space, such as verifying the access of other users, or sending and receiving messages or files between individuals or groups [8]. Mobile Instant Messenger refers to instant messaging using mobile devices, such as

cell phones or PDA. It enables users to connect to the Internet through mobile devices in order to achieve interactive communication within cyber space, such as verifying access of other users, sending real-time messages, chatting one-on-one or multilaterally, or sending files [9]. Mobile Messenger is a service that has achieved improvement in terms of efficiency when compared to previous PC-based messengers as it provides communication that is more abundant and enables users to achieve quicker communication based on mobility. In addition, when compared with previous SMS or PC-based instant messaging, it can be said that Mobile Messenger allows users to abundantly express emotions in terms of a Social Presence [10].

Previous studies have observed user satisfaction as an important factor that affects the success and usage of the information system [11]. User satisfaction, which is one of the most commonly used dependent variables, is defined as the sum of attitudes or feelings regarding various factors that affect certain situations [12], and is regarded as the user's evaluative response on the information system [13]. As Mobile Messenger is a service that was also created by using information technology, an analysis on the factors affecting satisfaction of Mobile Messenger users is required in order to move along discussions on success factors and on expanding the user base of Mobile Messengers. This study established a research model shown in Fig. 2 in order to analyze the factors affecting the satisfaction of mobile messenger users.

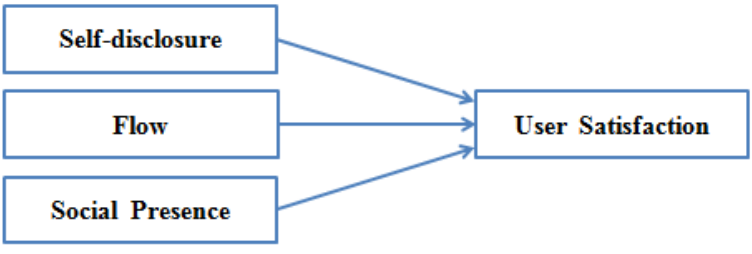


Fig. 2. Research model

2.2 Self-disclosure and User Satisfaction

Self-disclosure is defined as a process that enables others to verbally read one's private information, such as personal emotions, experiences, or opinions [14]. Self-disclosure is a important communication behaviour for establishing and maintaining human relationships [15]. Exchange of self-disclosure has been verified as an important factor in not only enhancing good feelings between two people [16], but also in forming satisfactory long-term relationships [17]. People have a tendency to develop favourable feelings toward the people they have exposed themselves [16]. Furthermore, self-disclosure not only helps establish relationships with others, but it provides a catharsis for exposing personal information [18]. Thus, the exchange of self-disclosure is regarded as an extremely significant variable in achieving satisfaction through interaction in relationships [17][19]. Schau and Gilly (2003)

stated that the concept of self-expression can be an important factor in researching SNS services, such as micro-blogs [20]. This study established the following hypothesis based on such previous studies.

H1: Self-disclosure has a positive effect on user satisfaction.

2.3 Flow and User Satisfaction

Flow refers to the state of being completely absorbed in a certain act, for example, losing the sense of time and becoming unaware of surrounding situations. Situational flow through interactive media refers to the cognitive state experienced during activity in computer-interactive environment [21]. Experienced by participants deeply related with certain situations, goals, and activities, this cognitive state refers to the status of being completely absorbed in the situation [22]. Flow is the factor that affects satisfaction and performance of activities in various areas, such as sports, leisure, sociology, and business [23]. The importance of flow has been emphasized in relation to satisfaction and performance of services using information technology. Flow enables users to discover enjoyment in the process to show more active participation in the activity to gain greater satisfaction. Various studies have verified that flow in cyber environments is a variable that significantly affects satisfaction [24]. Flow is a positive strengthener that increases future use of the information system in a computer environment [25]. This study presents the following hypothesis based on such previous studies.

H2: Flow has a positive effect on user satisfaction.

2.4 Social Presence and User Satisfaction

As a concept for explaining the psychological experience formed by interaction between communicators, social presence can be defined as the "salience of the other in communication interaction" [26]. Steuer (1992) defined social presence as the sense of being within a certain environment, stating that the sense of being in a distance place is the experience of sensing each other's existence in a certain environment through the communication medium [27]. It has been verified that social presence in cyber space can form positive relationships with the creation of results, such as satisfaction [28]. On the other hand, instant messenger has been verified to facilitate real-time dialogues to provide social presence to users, thus enabling users to forget about the computer medium and feel as if he or she is actually existing in the same space as the other user [29]. Certain studies predict that the mobile messenger will provide abundant emotional expressions in terms of the social presence, as the new mobile messenger can be accessed anywhere and anytime unlike previous SMS, PC-based instant messengers. However, this hypothesis has not been empirically proven by research results. As the appearance of smart phones have enabled users to exchange more information through mobile messengers, social presence, in relation to sharing emotions with others and experiencing co-existence with other distant users, can be regarded as a factor of greater significance for mobile messenger users. Thus, this study presents the following hypothesis.

H3: Social presence has a positive effect on user satisfaction.

3 Methods

In order to prove the research model and hypothesis, self-disclosure was defined as the act of revealing personal information, such as private emotions and opinions, and was composed of 4 items based research by Kendzierski and DeCarlo (1991). Absorption was defined as the state of being completely absorbed during use of the mobile messenger service and was composed of three items based on research by Webster and Ahuja (2006). Social presence was defined as the state of feeling as though the user has personally met the other user during the use of the mobile messenger service, and was composed of four items based on research by Dodds & Monroe (1991). Satisfaction for user experience was defined as the degree of satisfaction when using a mobile messenger and was composed of four items based on research by Cronin Jr., Brady, & Hult (2000) and Thomson (2006).

A survey was conducted on users of smart phone messengers for approximately 4 weeks in October 2010. A total of 220 surveys were collected, and 202 surveys, excluding 18 inaccurate surveys, were used in the data analysis. The 202 respondents were composed of 64 women (32%) and 138 men (68%), with 162 people in their 20s (80%) and 40 people in their 30s (20%). Also, the survey results showed that 90 people used a mobile messenger for less than 30 minutes per week (45%), 76 people used a mobile messenger for 30 min ~ 1 hour per week (38%), 30 people used a mobile messenger for 1 hour ~ 3 hours per week (15%), and 6 people used a mobile messenger for 3 ~ 5 hours per week (3%).

4 Results

Data analysis was conducted with SPSS ver. 12.0. The questionnaire asked participants to rate the extent to which they agree with each statement by circling a number from one to seven with fourteen items.

4.1 Factor Analysis

To test the construct validity of the measurement, we conducted a factor analysis on the survey questions using a principal component analysis, with a varimax rotation. Factor lading cutoff was set with value greater than 0.5. Factors with eigen value greater than 1.0 are extracted. As a result of factor analysis, two factors are exacted, named “Self-disclosure”, “Flow”, “Social presence”, and “ User satisfaction”. Results are shown in Table 2.

4.2 Reliability Analysis

A measurement instrument with a Cronbach’s α value of 0.6 or greater is generally considered satisfactory in terms of reliability. All factors show significant reliability level, ranged from .762 to .945 by Cronbach’s α in Table 1.

Table 1. Results of factor analysis and reliability analysis

Variables	Items	Factor1	Factor2	Factor3	Factor4	Cronbach's α
Self-disclosure	SEN3	.931				.945
	SEN2	.916				
	SEN1	.881				
	SEN4	.684				
Flow	FLO2		.849			.813
	FLO1		.838			
	FLO3		.705			
Social Presence	SPR3			.893		.859
	SPR4			.831		
	SPR2			.596		
User Satisfaction	SAT3				.788	.762
	SAT2				.779	

4.3 Regression Analysis

Findings of the regression analysis on the relationship between self-disclosure and user satisfaction, the relationship between flow and user satisfaction, and the relationship between social presence and user satisfaction show the significant associations, supporting the hypothesis 1, hypothesis 2, and hypothesis 3(R^2=.456). Details are shown in Table 2.

Table 2. Results of regression analysis

Variables	B	Std. Error	β	t	p	R^2
(Constant)	.855	.283		3.021	.003	.456
Self-Disclosure	.149	.066	.153	2.242**	.026	
Flow	.281	.059	.289	4.787***	.000	
Social Presence	.408	.071	.381	5.714***	.000	

5 Conclusions

This study conducted the survey on mobile messenger users to investigate the effect of self-disclosure, flow, and social presence on the user satisfaction of mobile messengers. The survey results showed that self-disclosure, flow, and social presence significantly affected user satisfaction; the influence shown in the order of magnitude from greatest to least was social presence, flow, and self-disclosure, respectively.

The theoretical contributions of this study are as follows. First, since this study conducted research on the user satisfaction of smart phone applications during a period that lacked various studies on smart phones, it can provide help to future researchers that will investigate smart phone applications. Second, since this study has clarified the factors affecting user satisfaction of mobile messengers during a period that lacked research on mobile messengers, it can contribute to future research on mobile messengers, to the satisfaction of messenger users, and to the use of mobile messengers.

The operational contributions of this study are as follows. First, since self-disclosure has been shown to be a factor affecting user satisfaction, industries should realize that the development of functions for facilitating easier self-disclosure of mobile messenger users should be considered during mobile messenger production and planning. Diversification regarding font, emoticon, or flashcon can be cited as a specific measure. Second, since absorption has been verified as a factor affecting user satisfaction, industries should realize that an environment for achieving better user absorption during the use of mobile messengers should be provided. Thus, factors hindering absorption, such as the internet cut-off phenomenon or, data transfer quality must be eliminated to maintain a stable internet status. In addition, a wide bandwidth for transmitting abundant information should be provided to users and an appropriate size for advertisements must be established to prevent ads from blocking the view of users during communications so not to disturb absorption. Third, since social presence has been presented as a factor affecting satisfaction, producers must consider functions for maximizing social presence, such as establishing the same background screen to make users feel as though they are in the same space or announcing the current status of users by displaying messages such as "the user is currently writing a message."

However, this study did not consider other important factors such as the utilitarian value, hedonic value, and functional characteristics of smart phones for mobile messengers. Future research must include the different effects on user satisfaction according to the various purposes of mobile messengers and should be done taking into consideration other important factors that influence user satisfaction, such as mobility, convenience of user interface, usefulness, dialogue transfer speed, diversity of sharable data, and financial risk, in addition to self-expression, flow, and social presence.

References

1. http://www.seri.org/db/dbReptV.html?g_menu=02&s_menu=0202&pubkey=db20100203001
2. Elkin, N.: Mobile Social Networks: Marketing by Location Shows Potential. eMarketer (2009)
3. Son, D., Park, K., Bae, S., Choi, H.: Instant Messenger System for Group Collaboration and Contents Sharing. In: Proc. of Korea Computer Conference, vol. 37(1), pp. 208–211 (2010)
4. http://www.fnnews.com/
5. http://media.daum.net/economic/industry/view.html
6. http://www.hankyung.com/news/app/newsview.php?aid=2011033071637
7. Bang, K.-C.: Intranet Based Intelligent Messenger connection research. Journal of Digital Contents Society 5(4), 283–288 (2004)
8. Park, C.E.: A Study on the Use Behavior of Instant Messaging: The Influence of Personal Origin and Social Network. MA dissertation, Yonsei University (2003)
9. http://ikis.re.kr/jsp/add/PublishGisaPreView_5.jsp?orgCode=AG&dTypeCode=10&jYSerial=2001MAG&jASerial=002&jLanguage=kr&cntFlag=2&filename=null

10. Hedbring, S.: Mobile Messaging Usability – Social and Pragmatic Aspects. Royal Institute of Technology, Sweden (2002)
11. McKinney, V., Yoon, K.-h., Zahedi, F.: The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach. Information Systems Research 13(3), 296–315 (2002)
12. Galletta, D.F., Lederer, A.L.: Some Cautions on the Measurement of User Information Satisfaction. Decision Sciences 20(3), 419–434 (1989)
13. DeLone, W.H., McLean, E.R.: Information Systems Success: The Quest for the Dependent Variable. Information Systems Research 3(1), 60–95 (1992)
14. Derlega, V.J., Metts, S., Petronio, S., Margulis, S.T.: Self –disclosure. Sage Publications, Newbury Park (1993)
15. Altmar, I., Taylor, D.A.: Social penetration: The development of interpersonal relationships. Holt, New York (1973)
16. Collins, N.L., Miller, L.C.: Self-disclosure and liking: A meta-analytic review. Psychological Bulletin 116(3), 457–475 (1994)
17. Sprecher, S., Hendrick, S.S.: Self-disclosure in intimate relationships: Associations with individual and relationship characteristics over time. Journal of Social and Clinical Psychology 23(6), 857–877 (2004)
18. Cozby, P.C.: Self-disclosure: A literature review. Psychological Bulletin 79(2), 73–91 (1973)
19. Jourard, S.M.: The transparent self. Van Nostrand Reinhold, New York (1971)
20. Schua, H.J., Gilly, M.C.: We Are What We Post? Self-Presentation in Personal Web Space. Journal of Consumer Research 30(3), 385–404 (2003)
21. Csikszentmihalyi, M.: Flow: the psychology of optimal experience. Harper Perennial, New York (1990)
22. Novak, Hoffman, D.L.: Measuring the Flow Construct in Online Environments: A Structural Modeling Approach, working paper (1999), http://www.elab.vanderbilt.edu
23. Kim, N.-Y.: The Structural Relationship among Academic Motivation, Program, Organizational Support, Interaction, Flow and Learning Outcom. In Cyber Education. Ph.D. dissertation of Dept. of Education Technology, Ewha University (2009)
24. Shin, N.: Online learner's 'flow' experience: an empirical study. British Journal of Educational Technology 37(5), 705–720 (2006)
25. Hackbarth, G., Grover, V., Yi, M.Y.: Computer Playfulness and Anxiety: Positive and Negative Mediators of the System Experience Effect on Perceived Ease of Use. Information and Management 40(3), 221–232 (2003)
26. Short, J.A., Williams, E., Christie, B.: The social psychology of telecommunications. John Wiley & Sons, Ltd., London (1976)
27. Steuer, J.: Defining virtual reality: dimensions determining telepresence. Journal of Communication 42(4), 73–93 (1992)
28. Hackman, M.Z., Walker, K.B.: Instructional communication in the televised classroom: The effects of system design and teacher immediacy on student learning and satisfaction. Communication Education 39, 196–206 (1990)
29. Hwang, H.-S.: Development of Social Presence Measurement of Mediated Social Interaction: A Case Study of Instant Messaging. Journal of Communication Science 7(1), 529–561 (2007)

Analysis of the RPS System in Korea Based on SCP Framework

Giseob Byun, Taisiya Kim, Ki Youn Kim, and Bong Gyou Lee

Graduate School of Information, Yonsei University
#410-2 New Millennium Hall, 262 Seongsan-no, Seodaemun-gu, Seoul 120-749, Korea
{tgbgs,lucky8619,seansky,bglee}@yonsei.ac.kr

Abstract. The purpose of this study is to examine why a separate quota was assigned for Photovoltaic power generation when introducing the RPS (Renewable Portfolio Standards) in Korea, as well as the implications of the new system to relevant policies. This study has examined the structure of the Korean new and renewable energy industry using the SCP(Structure-Conduct-Policy) framework. The results of the study show that though the standard currently has a high generation cost and is therefore comparatively lacking in competitive edge, the overall growth trend of the industry suggests high future potential. The study is expected to be utilized as data for other countries that are considering the application of, or are currently running RPS systems.

Keywords: ICT Policy and Services, Renewable Energy industry, Renewable Portfolio Standards, Industrial Organization, Photovoltaic power generation.

1 Introduction

Recently, energy depletion and global warming have received worldwide attention. Furthermore, as the price of oil has spiked up, major foreign countries have begun to push ahead various policies such as the FIT(Feed-In Tariff) and RPS(Renewable Portfolio Standards) systems for the development of new and renewable energy sources. In particular, the generation capacity of the Photovoltaic power generation industry has exhibited a high growth rate of 38.4% (2000~2005) due to an increase in the technical economy of the method, as well as subsidizing policies from various governments since 2000. In 2010, Photovoltaic power generation is expected to reach a generation capacity of 16.6GW, which is an expansion of 127% compared to the previous year, thus making Photovoltaic power generation the fastest growing new and renewable energy source [1, 2].

Nevertheless, Korea consumes the 9th most energy in the world, and its dependency on foreign energy is still high at 97% [3]. Also, according to the IEA (International Energy Agency), the average distribution rate of new and renewable energies for OECD countries in 2007 was 6.7%, while that for Korea is a mere 2.5%, which merits immediate attention. The Korean government has enacted the FIT system since 2002 in order to grow new and renewable energy industries. This has resulted in an annual average growth of 145% for the last five years in the domestic

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 288–295, 2011.

market for Photovoltaic power generation and the percentage of exports is rising every year with the expansion of global demands [2]. However, as the rapid building of power generators has revealed weaknesses in the system, in the forms of limited budget and lacking functionality in terms of stimulating technical development, the Korean government is scheduled to discontinue the FIT system in 2011, and introduce the RPS system on 2012 [4].

In order to introduce RPS effectively, there is a need to prepare and analyze such data as an analysis and mid-to-long-term projection of the status of relevant domestic industries [5]. Furthermore, a study on the goal-setting methods and application cases of RPS systems also states that there is a need to consider the fact that the structure and methods of the RPS system varies according to the institutions, markets, and other conditions of each country [6]. Therefore, this study attempts to analyze Korea's new and renewable energy industry using SCP(Structure-Conduct-Performance), which is an established analytical tool in industrial organization theory, and to examine the reasoning behind Korea's decision to designate a separate quota for Photovoltaic power generation in particular.

2 Theorical Background

2.1 FIT and RPS

The largest reason behind the success that major foreign countries have experienced in the development in distribution of Photovoltaic power generation technologies is because their governments have invested interest in subsidies to the growth and market creation of the Photovoltaic power generation industry [7]. This is because the costs for early management and installation are very high in new and renewable energies such as Photovoltaic power generation. Therefore, in order to supplement the lacking economic efficiency of new and renewable energies and to facilitate the distribution of power generation using those sources, developed nations in Europe and elsewhere have introduced various policies. The FIT and RPS system can be said to be the representative policies enacted for the distribution and early market creation of new and renewable energies.

Under the FIT system, when the government obligates a total purchase of the power generated with renewable energy, the operator purchases new renewable power from the generator businesses at a fixed price dictated by the policy and then transfers the remaining costs to the end customer. This is a method in which the government sets the standard price and the market sets the amount of power to be generated, and is adopted primarily by European countries such as Germany, Denmark, and the UK.

The RPS system sets an obligatory level to be supplied by new and renewable energy sources on the generator businesses, and provides RECs (Renewable Energy Certificates) to those companies corresponding to the amount that they have supplied power using new and renewable energy generation, which can be sold to companies that have fallen short of the obligatory standard, thus aiming for the distribution of new and renewable energy [8]. Korea is set to switch to the RPS system on 2012 from the existing FIT system.

2.2 SCP Framework

The SCP framework is an analytical method that explains the interrelations between achievements by analyzing industrial organizations from the three aspects of structure, conduct, and product.

Structure refers to the factors that determine a market's structural attributes, and includes the number of firms within the same industry, the distribution of company size, differentiation between products, and terms of entry. Conduct refers to behavior such as price differentiation, advertisement, and strategic responses that are influenced by market structure. The most well-known of these behaviors is collusion, which refers to several companies collectively adjusting their supply amount and thereby increasing market prices. The product of an industry is determined by how much efficiency has been achieved in resource distribution, which can be attained when considering the point at which price equals marginal cost in a market of perfect competition [9]. Recently, a newer version of the SCP framework, which is not a one-way framework in which the established market structure influences conduct, which then influences product [10], but also considers basic conditions and policies, as well as the feedback loop between structure, conduct, and product.

This study utilizes this new SCP framework which considers the influence of policies in analyzing the structure of the domestic new and renewable energy industry in order to find out the reasons why Photovoltaic power generation was imparted a distinct obligatory quota.

3 Analysis of Photovoltaic Power Generation Industry in Korea

3.1 Structure

The Korean government is executing various policies to reduce the usage of fossil fuels and to attain low carbon green growth, and is striving to raise the distribution rate of new and renewable energies, which was 2.5% as of 2009, to 11% until 2030 [11].

However, the Photovoltaic power generation industry is distinct from the rest of the new and renewable energy industry in that, it has exhibited a yearly average growth rate of 145% for the last five years and has exceeded 257MW in accumulated facility capacity as of 2008 [12]. Currently, the number of companies participating in the Korean market for Photovoltaic power generation is 260 [13] and they are active considering the size of the overall market. Although in its early stages, the Korean market for Photovoltaic power generation was formed by small companies with technological advantages entering the market with subsidies from the government, recent developments have raised the business potential of the Photovoltaic power generation industry to a much higher level [2, 14].

The technological level of Korea's new and renewable energy has been assessed as being under 50% that of developed countries, but has now reached an average of 70%. Also, technological autonomy has been attained at a percentage of 69.8% in the case of design, and 69% in production [1, 15]. In the case of Photovoltaic power generation, we have attained a technological level reaching 78.1% of that of developed countries, and have achieved a comparative advantage of 114.1 against our competitor China [16].

However, when compared to the industrial and technological level, the price of generation is high among new and renewable energy sources, being at 3~10 times that of wind energy, and also has the drawback of having a high installation price, which incurs higher initial costs [17]. Photovoltaic power generation's various strengths and its infinite availability makes it fit for expansion in its applications and gives it a high chance for attaining improved efficiency through technological innovations, which overall points to its development being highly probable [18].

3.2 Conduct

The supply percentage of new and renewable energy in Korea was 1.24% in 2001, 2.08% in 2004, and has expanded to 2.5% in 2009, and the current capacity of its facilities is at 354,861KW, with a total generation amount of 855,037MWh [12, 26]. In 2010, 937.6MW's worth of new and renewable energy facilities are being installed and operated in 1686 places, of which Photovoltaic power generation facilities comprise 450MW's worth, leading the new and renewable energy industry in terms of supply capacity [19].

Also, as the Korean market for Photovoltaic power generation is small compared to its generation capacity, 60%~70% of its revenue is being obtained from exports. Although the installed domestic Photovoltaic power generators only amount to 450MW, the production capacity of Korea's Photovoltaic power companies are estimated to exceed 700MW (Important_HanaFinance_2010). This is because, with the expansion in the global demand for Photovoltaic power generation, the percentage and amount of exports have increased annually, and the domestic sales of Korean machinery, materials, and products have been reduced to 50~60% due to increased usage of low-end Chinese imports. In 2009, the exports of Korean new and renewable energy amounted to 2 billion 4 thousand dollars, of which 1 billion 308 million dollars were accounted for by Photovoltaic power generation [20].

Korean Photovoltaic power generation market is expected to become dominated by large companies who are capable of securing surplus capital for investment in production facilities, and can also use M&A and their sub-companies to achieve vertical systemization [2]. This, in turn, is expected to raise Korea's share in the global new and renewable energy market.

3.3 Performance

The product of Korea's new and renewable energy industry, its earnings have increased from 139.4 billion won in 2004 to 4 trillion 275 billion won in 2009 thanks to government subsidization policies, and the annual rate of growth amounts to 102%. Also, the earnings of 2010 are projected to increase 100% compared to the previous year to reach 8 trillion 69.9 billion won due to the influence of the European boom in Photovoltaic power generation. Employment is expected to increase by 28% to reach 11,715 people, and civilian investment by 27% to reach 3 trillion 927 billion won [20].

A close look at the changes in earnings for each energy source shows that until 2007, the wind energy industry contributed 50% of the total sales, but as the Photovoltaic power generation industry grew rapidly after 2008, its earnings have risen from 33.2 billion won in 2004 by 72 times to 2 trillion 376.5 billion won in 2009, with the employment having risen by 33% compared to the previous year. Also,

the earnings for 2010 are expected to reach 5 trillion 373.6 billion won, which is a 126% increase compared to 2009 [21].

It was that Continuous governmental support such as the FIT system was behind this rapid growth in the new and renewable energy industry. The amount of governmental subsidy for Korea's new and renewable energy sector increased by 6.6 times from 118 billion won in 2003 to 808.4 billion won in 2010. The amount of subsidies for technological development also rose by 6.8 times from 37 billion won in 2003 to 252.8 billion own in 2010. The results of these investments can be confirmed in the rapid growth of Photovoltaic power generation, which benefited the most from the FIT system due to its higher cost of production compared to the other new and renewable energy sources [22].

3.4 Policy

The Korean government first suggested a detailed goal for the development of new and renewable energy when it released the second-stage basic plan for new and renewable energy in 2003. The basic direction of the policy involves expanding the new and renewable energy market to induce civilian investment, which is being realized with such direct subsidizing projects as the 100 thousand house distribution plan for Photovoltaic power generation, distribution subsidy plans, and regional distribution projects, as well as with policies for motivating investment such as loan support and the FIT system, and also by obligating the usage of new and renewable power in public facilities, and with the Renewable Portfolio Agreement [1,23,24,25]. In particular, it has enacted the FIT system focusing on electric energy, which contributed to the establishment of the market and institutions for new and renewable energy during its conception. However, there were also drawbacks to this intervention in that the insufficient preparation of institutions caused a bias towards certain energy sources.

Due to this, the Korean government disclosed a plan for replacing the current FIT system with a RPS system effective 2012. While the FIT system was a distribution system in which the government ensured that the market price stayed at a certain level for a certain period of time, the RPS system is different in that the government obligates the supplier and creates markets, but leaves the price to be determined by market forces. The RPS system to be introduced obligates generator companies with facility capacities of over 500MW, which collectively generate 98.7% of the country's power, to generate a certain percentage of their energy using new and renewable energy sources. Also, Korean government has come up with an incentive system rewarding the usage of Photovoltaic power generation, which has a high cost of generation. According to this incentive system, the obligated generator companies must newly distribute 200~300MW of Photovoltaic power generation facilities annually for the first five years of RPS, with the goal of achieving 1,200MW, which is the goal until 2022, within those five years. Considering the fact that in 2004~2009, the domestic capacity for Photovoltaic power generation distribution has increased by 90MW annually supported by government subsidies, the increase in Photovoltaic power generation distribution between 2012~2016 when the incentive system will be in effect is expected to remain at two times the current rate. This is considered to be a means for supporting the industry for more rapid technological development [2, 11].

4 Result of Analysis

The results of the analysis of the new and renewable energy using the SCP framework are as follows.

First, from the market structure aspect, it was shown that Photovoltaic power generation is in a stage of growth where many smaller companies are participating in the market and forming a structure of competition. Also, while the technical level of the other new and renewable energy sources were at 70% of that of developed countries, Photovoltaic power generation was at a higher 78.1%. However, in terms of costs of generation, Photovoltaic power generation was still higher than the other new and renewable energy sources, which implies a need for a separate subsidy policy under the RPS system as well.

In terms of conduct, the percentage of Korea's Photovoltaic power generation industry against the production and export of the whole of its new and renewable energy industry was found to be high, and unlike the other new and renewable energy industries which are in their beginning stages, Photovoltaic power generation was shown to be attaining vertical systemization with the entrance of large companies into the market.

In terms of product, the results were not very different from the above in that the percentage of Photovoltaic power generation amongst the total 4 trillion 27.5 billion won worth of earnings of the new and renewable energy industry was high, and the increase in employment was thus also high.

In terms of the government policies, it was found that the Korean government had run separate operations for the distribution of Photovoltaic power generation specifically among its many policies for the growth of new and renewable energies. The results of these analyses show why the government assigned a separate quota for Photovoltaic power generation in introducing the RPS system. The separate standard is a shield against the projected reduction in investment for Photovoltaic power generation due to its higher costs of generation once the RPS system goes into effect, and also a means for focusing on the growth of Photovoltaic power generation for its own merits that are being shown currently through the rapidness of its growth in terms of technology and industry.

5 Conclusion

Currently, Korea is lacking in new and renewable energy reserves, technological achievements compared to developed nations, and domestic infrastructure of related industries. Also, Korea lacks the production capability for key facilities in fuel batteries, Photovoltaic and wind power, which has led to higher dependency on imported products [23].

Recently, the Korean government announced that it would replace the FIT system which had been in effect since 2002 with the RPS system effective 2012. The two systems each have their ups and downs and one cannot be definitively said to be the more effective. However, it is somewhat established that FIT is more fit for growing markets in their initial stages, while RPS may be more effective once the market gains some measure of maturity in that it induces competition. However, as the government

has made clear its intentions of executing the RPS systems through orders involving the assignment of a separate quota for Photovoltaic power generation, there is a need for researches assessing the appropriateness of this decision.

Therefore, this study has analyzed the new and renewable energy industry of Korea and investigated the reasons behind the government's decision to assign a separate quota for Photovoltaic power generation. The results show that the decision is both for the protection of Photovoltaic power generation, which currently lacks competitive edge due to its higher generation costs and also because of its high potential for future growth. Although the RPS system in this form seems appropriate for Korea's current new and renewable energy industry, there is a need for more detailed and specific analyses on the introduction of the RPS system when considering the cases of foreign countries that have failed in their attempts to utilize the RPS system.

The limits of this study lie in the fact that it has used data released by the government and other research institutions due to the difficulties in obtaining clear data reaching to the present time for statistics on new and renewable energy. Also, it is lacking in that it does not provide a comparison between foreign cases with Korea's version of the RPS system. Therefore, additional research is needed. This study will be able to be used as basic data for countries considering introducing, or currently operating the RPS system.

Acknowledgement

This work was supported by the Human Resources Development of the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Ministry of Knowledge Economy, Republic of Korea (No. 20094010100130).

References

1. Lee, M.S.: Global Trends and Implications in the Photovoltaic Power Generation Industry. KDB Research Institute's Facility Investment Team (2008)
2. Ah, H.Y.: Domestic and International Photovoltaic Power Generation Market Forecast. Industry Research 16 (2010); Hana Institute of Finance
3. British Petroleum: BP Statistical Review of World Energy 2009 (2009)
4. Lee, M.S.: Comparison of the FIT and RPS Systems. KDB Research Institute's Facility Investment Team (2009)
5. Jang, H.S., Choi, K.R., Kim, S.D.: Research on the Implications of Domestic Adoption of the RPS System for Renewable Energy. Energy Engineering 14(2) (2005)
6. Jo, I.S.: Methods and Issues on Shaping the Goal of Renewable Energy Portfolio. Korean Institute of Electric Engineers, Society of Electrical Technology, pp. 264–266 (2007)
7. Park, K.E., Kang, K.H., Kim, H.I.: Status of Photovoltaics in the World, Korean Institute of Electric Engineers. In: Summer Conference Journal, pp. 1275–1276 (2010)
8. Korea Power Exchange, Outline of the RPS System and Plan for its Adoption (2008)
9. Jung, G.Y.: Industrial Organization. Parkyoungsa, 1–10 (2010)
10. Bain, J.S.: Industrial Organization, p. 79. John Wiley & Son, Chichester (1959)

11. Ministry of Knowledge Economy, Republic of Korea: Revision of the Renewable Energy Promotion Legislature Regarding a Detailed Plan for RPS Adoption, Press Release (2010)
12. Renewable Energy Center: 2009 Statistics on Supply of New and Renewable Energy, Korea Energy Management Corporation (2010)
13. Shin, Y.S.: Present Situation and Utilization-Rate of a Mass Solar-Powerplant in Korea, Korean Institute of Electric Engineers. Society of Electrical Technology Fall Conference Journal (2008)
14. Lee, M.S.: Changes in the Solar Battery Industry and Policy Trends in Key Countries. KDB Research Institute (2009)
15. Lee, C.H.: Domestic Establishment of Renewable Energy RPS System Operations, Korea Electro technology Research Institute, Ministry of Knowledge Economy, Republic of Korea (2009)
16. Kim, S.W.: An Analysis on the Actual Status of SMEs and Supply Chain in Renewable Energy Sectors. Korea Small Business Institute (2010)
17. Sung, D.W.: Trends and Forecasts of the Photovoltaic Power Generation Industry. Korea Exim Bank Overseas Economic Research Institute, Issue Briefing (2009)
18. The Federation of Korea Industries, Policies Regarding the Photovoltaic Industry in Key Countries and their Implications, FKI report, vol. 169 (2010)
19. Lee, C.H.: Review of the Detailed Plan for Renewable Energy RPS System, Presented Material from the Public Hearing. Korea Electro Technology Research Institute (2010)
20. Ministry of Knowledge Economy, Republic of Korea: Results of a Total Inspection on 150 Businesses in the Renewable Energy Industry, Press Release (2010)
21. Lee, S.B.: Sales in the Renewable Energy Industry Increased by 30 Times in the Past 5 Years. Energy & Environment News, Web site: http://www.e2news.com/news/articleView.html?idxno=36484 (retrieved January 7, 2011)
22. Jang, H.J.: Head of Green Growth of Korea Mentions FIT, The Energytimes, Web site: http://www.energytimes.kr/news/articleView.html?idxno=5423# (retrieved January 7, 2011)
23. Kim, H.J.: Analysis on the Effects of the Renewable Energy RPS System on Domestic Industries. Korea Energy Economics Institute (2009)
24. Yun, S.Y.: Status and Forecast of the Renewable Energy Industry. The Korea Gas Union Journal, 20–26 (2007)
25. Maeng, S.J.: Photovoltaic Power Generation Subsidization in Korea, FSB Research Center. Future Strategy & Business Development, 22–25 (2009)

A Design of Short-Term Load Forecasting Structure Based on ARIMA Using Load Pattern Classification

Kyoung-Mi Im[1] and Jae-Hyun Lim[2]

[1] Green Home Energy Technology Research Center, Kongju National University
[2] Computer Science Department, Kongju National University
Budeadong 275, Cheonan, Chungnam, Korea
{omnibus,defacto}@kongju.ac.kr

Abstract. This paper suggests a short-term electricity load prediction structure for collection of most efficient energy operating plan by predicting the next day's demand from separately measured past electricity demand. The present short term load prediction methods predict the next day's electricity expenditure using a probability method, time-series method or a nerve network model without separation between similar data in a traditional way or uses a standard method of season separation or day separation. However if data separation is not done, fluctuating parameter that affects load like season and day cannot be responded to and in the case of the latter there is a problem caused due to the recent increased fluctuation of climate data - separation without similarities arises and so the prediction rate decreases. This paper separated collected electricity expenditure into months, holidays, Monday, Saturday and weekdays by high similarities and short term electricity load prediction is carried out based on time-series prediction method ARIMA. Also, to evaluate the accuracy of the prediction simulation will be done based on the collected data.

Keywords: Short-Term Load, ARIMA, Pattern Classification, Power Load.

1 Introduction

Recently, due to the increase in the importance of climatic environment problem due to global warming, energy management research to decrease production of greenhouse gases is being actively carried out.

For this, research is being carried out to establish a most efficient energy operating plan by predicting production due to electricity expenditure. In electricity load prediction research there are long term load prediction research about annual maximum electricity load and hourly or daily electricity load, and the accuracy of load prediction especially has a close relationship with the reduction of cost of generation and amount of green house gases production. If the load prediction amount is bigger than the actual load amount unnecessary electricity cost is created and amount of green house gases produced will also increase as a result of this, and if it is small the driving reserve is small and the possibility of a power cut will arise [1].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 296–303, 2011.

In the short term load prediction method used until now there are prediction method using data very similar to simple prediction method types and more. In simple prediction method there are 3 types of methods using regression analysis, time series analysis and ANN, and in the case of the first regression analysis many variables such as recent electricity load status and weather are reflected to consider the factors affecting load fluctuation but there is a negative factor to it that the analysing process is quite difficult and that editing of variables depending on the fluctuation of load is necessary. Secondly, time series analysis is a method that predicts load based on past data collected where ARIMA(AutoRegressive Integrated Moving Average) analysis and analysis using ESM(Exponential Smoothing Method) is used. However, for these types of time series analysis, it is hard to change many variables that affect the characteristics of load and if the model's degree increases, the reliability of the prediction decreases. Finally prediction method using Neural Networks is a method to solve nonlinear functions and because of the clear model, easy realization and good performance it is currently receiving the most attention. But in the case of Neural Networks if the filtering of collected data is not properly carried out the correct predicted value cannot be obtained and because of this the need for deletion of irregular load data arises.

This paper suggests a short term load prediction structure for the most energy efficient operation from collected electricity data. Here, the collected electricity data uses BP(Back-Propagation) algorithm in advance and separates by months of high similarities and after weekly analysis again, separated by data of high similarities. The sorted data is predicted using ARIMA,. a time series model and is judged in comparison with the present short term electricity load prediction method.

2 Related Works

2.1 ARIMA Model

Observed data can be formed into a model by discovering a rule and changing it from the past pattern of the assumption that past patterns will not change and be repetitively continued [4]. The ARIMA model deletes trend or seasonality if the time-series data has either of these and changes them to stationary and chooses the most appropriate model from AR; autoregressive, MA; moving average or ARMA; autoregressive moving average. Formula 1 shows a AR model that predicts the future based on past observed values and Formula 2 shows a MA model that predicts the future based on past erroneous values [7].

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots + \Phi_p y_{t-p} + \varepsilon_t \tag{1}$$

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \tag{2}$$

The ARIMA model predicts the future for both past observed values and past erroneous values and is denoted like Formula 3.

$$y_t = \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \tag{3}$$

3 Short-Term Electricity Load Forecasting Structure

The short-term electricity load prediction system suggested in this paper separates electricity data collected from each substation into seasons and days using ANN's BP algorithm. The separated data predicts daily electricity expenditure using ARIMA. [3]

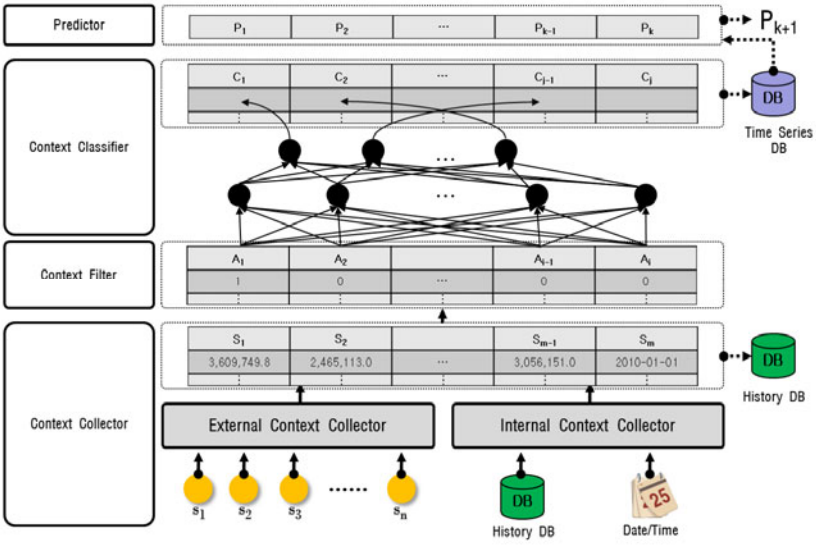


Fig. 1. Short-term load forecasting structure

3.1 System Structure

Fig.1 shows the suggested short term electricity load prediction structure and each step is as follows.

- Context Collector: Context Collector is a step necessary for electricity prediction and it is separated into External Context Collector which is collected from physical sensor according to collection route and Internal Context Collector which uses an external source. External Context Collector collects daily electricity data from sensors installed in each substation and Internal Context Collector collects date/time data collected from history database and system which were data already collected before. The raw data collected is stored in the history database.
- Context Filter: Context Filter is a step that processes data extracted from Context Collector and addition of electricity expenditure input, date data and other types of data are produced.
- Context Classifier: Context Classifier is a step that separates data processed in the Context Filter by similarity and in this paper, ANN's BP algorithm is used to separate by months of similar electricity expenditure and the sorted data is separated into days.
- Predictor: Predictor extracts data from time-series database according to the following day's data type and predicts electricity load amount due to ARIMA.

3.2 Separation of Data

If the pattern of the electricity data measured in the past is looked into, it can be known that electricity expenditure shows different flows according to the seasons or days of the week [5]. This paper aims to separate data before prediction according to seasons and then days.

3.2.1 Seasonal Separation

Fig.2 is a graph showing the electricity usage of Wednesdays during January-December of 2009 and it can be seen that the graph appears differently according to the months. Also, depending on the usage of air conditioning, which makes up 55% of the building's total electricity usage, electricity usage during summer and winter increases.

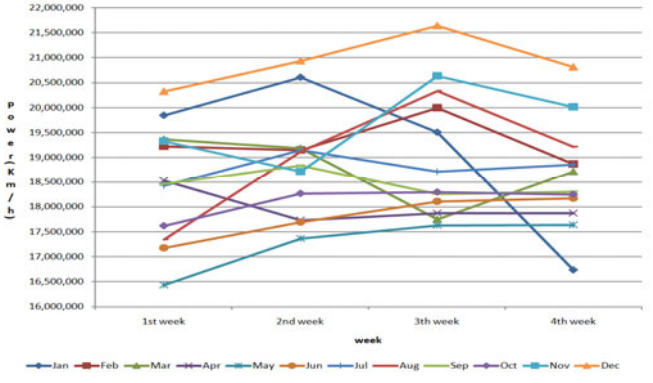


Fig. 2. 2009 Weekday(Wednesday) electricity usage

In the case where there are four seasons like in Korea, the electricity expenditure is not constant. If monthly average electricity usage and monthly average temperature during 2008-2009 are looked at, it can be seen that January-February and November-December, March and July-September, and April to June and October have similar patterns. Thereby this paper does not separate into the usual seasonal groups, spring, summer, autumn and winter, but into three types.

3.2.2 Daily Separation

Fig. 3 is a graph showing the daily electricity expenditure of the eight weeks during 5th January to 1st of March of 2009 and although there was a small change in the weekly electricity expenditure, it could be seen the Monday's load pattern was somewhat decreasing compared to other weekdays, and Tuesdays, Wednesdays, Thursdays, Fridays' load pattern, Saturdays' load pattern and Sundays' load pattern are similar weekly. Graph 4 has a completely different load pattern as 2009 January 26th and January 27th of the fourth week which were a Monday and a Tuesday, and were holidays unlike the other weeks. Also Fig. 4 is a graph of electricity usage from May

4th to May 31st of 2009 and in the case of Tuesday May 5th, a different pattern to weeks 2-4 is shown as it was a holiday, and in the case of Monday May 4th of 2009, it was a sandwich day between two holidays and so shows a different load pattern. Therefore in this paper even if it is a weekday, in the case of sandwich days between two holidays, it is regarded as a holiday for the prediction of electricity usage.

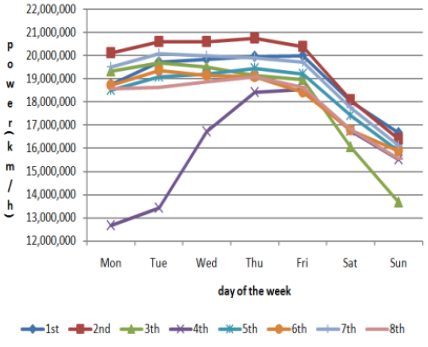


Fig. 3. Weekly electricity usage pattern

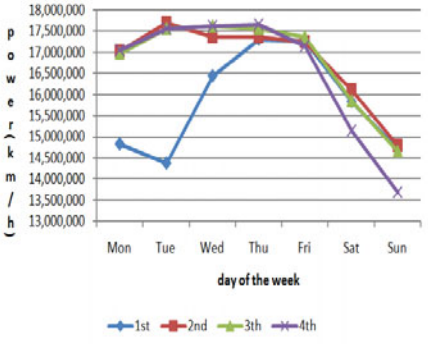


Fig. 4. Sandwich date's total electricity Usage pattern

4 Prediction Structure's Simulation and Evaluation

This paper was tested using 731 daily electricity usage measured from 2008 January 1st-2009 December 31st from 11 branches of substations situated in Cheonan, Korea. Of these, 366 of the data were used as data for education and 365 were used as data for evaluation.

4.1 Data Separation Using BP

BP algorithm is one type of neural network and prediction algorithms and is made up of MLP(Multi-Layer Perceptron), has a three steps structure where it has hidden layer between input layer and output layer [2].

The turning step suggests input pattern of neural network and uses input function and activation function on each node and produces a calculation. Activation function used in calculation of connective weighting value uses a sigmoid function like Formula 4.

$$\log sig(x) = \frac{1}{1+\exp^{-x}} \tag{4}$$

Internal context information and external context information collected from Context Collector are separated seasonally and daily like in Table 1. Here, 3 stepped MLP network made up of 9 input layers, 7 hidden layers and 4 output layers, contingent upon seasons, days and sandwich days, and each type is formed into data like Table 2.

Table 1. Classification of date

Input Pattern	Standard of Classification	
Season	Standard	Month
	1	Jan, Feb, Nov, Dec
	2	Mar, Jul, Aug, Sep
	3	Apr, May, Jun, Oct
Day of the Week	Standard	Day of the week
	1	Mon
	2	Weekday (Tue, Wed, Thu, Fri)
	3	Sat
	4	Sun

Table 2. Desired Output

Type	Desired Output	Type	Desired Output
1	0001	7	0111
2	0010	8	1000
3	0011	9	1001
4	0100	10	1010
5	0101	11	1011
6	0110	12	1100

4.2 Forecasting Using ARIMA Model

Data separated from Context Classifier predicts the next day's electricity expenditure during the predictor process.

Fig. 5 only shows data of weekdays during November-February out of the classified electricity data as a time-series graph and it can be seen that the data is non-stationary using the reference line that shows the average as a standard. By defining variable w_t and through $y_t - y_{t-1}$ primary differencing is carried out and is converted into stationary time-series data. Fig. 6 shows electricity expenditure that has been turned stationary through primary differencing.

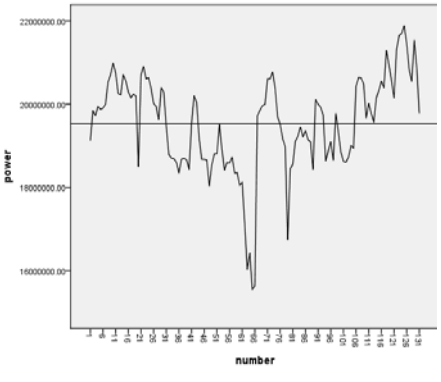


Fig. 5. Actual electricity expenditure of weekdays of Nov-Feb

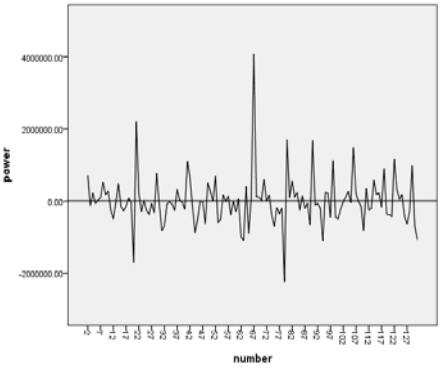


Fig. 6. Electricity expendirue of weekdays of Nov-Feb after primary differencing

The degree with the optimum auto-correlation is found by comparing auto-correlation function and partial auto-correlation function based on differed load and by changing the degree. ARIMA (3,2,0) is established and predicted as auto-correlation after secondary differencing was judged to be the optimum model than auto-correlation after primary differencing, like shown in Fig. 7.

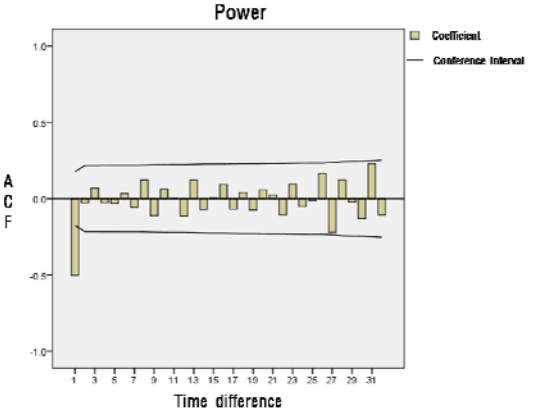


Fig. 7. Auto-correlation function that has been secondary differncing

Table 3 shows the results of the value predicted with ARIMA model without data separation from the actual data, the value predicted with the suggested model and the prediction error value, of the weekday data during November-February. The predicted error value of traditional ARIMA model and the suggested model was found to be 4.19% and 3.19% respectively and through this it could be known that the error value of the suggested model is smaller.

Table 3. Comparison of daily forecasting data

Raw Data	ARIMA	Proposed Method
…	…	…
15635915.4	12882431.96	14858824.73
19722010.8	21190380.43	15305717.57
19850218.6	19883335.95	20389127.51
19964274.7	20944088.19	20662506.9
19996143.5	21025924.81	21339272.28
20604506.3	21476965.31	21002502.13
20605003.7	19874221.78	20790594.68
20772173.4	21397974.64	20772404.31
20386530.4	21669121.42	20995520.53
19687992.1	20241005.39	20472524.72
…	…	…
Mean Error Rate	4.19%	3.19%

Table 4. Comparison of Holiday forecasting data

Date	ARIMA	Proposed Method
…	…	…
04/19	14337850.1	14983795.93
04/26	14545223.22	15307169.35
05/02	14699473.9	14015259.56
05/03	13896395.63	13608113.02
05/04	13399840.58	12583624.55
05/05	14646105.82	15035415.13
05/10	14405875.13	15709795.6
05/17	14963438.56	15413539.78
05/24	15287548.55	15025810.52
05/31	14690082.37	14094657.78
…	…	…
Mean Error Rate	5.58%	5.33%

In the case of holidays, the result of prediction based on the data of April-June and October of 2009 showed a format like in Table 4, and the result of prediction error value calculation the traditional ARIMA model and suggested model each showed

5.58% and 5.33% respectively. Especially for a sandwich day like the May 4th the traditional ARIMA model and suggested model gave 15.22% and 9.73% of predicted error values respectively, and therefore this shows that classifying electricity data between two holidays as a holiday produces a better prediction result.

5 Conclusion

This paper suggests a structure that separates data of high similarities before prediction to increase the prediction rate of the time-series prediction method, ARIMA model which is one of the methods used in present electricity load prediction. Classification of data was carried out through ANN's BP algorithm based on data collected from each substation. Electricity usage pattern was separated into three types based on months of similar electricity usage for the seasonal separation of data, and this was separated again according to days of the week. into Mondays, Weekdays, Saturdays and holidays. Also, a sandwich day which may be a weekday but in between two holidays, was classified in the same way as a holiday and the prediction rate was improved. The suggested model carried out the simulation based on the collected data and made comparisons with the traditional ARIMA model and confirmed the superiority of the structure.

In the case of seasonal separation carried out in this paper it could be seen that it has a close relationship with temperature and in the future, electricity load prediction system is planned to be researched based on the measuring of parameters that have context with electricity data usage.

References

1. Jun-Ho, P., Gill-Jae, S., Hwa-Suk, L.: A study on the Electrical Load Pattern Classification and Forecasting using Neural Network. The Transactions of the Korean Institute of Electrical Engineers, 39–42 (1991)
2. Gwi-Im, J., Sang-Sung, P., Young-Geun, S.: Construction of Personalized Recommendation System Based on Back Propagation Neural Network. The Korea Contents Society, Journal of the Korea Contents Association 7(12), 292–302 (2007)
3. Kyoung-mi, I., Jae-hyun, L.: A Design of Short-Term Load Forecasting Structure using Data Classification. Korean Society For Internet Information, 69–73 (2010)
4. Dong-Bin, J.: Demand Forecasting of Time Series, Hnnnarae (2009)
5. Sang-Yule, C., Hyoung-Joong, K.: Short-term demand forecasting Using Data Mining Method. The Korean Institute of Illuminating and Electrical Installation Engineers, Journal of the Korean Institute of Illuminating and Electrical Installation Engineers 21(10), 126–133 (2007)
6. Lincoln, W.P., Skrzypek, J.: Synergy of clustering multiple back-propagation networks. In: Touretzky, D.S. (ed.) Advances in Neural Information Processing Systems, vol. 1, pp. 650–657. Morgan Kauffman, San Mateo (1990)
7. Wang, X., Liu, Y.: ARIMA Time Series Application to Employment Forecasting. In: Proceedings of 2009 4th International Conference on Computer Science & Education.

Comparative Analysis of Learning Effect on Lexical Recognition in the e-Learning and s-Learning*

JaeChoon Jo and HeuiSeok Lim

Dept. of Computer Science Education, Korea University
{jaechoon,limhseok}@korea.ac.kr

Abstract. The purpose of this study is to analyze learning effects of e-learning and s-learning on learner's brain. The subjects of this study were twenty elementary school students. The experimental group was divided into two groups and both groups learned the same content. As a result of the research, both groups showed improvement. But s-learning is higher than e-learning in the effect of learning based on the brain.

Keywords: e-Learning, s-Learning, Smart Learning, Brain cognition, cognition ability.

1 Introduction

An advance in information and information communication technologies to the mobile environment has evolved from the Internet environment. Accordingly the educational environment has changed from e-learning to m-learning. Recently Smartphones have attracted a lot of interest to the learner and they have been applied to education. We called s-learning that education used in smartphone or smart device. Furthermore s-learning has been continuous research and development in ubiquitous environment.

According to the rapid development of information communication technology, the field of teaching is researching and developing to create a learning environment containing new information communication technology. Among them e-learning showed the positive effects of education for long time and constant e-learning has been developed. m-learning and s-learning have also shown a positive effect of education in addition to e-Learning.

Thus in this study, we will analyse e-learning and s-learning and how they show a positive learning effect through the Lexical Decision Task (LDT) based on the brain.

LDT is a procedure for measuring how quickly people classify proposed contents as correct or incorrect word and the reaction times and the error rates in the various conditions of presented words are measured for proficiency [1].

We will use a LDT Test for measuring a learner's proficiency.

The present paper takes it as a focal point of the learning effect based on brain science.

* This research is supported by the KOCCA(Korea Creative Content Agency) funded by the Korean Government [2-10-7602-001-10748-07-003, Technology of the Emotion Measurement and Standardization on Functional Contents].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 304–308, 2011.

2 Background

The traditional education is made in classrooms where the teacher presents the learning material to a group of students. The educational technology depends mainly on the teacher and the students must physically participate in the learning process. Regardless of the obvious advantages with direct contact between teacher and students and also the immediate feedback the traditional classroom setting has many disadvantages [2].

E-learning, m-learning, s-learning and u-learning are methods of learning and utilizing information technology to overcome disadvantages of traditional teaching methods with the latest information technology to improve the efficiency of learning.

2.1 e-Learning

We will call e-Learning all forms of electronic supported learning and teaching. These are procedural in character and aim to effect the construction of knowledge with reference to individual experience, practice and knowledge of the learner. Information and communication systems, whether networked or not, serve as specific media to implement the learning process [3].

Currently e-learning is used in learning of formal as well as in formal.

Traditional e-learning was implemented to provide simply content of education to students from education of environment offline to education of environment online.

Nowadays, learners are receiving motivation, Q&A, interaction and many methods of teaching by e-learning by advancement of a lot of e-learning.

Experimental content of e-learning was developed by the traditional way of e-learning that simply provides content of learning to students.

In this study, developed experiment content was used in experiment of e-learning.

2.2 s-Learning

We called s-learning when learning by smartphone. That is not all; s-learning beyond learning with a smartphone is using the smart device between learners. Content is a feature that allows ICT skills that communication, collaboration and participation. Also it is a holistic approach that horizontal, interactive, participatory, intelligence, and an interactive way to increase the learning effect.

In this study, experimental content of s-learning was used in s-learning of suitable characteristics.

2.3 Other Learning

A ubiquitous learning environment is any setting in which students can become totally immersed in the learning process. So, a ubiquitous learning environment is a situation or setting of pervasive or omnipresent education or learning. Education is happening all around the student but the student may not even be conscious of the learning process. Source data is present in the embedded objects and students do not have to DO anything in order to learn. They just have to be there [4].

The term M-Learning, or "mobile learning", has different meanings for different communities. Although related to e-learning and distance education, it is distinct in its focus on learning across contexts and learning with mobile devices. One definition of mobile learning is: Any sort of learning that happens when the learner is not at a fixed, predetermined location, or learning that happens when the learner takes advantage of the learning opportunities offered by mobile technologies. In other words mobile learning decreases limitation of learning location with the mobility of general portable devices [5].

All methods of learning were advanced from learning offline or traditional to e-learning online environment through development in information communication technology and each learning is related in characteristic.

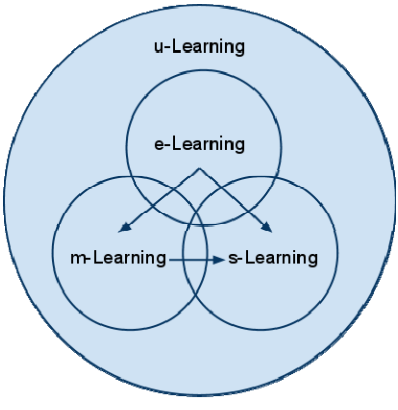


Fig. 1. Domain and relation of Learning

3 Experimental Method

The subjects of this study were 20 elementary school students. The students were divided into two groups and given the same learning content. One group, the e-learning group, consisted of ten students and second group, the s-learning group, also consisted of ten. The e-learning group used content in e-learning developed for this study and s-learning group used content in s-learning, downloaded from the app store. The learning times were exactly the same.

Each participant had a pre-written test and were pre-tested before the experiment. Experiment (learning - Evaluation) was tested twice.

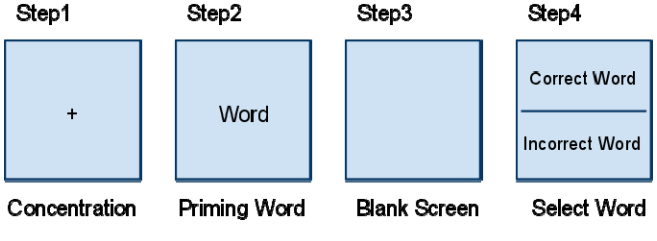


Fig. 2. Test Process

As shown in Figure 2, the test consists of four steps; all steps were repeated according to the number of stimulus words. In the first step, a small black cross is presented in the center of the screen for 150ms to induce the learner's concentration. The second step presents a word as a primer for 250ms; the third step presents a blank for 150ms. Lastly, in the fourth step, a target word is presented and the learner decides whether it is a correct word or incorrect word by using the click.

4 Result

The results of the experiment were positive. Remembrance of the words in both groups showed the rates rise. The e-learning group improved from 51 points to 83 points and the s-learning group improved from 58 points to 85 points.

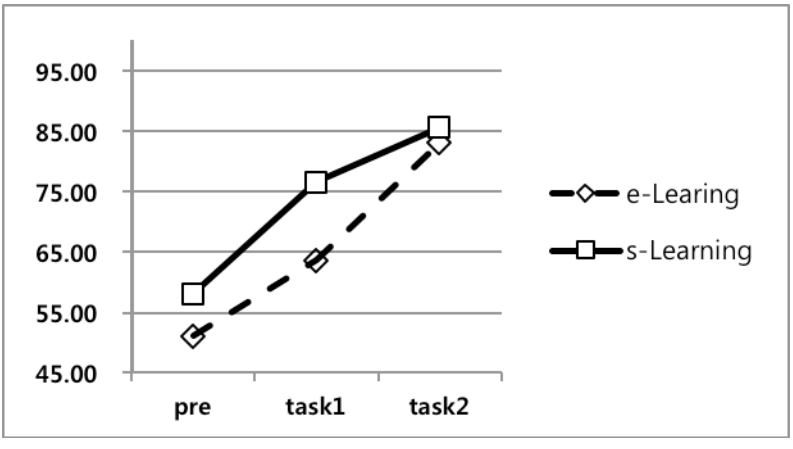


Fig. 3. Effect of Learning

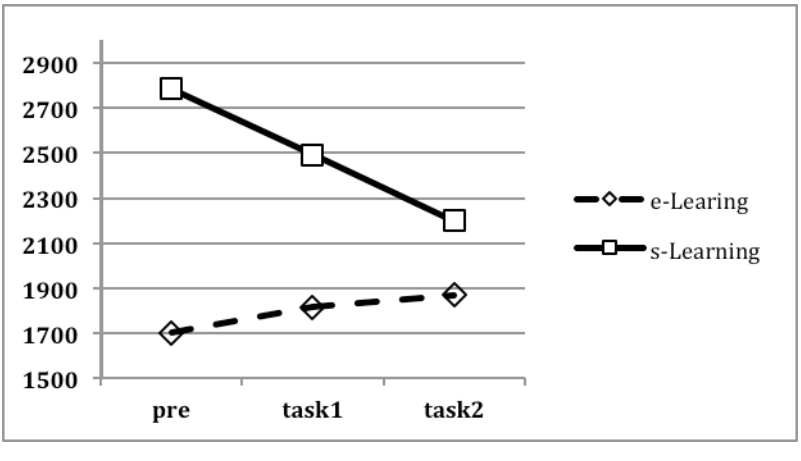


Fig. 4. Response of Cognitive Ability

As shown in Figure 3, a change in the average scores was shown. Both groups had risen their average score and there was no significant difference between the two methods of learning.

Shown as in Figure 4, evaluation of the cognitive responses based on the brain showed significant levels on .05 in the learning effect between the two groups. While the e-learning group was constant in response rates between tasks, the s-learning group showed decreasing response rates between tasks. Therefore, we have found the s-learning to be higher than e-learning in the effects of learning based on Brain Science.

5 Conclusion

According to the rapid development of information communication technology, the field of teaching is researching and developing to create a better learning environment to ensure a self-directed learning environment and to improve learning effect. The learning environment showed the positive effects of education. m-learning, s-learning and u-learning has also shown a positive effect of education in addition to e-learning.

In this study, we analysed e-learning and s-learning in showing a positive learning effect through responses of cognitive ability based on the brain.

As a result of research, both groups showed a positive effect of learning. But the s-learning was higher than e-learning in the effects of learning based on brain.

Therefore, we have found s-learning higher than e-learning in the effects of learning based on Brain Science because s-learning is concentration from interesting and interacting.

In this study, e-learning was used in content of the traditional way, whereas s-learning was used in content of interaction.

But the current e-learning had a lot of technological advances. So we can expect the effect to be higher.

References

1. Kwon, D.-Y., Lim, H.-S., Lee, W., Kim, H.-C., Jung, S., Suh, T., Nam, K.: A personalized English vocabulary learning system based on cognitive abilities related to foreign language proficiency. KSII Transactions on Internet and Information Systems 4(4) (August 2010)
2. Georgiev, T., Georgieva, E., Smrikarov, A.: M-Learning: a New Stage of E-Learning
3. Tavangarian, D., Leypold, M., Nölting, K., Röser, M.: Is e-learning the Solution for Individual Learning? Journal of e-learning (2004)
4. wikipedia, http://en.wikipedia.org/wiki/Ubiquitous_learning
5. wikipedia, http://en.wikipedia.org/wiki/M-learning
6. Lee, S., Lee, W., Kim, H.-C., Jung, S.-Y., Lim, H.: A Personalized CALL System Considering Users Cognitive Abilities. In: Taniar, D., Gervasi, O., Murgante, B., Pardede, E., Apduhan, B.O. (eds.) ICCSA 2010. LNCS, vol. 6019, pp. 376–384. Springer, Heidelberg (2010)
7. Park, C., Nam, s.: A Study on U-Learning

Predicting Text Entry for Brain-Computer Interface

Saebyeok Lee and Heui-Seok Lim

Computer Science Education, Korea University
Seoul, South Korea
saebyeok@blp.korea.ac.kr,
limhseok@korea.ac.kr

Abstract. A brain-computer interface (BCI) is a communication system that translates brain activity into commands for a computer or other devices. In other words, BCIs create a new communication channel between the brain and an output device by bypassing conventional motor output pathways consisting of nerve and muscles. This is particularly useful for facilitating communication for people are totally paralyzed. However, it is hard to practical use because low communication rate. Therefore, in this paper, we propose a method to speed up text entry in BCI using a statistical language model.

Keywords: Brain-Computer Interface, Predicting Text Entry.

1 Introduction

Brain-computer interface (BCI) is a new communication channel in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerve and muscles [1]. In order to detect the brain's operation, it is used to recording methods such as electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). However, MEG, PET, fMRI are still technically demanding and expensive, so that EEG used generally recording method, which have relatively short time constants, and require relatively simple and inexpensive equipment.

The input methods in EEG based BCI systems depend on measurable mental states but it is very limited. Generally, it is used to evoked potential or mu- and beta-rhythms. The evoked potentials involve two states: frequent or infrequent features. The mu- and beta-rhythm is able to involve more various states: left or right hand movement (or movement imagery), foot movement, tongue movement. However, if systems use a lot of mental states, it causes row accuracy of classification [1]. Though these limited input methods, BCI is helpful to be able to provide communication channel to people who are totally paralyzed such as amyotrophic lateral sclerosis (ALS), brainstem stroke, spinal cord injury and muscular dystrophies.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 309–312, 2011.

Actually, BCI based communication system such as P300 speller [2] and hex-o-spell [3] is used often for conveying messages and commands to the external world to these disorders. Typically, research into BCI is important to measure various mental states correctly. Therefore, much of the research on BCI has focused on signal processing or classification. However, it is possible to improve communication rate though the efficient interface that involve the smart or intelligence interface. Therefore, in this paper, we propose predicting text entry interface that often used in mobile phone in order to improve the communication rate.

2 Methodology

Limited interface in the mobile was successfully applied to use word prediction method [4]. Word prediction method is to predict by button array that user press the buttons so that a button which have more 3 characters only press one time. The probability equation of these processes is as follows:

$$\begin{aligned} \mathrm{W}(b) &= \mathrm{argmax}_{x\in X} P(w_x|b) = \mathrm{argmax}_{x\in X} P(w_x)P(b|w_x) \\ &= \mathrm{argmax}_{x\in X} P(w_x) \end{aligned} \tag{1}$$

A set of X is writable words given inputted button array b. $\mathrm{W}(b)$ is most likely word w_x in a set of X. Here, buttons is not required to consider calculating because it depends on the words (therefore, $P(b|w_x)$ is 1).

If systems known statistic information of possible words, it may be load stored database. However, it is impossible to extract probabilities about all possible words and if it is possible, it is difficult matter because it is to require huge storage space. Generally, it is to use a corpus in order to solve this problem. The corpus is a collection that people who are linguistic researcher have collected various texts in newspapers, magazines or other documents. Therefore, in this paper, we use bag of words that are extracted in corpus for language modeling.

3 Predicting Text Entry for BCI

Here we present the example that the predicting text entry system is controlled by the two mental states right and left hand movement imagery, right hand and foot movement imagery or others that is possible the binary decision though BCI. In this paper, it is no consider that mental state is extracting to feature extraction or classification and then we deals the visual interface for communicate to people who are able to help to a patient. Visual interface consist of 3 parts: text entry, symbols selection, and feedback. With this entry system, user interface is provided with a general entry method and result of prediction given the symbols which is a fixed button.

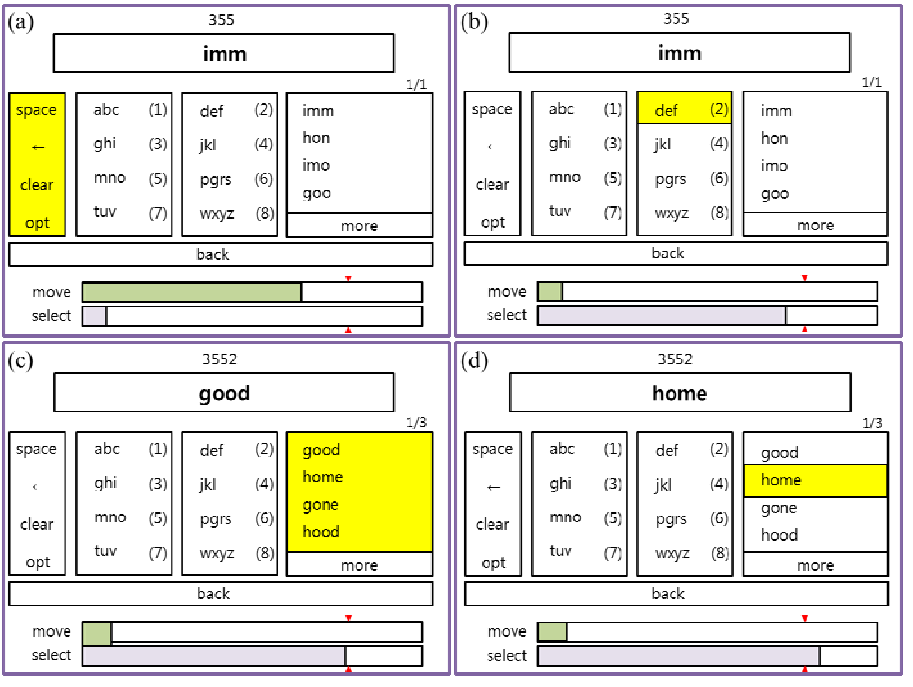


Fig. 1. This figure presents a sequence of interface of text entry system. 4 box of middle is symbols and the activation (yellow) move from left to right. The yellow boxes present a current state. A move and select in the bottom is feedback of system that measured user's current state. And, text box of top is inputted text by user and upper numbers is the set of unique number of selected symbols.

4 Discussion

When inputting text, if the system can predict or otherwise reflect user's intention, the process of writing can be speeded up. Previous studies concerned with feature extractions and translation algorithms for boosting communication rate in BCI. Therefore, our proposed system that is to improve though user interface have a difference from they. We believe that it also be able to boost communication rate. However, it remain that we should solve two problem: one is to solve the disambiguation that when user input the buttons, the predicting result is not involved whole characters that could have some mean although it is non-words. The others is that it is no verify to experiments. In order to verify this system, it need to people, who need a BCI system, and more specific BCI technique that we are no concerned in this paper.

Acknowledgments. This work was supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2010.

References

1. Wolpaw, J.R., Birhaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-Computer Interfaces for Communication and Control. Clinical Neurophysi. 133, 767–791 (2002)
2. Lu, S., Guan, C., Zhang, H.: Unsupervised Brain Computer Interface based on inter-subject information and Online Adaptation. IEEE Trans. on Neural Sys. and Rehab. Eng. 17(2), 1–11 (2009)
3. Williamson, J., Murray-Smith, R., Blankertz, B., Krauledat, M., Müller, K.R.: Designing for Uncertain, Asymmetric Control: Interaction Design for Brain-Computer Interfaces. Int. J. Human-Computer Studies 67, 827–841 (2009)
4. Silfverberg, M., MacKenzie, I.S., Korhonen, P.: Predicting Text Entry Speed on Mobile Phones. In: Proceedings of the ACM Conference on Human Factors in Computing System (2000)

A Safe Driving Support System for Motor Vehicles with DSRC

Myung Hwa Hyun[1], Jisu Park[2], Jingon Shon[3], and Kwang Sik Chung[4,*]

[1] MANDO Co., R&D Administration Department, R&D Planning Team, R&D Division, 413-5,Gomae-Dong, Giheung-Gu, Yongin-Si, Kyonggi-Do, Korea 446-901
mhhyun@mando.com
[2] Dept. of Computer Science Education, Korea University
Anam-dong Seongbuk-Gu, Seoul, Korea
bluejisu@comedu.korea.ac.kr
[3] Dept. of Computer Science, Korea National Open University
169, Dongsung-dong, Jongro-Gu, Seoul, Korea
{jgshon,kchung}@knou.ac.kr
[4] Dept. of Computer Science, Korea National Open University
169, Dongsung-dong, Jongro-Gu, Seoul, Korea

Abstract. In this paper, a safe driving support system for motor vehicles is disclosed. The system is designed to prevent traffic accidents via inter-vehicle communication even in the case where a driver cannot obtain sufficient information of a region, such as curved roads, crossroads, and the like, using only auditory and sight senses. The system is operated as follows via an antenna and communication module mounted on a vehicle. When receiving information of an opponent vehicle on the road via inter-vehicle communication, the system compares conditions of a host vehicle with those of the opponent vehicle and determines a proper service for the host vehicle. Then, a warning for a driver of the host vehicle was generated as a front-vehicle warning, rear-vehicle warning, side-vehicle warning, host-vehicle warning, etc. in accordance with the determined service. In the experiment for checking the operation of the system, the prototype vehicle sent and received vehicle information, such as vehicle speeds, vehicle positions, parking and stopping of a vehicle, driving conditions, driving directions, and the like, and provided the information to the driver of the host vehicle. Further, the system calculated a vehicle-to-vehicle distance, and processed information of curved roads and crossroads and alarm generation in accordance with a degree of danger through the vehicle terminal.

Keywords: economic scheduling, workflow, mobile grid, maximum flow.

1 Introduction

With increasing replacement of mechanical devices by electronic devices along with increasing installation of convenience devices to provide more intelligent motor

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 313–320, 2011.

vehicles, consumers demand motor vehicles that ensure a higher degree of safety. In our daily life, it frequently happens that a driver experiences an accident or embarrassment due to an abrupt situation in traffic. In this case, if the vehicle under such a situation is provided with means for suitably informing other surrounding motor vehicles of the emergency situation, it may be significantly conducive to achieving traffic safety. Particularly, even in the case where a driver is not given a sufficient visual field due to natural or geographical causes, the traffic accident rate will be significantly reduced if the driver is provided in advance with information of the area in front of a driver's motor vehicle[1]. Most traffic accidents occur due to failure of rapid and accurate judgment on emergency situation around a driver's vehicle. While driving on the road, a driver generally determines, based on auditory and sight senses, information elements, such as a surrounding climate, road shape, road conditions, traffic laws, traffic current, vehicle defects, and the like. As a result, the driver operates the vehicle by stepping on a brake pedal or changing lanes without accurate judgment as to the information elements. Therefore, if the information elements are provided in advance to motor vehicles on the road, a remarkable reduction in traffic accident rate can be achieved.

In this paper, a safe driving support system for motor vehicles based on 5.8 GHz-band dedicated short range communication (DSRC) is suggested. This safe driving support system is designed to prevent traffic accidents via inter-vehicle communication technology even in the case where a driver cannot obtain sufficient information of a region, such as curved roads, crossroads, and the like, using only auditory and sight senses. After receiving signals from an opponent vehicle via inter-vehicle communication, the system compares conditions of a host vehicle with those of the opponent vehicle and determines a proper service for the host vehicle. Then, the system suggests a method of warning for a driver of the host vehicle and a processing procedure based on the determined service. Further, an experiment is performed using prototype vehicles for testing the system.

2 Related Works

Studies on collision prevention and safety of motor vehicles have mainly focused on a motor vehicle, a driver and passengers, that is, capability and mechanism of absorbing collision with respect to the vehicle and the persons therein. That is, the vehicle is designed to ensure not only avoidance of physical contact between a vehicle body and persons in the vehicle due to deformation of the vehicle body for protection of the persons upon collision, but also suitable absorption and reduction of the collision, thereby securing collision safety. However, although this technology can minimize impact by the collision, it cannot serve to prevent the collision[2][3][4][5]. Collision prevention systems include a collision alarm and avoidance function, a night pedestrian and object detection function, a front/rear and side obstacle detection function, a vehicle-to-vehicle distance control function, or the like, which are provided by generalization of an advanced safety vehicle (ASV) concept. These functions are commonly based on distance calculation during driving on the road. Particularly, collision during driving at high speeds or collision with a pedestrian or object at night lead to a much higher death rate than other traffic accidents. Thus, a

function for preventing such accidents is important for the safe driving support system, and can be realized through positional recognition with respect to front vehicles or obstacles, rearward and side vehicles, and pedestrians[5][6]. A vehicle collision prevention algorithm includes an algorithm for measuring vehicle-to-vehicle distances and speeds using sensors, cameras and radar, an algorithm for warning of collision between motor vehicles, and an algorithm of the collision prevention system. The algorithm of the collision prevention system is used as grounds for judgment of danger by an electronic control unit (ECU) in accordance with signals input from the sensors and is believed to be the most important part of the collision prevention system. When setting a short danger distance leads to frequent alarm generation, a driver is liable to become insensitive to the alarm and may ignore an alarm informing of an emergency situation. Further, if an alarm distance is set only for the emergency situation, a driver can be put in danger due to an insufficient time for the driver to cope with the emergency situation. The algorithm of the collision prevention system serves to set a limit of the alarm distance and a limit of a stop distance, with which a driver can cope, based on vehicle speeds, relative speeds, and various other experimental data[7][8].

3 Design of Safe Driving Support System for Motor Vehicles

The safe driving support system for motor vehicles in this paper is designed using DSRC technology. Hereinafter, the system will be described in terms of object, configuration and function, communication method and processing procedure thereof, driver warning method, calculation for prevention of vehicle accident, and the like.

3.1 System Configuration

The safe driving support system includes four parts, that is, an antenna and communication module, a main controller, an alarm device, a chassis, and electronic equipment.

The communication module is a short-range communication module constituted by integrating a Global Positioning System (GPS) module, a DSRC module, and other communication modules, which communicate via an antenna and an RF (radio frequency) modem. The antenna and communication module serves to enable communication between motor vehicles and to send and receive a variety of state information, such as positions and directions of surrounding motor vehicles, various road conditions, and the like. The main controller is provided as a central processing unit, which processes data transmitted and received by the short-range communication module and sends the processed data to the chassis and the electric assembly. Furthermore, the main controller serves to obtain state information of a host vehicle and send it to the alarm device to improve positional accuracy. The alarm device provides a suitable alarm to a driver in response to the information sent from the main controller. The alarm can be provided to the driver through voice service, alarm light, alarm sound, a navigation terminal, a display (LCD monitor), a vibrating seat, a seat belt, and the like. The chassis and the electronic equipment communicate with the main controller via a controller area network (CAN) to provide intra-vehicle information and to correct vehicle-to-vehicle information, such as vehicle-to-vehicle

distances (brake distance), vehicle speeds, and the like, through comparison of the host vehicle information with the data received from the short-range communication module. A communication system between components in a motor vehicle is illustrated in Figure 1, in which the chassis and the electronic equipment communicate with the ECU and the main controller via the CAN in the safe driving support system. The modem (IEEE 802.15.4), GPS, and the main controller are connected to each other via RS232, and the alarm device is composed of I/O units or the like.

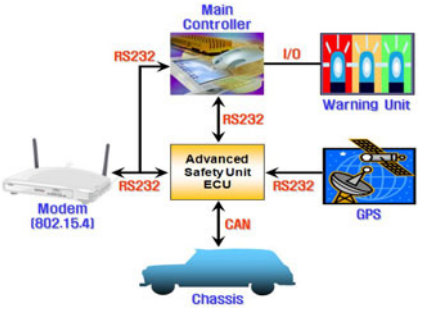


Fig. 1. Configuration view of vehicle communication system

In a vehicle communication procedure, information packets are first received by hardware and used to measure, determine and calculate the vehicle-to-vehicle distances, the vehicle speeds, etc. by software, and an alarm is provided to a driver. After warning to the driver, the hardware sends information packets to the host vehicle and surrounding vehicles. The vehicle communication procedure includes the steps of information receipt, wireless communication, vehicle-information processing, recognition, warning, and information transmission. Figure 2 is a block diagram of vehicle communication.

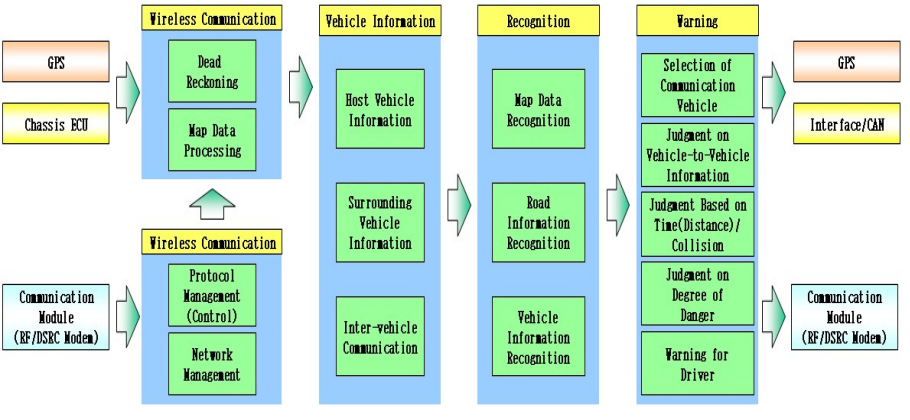


Fig. 2. Block diagram of vehicle communication

3.2 Calculation Method for Prevention of Vehicle Accident

In this paper, a calculation method for prevention of vehicle accident based on time and collision possibility is suggested to overcome drawbacks of conventional methods using ultrasonic waves, lasers, and other distance calculation methods. The calculation method is based on values obtained by the following Expressions 1, 2 and 3. An alarm is provided based on the calculation to inform a driver of a safe state (green), a caution state (yellow), or a danger state (red). As for a vehicle communication range, the communication module of a host vehicle communicates with an opponent vehicle in the maximum communication distance of 600 meters when driving at high speeds, in a communication distance of 200~400 meters when driving at medium speeds, and in a communication distance of 200 meters or less when driving at low speeds.

The calculation method for prevention of vehicle accident based on time is as follows:

$$d = c \cdot v \qquad \text{<Expression 1>}$$

d : Distance Gap(m)

c : Time gap, Driver safety and accident coping capability (1~3 sec., average 2 sec.), time that a driver believes safe (c = 2 sec.)

v : Host Vehicle Speed(m/sec), vehicle speed function, host vehicle speed

The calculation method for prevention of vehicle accident base on the collision possibility is as follows:

$$Pc > Kp \quad \text{<Expression 2>}$$

$$Pc > Kd \quad \text{<Expression 3>}$$

Pc : collision index (collision regulation of the IEEE standard)[9].

Kp : value obtained by dividing relative speed by vehicle-to-vehicle speed

Kd : value obtained by dividing relative acceleration by vehicle-to-vehicle speed

According to whether the collision index is greater than or equal to a predetermined value, it is possible to recognize the accident or dangerous situation. The degree of danger increases with decreasing time difference, and when the collision index is greater than a value obtained by dividing a relative speed by a vehicle-to-vehicle distance and a value obtained by dividing a relative acceleration by the vehicle-to-vehicle distance, it is determined that the degree of danger and the possibility of collision are high.

4 Experiment of System

This experiment for inter-vehicle communication was performed to warn of a dangerous situation when a driver faces a front sight-field non-securing region, such as a sloped road or a curved road, during driving. In this experiment, an opponent vehicle stopped on a sloped road and a driver in a host vehicle couldn't see the opponent vehicle due to the slope which shielded the host vehicle driving on the sloped road. That is, the experiment in the front sight-field non-securing region was performed under conditions that the opponent vehicle (red color on the detection

screen) stopped on a sloped road and the driver in the host vehicle (green color on the detection screen) couldn't see the opponent vehicle due to the slope which shielded the host vehicle driving on the sloped road. Conditions for the experiment and results thereof are shown in Figure 3, and details of the experiment are as follows.

(1) The opponent vehicle stopped on the sloped road and the driver in the host vehicle couldn't see the opponent vehicle due to the slope which shielded the host vehicle driving on the sloped road.

(2) The opponent vehicle remained stopped on the slopped road and the host vehicle was driving on the road.

(3) The host vehicle was approaching the opponent vehicle while driving on the road. Communication between the host vehicle and the opponent vehicle was in an enabled state and a safety alarm (that is, a green alarm generated based on judgment of a distance between the host vehicle and the opponent vehicle) was provided to the driver of the host vehicle by comparison of a calculated vehicle-to-vehicle distance with a predetermined value.

(4) The host vehicle was still approaching the opponent vehicle while driving on the road. The communication between the host vehicle and the opponent vehicle was in an enabled state and a caution alarm (that is, a yellow alarm generated based on judgment of the distance between the host vehicle and the opponent vehicle) was provided to the driver of the host vehicle by comparison of a newly calculated vehicle-to-vehicle distance with the predetermined value, and the alarm device was operated.

(5) The host vehicle was still approaching the opponent vehicle while driving on the road. The communication between the host vehicle and the opponent vehicle was in an enabled state and a danger alarm (that is, a red alarm generated based on judgment of the distance between the host vehicle and the opponent vehicle) was provided to the driver of the host vehicle by comparison of a newly calculated vehicle-to-vehicle distance with the predetermined value, and the alarm device was operated. Further, the system informed the driver of the most dangerous state and demanded that the driver perform proper manipulation of the vehicle.

(6) The host vehicle stopped near the opponent vehicle.

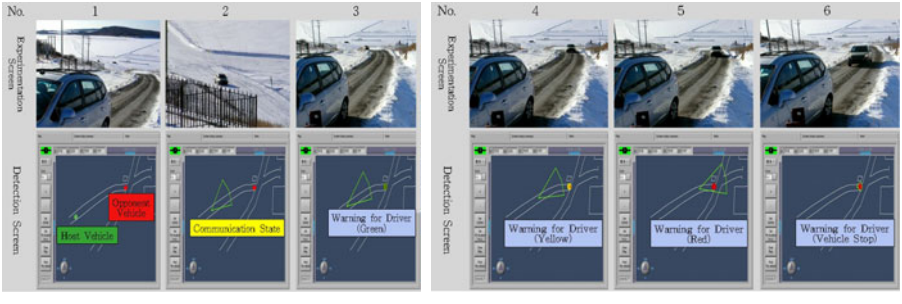


Fig. 3. Experiment in a front sight-field non-securing region and detection screen

5 Conclusion and Study Plan

In this paper, the safe driving support system for motor vehicles with DSRC, which is an inter-vehicle communication technology, is suggested and results of experiments for testing the system using the prototype vehicles are disclosed. The safe driving support system suggested in this paper includes four parts, that is, the antenna and communication module, the main controller, the alarm device, the chassis, and the electronic equipment. With the system configuration as illustrated in Figures 1, 2 and 3, the experiments for testing the system were carried out. In other words, when receiving signals sent from an opponent vehicle via inter-vehicle communication, the system compared the conditions of the host vehicle with those of the opponent vehicle and determined a proper service for the host vehicle. Then, a warning for the driver of the host vehicle was generated as a front-vehicle warning, rear-vehicle warning, side-vehicle warning and host vehicle warning in accordance with the determined service. The alarm device was operated in accordance with a degree of danger judged based on time and collision index.

The results of the experiments showed that the inter-vehicle communication was performed with reference to the host vehicle communicating with the opponent vehicle, a single node occupation ratio was 30%, and the maximum communication distance was 600 meters. Further, the system response rate was up to 20 Hz and the system detention time was up to 2,000 ms. Moreover, the vehicle-to-vehicle distance, vehicle speeds, vehicle positions, driving directions, road shape (curved road and cross road) were obtained as the vehicle information, and the alarm was provided to the driver in accordance with a degree of danger judged via the inter-vehicle communication. Based on such results of the experiments, it is planned to realize the system by establishing several nodes while using various communication media in order to deduce grounds for judgment of danger in various ways. If some restrictions can be solved, it is believed that the system will be operated at a single node occupation ratio of about 5%, a system detention time of up to 1,000 ms, and a communication frequency of 4 Hz under conditions that a communication error rate can be directly measured through experimentation.

Additionally, it is planned to realize and evaluate a vehicle network system corresponding to a multi-vehicle environment. Further, it is planned to design and realize an automatic automobile control system in association with C2C (car-to-car) communication and C2I (Car-to-Infrastructure: signal light node, display board node, road-side equipment, etc.) communication. Moreover, it is planned to realize a method of securing an integrated positional accuracy of communication modules for hardware stabilization and performance improvement, an ECU (a safety ECU of vehicles) for advanced safety equipment, and sensors.

References

1. Yu, S.D., Cho, K.H.: Vehicle Safety Communication Technology Trends. In: IITA Weekly Technology Treands, vol. 1292, pp. 26–28 (2007)
2. Lim, Y.-C., Ryoo, Y.-J.: Autonomous Land Vehicle Trends. Proceedings of KIEE 45(5) (1996)

3. Germann, S., Isermann, R.: Nonlinear distance and cruise control for Passenger cars. In: Proceedings of the American Conference, Seattle, Washington, pp. 3081–3085 (1995)
4. Lee, J.R.: Ultrasonic waves and Usability. SeaHwa Press (1997)
5. Cheol, K.S.: A Study Collision Avoidance system of Vehicle using Fuzzy Control Algorithms. Dept. of Electronics Engineering Graduate School, Dong-A University, pp. 1–4 (2001)
6. Moon, Y.J.: Telematic/ITS Service Standards and Development Case. TTA Journal 113, 60–61 (2007)
7. Kim, B.K., Yi, H.C., Han, S.C.: A Study on Rear Lateral Collision Warning System of Vehicle using Fuzzy Control Algorithms. Proceedings of Korean Society for Precision Engineering 24(2) (2007)
8. Yi, K., Jeoung, T.Y., Woo, M.S., Lee, S.C., Oh, T.I., Kim, S.H.: Vehicle Collision Warning/Avoidance Algorithms. In: Proceeding of the Korean Society of Automotive Engineers, pp. 30–36 (1997)
9. IEEE, IEEE Standard for Motor Vehicle Event Data Recorders (MVEDRs), IEEE Std 1616TM-2004, p. 32 (2005)

Changes and Directions of Mobile Technology and Infrastructure to Enhancing Public Education: Policy and Strategy Perspectives in South Korea

JiYoung Choi[1], Hye Jeong Kim[1], and WonGyu Lee[2,*]

[1] Department of Computer Science Education, Korea University, Seoul, Korea
{jiyoung.choi,hyejeong.kim}@inc.korea.ac.kr
[2] Department of Computer Education, Korea University,
Anam-dong, Seongbuk-gu, Seoul 136-701, Korea
lee@inc.korea.ac.kr

Abstract. In this study, we review national informatics policy and strategy whether the policy and strategy support the use of mobile technology in public education. The rapid growing of mobile technology and infrastructure changes social and educational paradigm. With regard to the technological changes, it is necessary to establish a sustainable national policy and strategy for supporting public education. For this, we review relevant documents and papers based on an analysis framework. In addition, we discuss the directions of national informatics policy and strategy in order to strengthen the capacity of public education in South Korea via mobile learning.

Keywords: mobile technology, education policy, mobile learning.

1 Introduction

The mobile devices are getting faster and smarter after the market introduction of the mobile phones in the 1990s. Also the users can be staying connected on the internet using cellular-based hotspots and mobile broadband cards or to a Wi-Fi network. Recently, a massive number of people are experiencing smart phones as a computer that fits in people's hand. Many people all over the world, they can access to tools for business, edit video or audio files, examine geological information, use social networking services, and use educational materials and environments, indeed, people can do many things like on a desktop [1]. It becomes a common learning and communication tools in our daily life. The mobile devices and capabilities can make them ideal tools for learner's communication and productivity in learning. For example, the use of mobile devices in higher education has been increased for more interactive learning environments. The school adopting the framework using the mobile devices on campus has been increasing for the purpose of using learning management system.

Particularly, the use of mobile devices having the function of wireless computing has many benefits for teaching and learning. McKenzie [2] asserted that the merits of

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 321–324, 2011.

wireless computers for teaching and learning include ease of movement, relaxed fit, strategic deployment, low profile, flexibility, cleanliness, convenience, simplicity, ad speed. In addition, Oku [3] identified the benefits of mobile computing includes overcoming learning hours and locations, enhancing personalized learning by individual needs and level of ability, and updating the contents of lecture.

To cope with ubiquitous society based on high-tech wireless internet technologies, it is necessary to consider various strategies for the use of mobile devices in school. In general, learners are using educational mobile contents via wireless internet and mobile console in the basis of web-based e-learning environment. The type of mobile-based learning is at the initial stage. However, the mobile-based learning has the potentials of ease to use for learning because of the flexibility of time and distance. Mobile learning means using hand-held technology, mobile communication technologies, enhancing learning and education. The mobile learning is expected to happen with relatively short time and quick adaptation. In addition, mobile learning can customize educational contents and include educational resources or personalized contents. In mobile learning, functionality is more important than the devices and about experiences than the mobile technologies.

There is a rapid increase in the market of mobile educational application. However, the use of mobile applications in school education is still limited. It may need to consider as the view of educational or information technology policies regarding mobile technologies instead of the issues of technological framework or devices (KERIS, 2010). To boost mobile learning in school and educational institute, Go(2008) reported that it is necessary to establish specific national policy or practice (34%) and to participate actively in mobile learning activities and active attitudes (44%), see table 1.

Table 1. The impacts of mobile learning and ubiquitous learning

	specific policy and practice	Participation and active attitude	Enhancing curriculum	others
Mobile learning	34%	49%	13%	4%
Ubiquitous learning	22%	64%	10%	4%

To accomplish the effective environment for mobile learning, it is necessary to consider policy regarding the relevant supports to school and educational institute. Also, to enhance public education, the infrastructure and technologies regarding mobile learning should be considered in the view of quality reconsiderations and change of education environments.

2 Policies on Mobile-Based Technology, Infrastructure, and Learning

To support the adaptation of mobile technologies in education, we examine the efforts of governmental institutions and government regarding national framework of mobile

technologies in education. We examine the information and communication technology infrastructure necessary for effective mobile learning in Korea. To improve ubiquitous connectivity for effective mobile learning, mobile communication networks need sufficient infrastructure to support the application of mobile devices using wireless data. In South Korea, the number of smartphone users rapidly increased and reached 10 million in this year. The annual presentation of Morgan Stanley, internet trends [4], predicts the increase of mobile users. The presentation describe the future trends of mobile uses that the number of mobile users will be more than 1.6 billion in 2014 and exceed the number of desktop users. For mobile-based learning, learners can access to the mobile applications in anyplace in South Korea.

In March 2011, the number of smartphone subscribers in South Korea increased over 10 million and the number of smartphone users is expected to grow rapidly to exceed the 30 million mark by this year, according to the latest report. The number account for 20% of the country's population. Also the report predicted that the rapid growth of smartphone users led to a rise in wireless traffic, which will jump more than 8 times as of the increase of wireless traffic from 5,496 terabytes in January 2011. One year ago, the wireless traffic was 449 terabytes. However, the increase of wireless network traffic will reach soon in limit of wireless network capacity by the end of this year. It may be possible to cause poor wireless quality.

The Korea Commutations Commission (2010) is preparing the policy to enhance (or Advance) broadband wireless infrastructure such as public WiFi, Wibro, or 4G and to increase the use of networks. With regard to mobile learning, the national informatics roadmap in education consists of multiple stages and projects for the informatics in education, such as infrastructure development in first stage (1996-2000), information and communication technology education in second stage (2001-2005), and adaptation of ubiquitous learning systems (2006-2010).

According to the Korea Research Institute for Vocational Education & Training (KRIVET), a significant barrier in mobile learning seems to be a lack of people who design and develop the mobile-based learning. It causes the lack of learning contents in mobile learning. According to a report of the KRIVET, it surveyed the current status of the development and management of mobile learning courses operating by e-learning institutions. The institute recognized that the infrastructure for mobile learning should consider human resources, contents, culture, and resources as fundamental components. However, the institutes responded that the level of readiness on the infrastructure for mobile learning is at the beginning stage.

However, there is a lack of human resources in mobile learning design. It is necessary to support educating experts in designing mobile learning by governmental institute. Currently the relevant policy is not presented for the development of human resources. According to the report of KRIVET, government and other institutes regarding mobile learning did not extract the features in education and training views, such as in sharing, collaboration, interactivity, and learning context. And then people who in charge of making decision in policy development and practice do not recognize the features of mobile devices and technologies.

According to a governmental report regarding the development, vision, and strategies of 2020 ubiquitous learning country [5], the government has a plan to establish mobile environments to support various operating systems in mobile devices

and browsers. Working in new device, the government will set up one source multi device system. Also, there are plan to establish ubiquitous learning zones and managing of them.

3 Discussion

In terms of our analyses, we reviewed current mobile technologies and infrastructure in the national level for the implementation of mobile-based learning in public education. Even though there are huge changes in mobile technologies and infrastructure, it is necessary to reconsider the change of relevant people's understanding of the effect of mobile devices in learning contexts. Also it should have a plan to train expert people in designing and developing the mobile contents in national level. Also the market of open platform-based educational content development should have to activate through educating human resources and market places. For this, South Korean government should also invest in developing an integrated network among school, society, and government. For the public education, still there are many things to be solved in mobile-based learning.

References

1. New Media Consortium: Horizon Report NMC (2010), http://www.nmc.org/publications/2011-horizon-report
2. Mckenzie, J.: The unwired classroom: Wireless computers come of age. Educational Technology 10(4) (2001)
3. Oku, M.: High-level educational performance brings innovations to corporate culture, http://www.sw.nec.co.jp/english/special/e_learning/tomess01.html (retrieved November 9, 2005)
4. Stanley, M.: Internet trends (June 2010), http://www.morganstanley.com/institutional/techresearch/pdfs/MS_Internet_Trends_060710.pdf
5. Kwak, D.H.: The report for, ubiquitous learning country system with visions and strategies. Presidential Council for Future & Vision (2010)

A Cognitive Ability Measuring System with Cognitive Loading Task

YeongWook Yang and HeuiSeok Lim

Dept. of Computer Science Education, Korea University
yeongwook@blp.korea.ac.kr

Abstract. This paper was developed a cognitive ability measuring system to investigate that give effect of a cognitive load capacity. This paper of a hypothesis is increase of cognitive abilities when cognitive load capacity is increase and then cognitive abilities are decrease when cognitive load overload. This system offers about cognitive abilities that are attention, memory, processing changes and processing difficulty. The process observes to change of cognitive ability through with 2minute of cognitive load task and 2minute of cognitive measuring task. This paper of result affects the attention ability to cognitive load capacity, but other cognitive abilities are not affects.

Keywords: Cognitive ability, Cognitive load.

1 Introduction

Currently studies try to develop new technology that valid field merger. In order to offer comfortable of technical as well as emotional, the field related cognitive ability is necessary.

The cognitive ability means skill to distinguish things and to conduct some work. The cognitive abilities include attention, memory, processing change, speed and others. Early, this is called 'IQ Test'[1]. If the people less cognitive abilities will be confronted with some problem that ADHD(Attention Deficit/Hyperactive-impulsive Disorder), AD(Alzheimer's Disease), language disorder and the others. Therefore, the cognitive abilities are indispensable in our daily life.

The cognitive ability and cognitive load closely related that cognitive ability increase or decrease[2][5][6]. This paper had a hypothesis that cognitive abilities are increase when cognitive load is increase, and that cognitive abilities are decrease when cognitive load overload.

This paper of proposed system is a cognitive ability measuring system with cognitive loading task. This system of purpose is to measure variation of cognitive ability according to cognitive load.

2 System Architecture

This system was developed using the Flash. The database constructs of the Mysql. This system consists of a cognitive load module, a measure module and database. The database store measure values and have stimulus data of processing difficulty cognitive load.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 325–328, 2011.

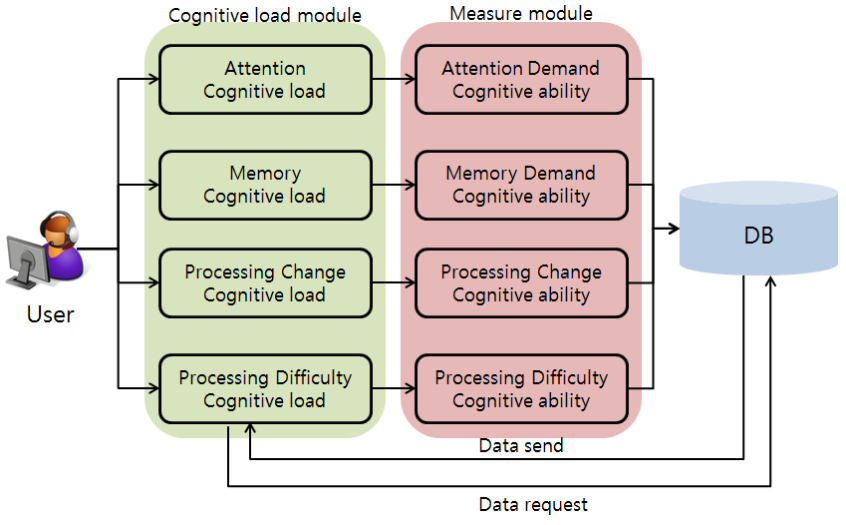


Fig. 1. This figure is system architecture that consists of two module(cognitive load and measure) and one database

2.1 Cognitive Load Module

Richard A. Block define cognitive load type: attention demands, response demands familiarity, memory demands, processing changes, processing difficulty[2]. This paper used attention demands, memory demands, processing change and processing difficulty.

Before cognitive ability measure, this module gives a stimulus that related attention, memory, processing change and processing difficulty.

Attention cognitive load task is to find an alphabet 'Z' at 5x25matrix. Other alphabets consist of 2 groups. One group include alphabet 'C', 'D', 'G', 'O', 'Q', 'R' and 'U'. Another group include alphabet 'E', 'I', 'M', 'V', 'W', 'X' and 'Y'. Each alphabets of position are random.

Memory cognitive load task is 2-back task. The 2-back task is to decide whether two number that current and 2-back number correct. If current number is 4 and 2-back number is 3, this case is not correct. But this case is correct when 2-back number is 4.

Processing change cognitive load task consists of two mode that length and scale. The length mode is to decide whether the number of characters are big than 5. The scale mode decides value of number.

Processing difficulty cognitive load task consists of two types. One type is to find that sub-words of some word. Another type is to find that the first spell of word include some character. This task has reference at [2].

2.2 Measure Module

This module is to measure cognitive ability. Total 5 cognitive abilities are attention demand, memory demand, processing change and processing difficulty.

Attention demand is to find a blue triangle. Shapes are triangle and rectangle. This part measures response time and correct rate.

Memory demand should be remembered number in reverse order. If this task have a sequence of numbers that 4, 3 and 1, you should be answer 1,3 and 4. This part measures capacity of memory.

Processing change consists of two parts. One part has an arrow that red color. Another part has an arrow that blue color. Red color means opposition direction. Blue color means same direction. If the arrow has red color and left, you should be push the right button. This task measures response time and correct rate.

Processing difficulty is to decide whether two sentences of mean are correct. This part measures response time and correct.

This module of result values sends the database. The result values are cognitive abilities measurement value.

3 Experiment

One hundred four observers between the ages of 20-30 participated. Each task consists of attention 25 observers, memory 25 observers, processing change 26 observers and processing difficulty 28 observers. This research is one trial that runs 2 minute with cognitive load task and cognitive measure task. It was performed total 4 trials.

4 Result

This research of result includes cognitive abilities capacity with attention, processing change and processing difficulty. The factors of cognitive abilities are response time, memory capacity and correct rate. This result was removed meaninglessness values.
The attention task values shows figure 2. Left is response time. And Right is correct rate. It is showed that first trial better than last trial. This result means increase attention cognitive ability.

But values of other tasks are not significant. It is because not affect cognitive load.

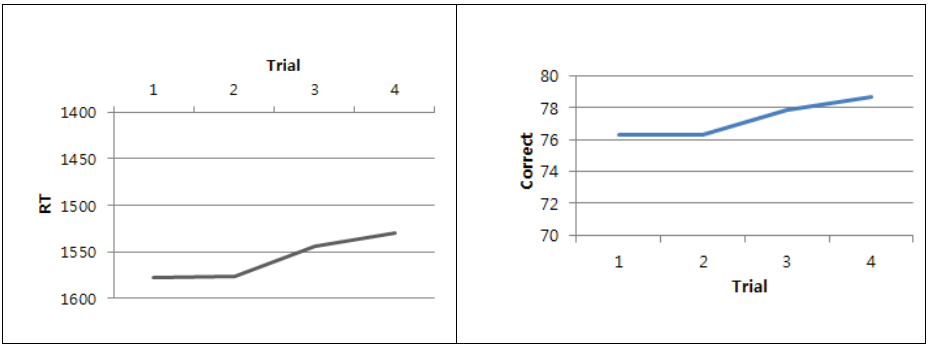


Fig. 2. This is value of attention task. Left is response time. And Right is correct rate. It is showed that first trial better than last trial.

5 Conclusion

Through this research, this examined change of cognitive ability according to the cognitive load. The attention task saw increases in value. But other tasks were almost no changes in value. It was may be wrong hypothesis or does not affect cognitive load. Through new approach, it should be identified as the cause.

The limitation in this research does not know the degree of cognitive load. We need to measure the degree of cognitive load to obtain accurate value. And we will try to experiment about all ages.

Acknowledgments. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST).(NO.2010-0014325).

References

1. Dickens, W.T.: Cognitive Ability. In: Durlauf, S. (ed.) Forthcoming in The New Palgrave Dictionary of Economics
2. Block, R.A., Hancock, P.A., Zakay, D.: How cognitive load affects duration judgments: A meta-analytic review. Acta Psychologica 134, 330–343 (2010)
3. Mahncke, h.W., Connor, B.B.: Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. PNAS 103, 12523–12528 (2006)
4. Jaeggi, S.M., Buschkuehl, M.: Improving fluid intelligence with training on working memory. PNAS (2008)
5. Barrouillet, P.: Time and Cognitive load in working memory. Journal of Experimental Psychology:Learning, Memory, and Cognition 33, 570–585 (2007)
6. Sweller, J.: Cognitive Load During Problem Solving Effects on Learning. Cognitive Science 12, 257–285 (1988)

Automatic Music Emotion Classification Using Chords and BPM

Sang-yong Park, Han-moi Sim, Mi-kyung Kwon, and Won-hyung Lee

Chung-Ang University, 221 Heukseok-dong, Dongjak-gu, Seoul, South Korea,
Daumsoft, 168-21 Samsung-dong, Kangnam-gu, Seoul, South Korea
deadcode@hanmail.net, kmusic12@naver.com, whlee@cau.ac.kr,
fanto@daumsoft.com

Abstract. Existing research on music emotion classification has used the mp3 files' meta-information and meaning of lyrics or users' feedback description to classify the music emotion after hearing the music. In this paper, we propose the new method that classifying the music emotion by extracting chords and using Beats per Minute information of digital music. Firstly, get the valence value by the used chords, and get the arousal value by using Beat per Minute information and the extract the emotion of the music by mapping such values with Russell's emotion model.

Keywords: Emotional Music Classification, Chords and Harmonies, Tempo and Beats per Minute.

1 Introduction

In Recent years, research in Music emotion classification has been very active within both of scientific and commercial area. After inventing mobile digital music players, digital music is major way to enjoy music. Moreover music listeners tend to have huge amount of music and also they listen many new songs every day. In this situation, music emotion classification techniques [7] such as music genre detection [8], speech/music analysis, spoken language analysis, is in the spotlight.

Music emotion classification is one of important methods to introduce music. Melon.com, a representative online music store in Korea, already providing the emotional classification for music and many music listeners use the classification when they select music. However, the method to classify music emotion is depends on experimental results such as users' feedback. In this case, providing an emotional classification for new music is meaningless, before enough data was collected and classified.

Our focus is to classify emotion of music by its own feature. Modern music mostly has certain typical features such as chords and BPM information. Classifying emotion of music by using these features can produce meaningful result. In this paper, we propose a method that automatically classify emotion of music by using their chords and Beats per Minute.

In section 2, we review a model of emotion and feature extraction method to apply into our system. By using selected two-dimensional model, we propose a classification

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 329–336, 2011.

system in section 3. In section 4, the results of the proposed music emotion classification method was evaluated.

2 Related Works

2.1 Model of Emotion

In order to select a suitable model of emotion, we discussed established model of emotion. The most common of the dimensional approaches to classify emotion is that James Russell's two-dimensional bipolar space.

Russell's model of emotion uses pleasant/unpleasant for one dimension and arousal/non-arousal for the second dimension. And emotional words have their own position in two-dimensional space of Russell's model of emotion.

In this paper, the proposed music emotion classification method uses two computable features of music independently. Therefore, Russell's model of emotion, shown in Figure 1 is suitable to map the two features on each dimension.

In order to evaluate the result of music emotion classification, categorization of music is demanded. Most of emotional words represented in Russell's model of emotion, is also suitable to categorize emotion of music. Although Russell's model of emotion is not perfectly appropriate for the proposed music emotion classification method, it can be useful with a few modifications.

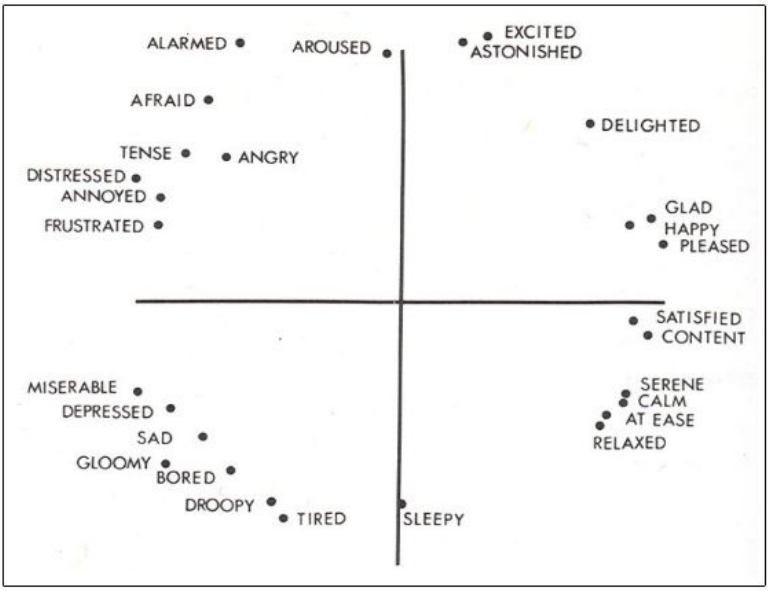


Fig. 1. Multidimensional scaling of Russell's model of emotion [1]

2.2 Music Theory

Harmony is "the combination of simultaneously sounded musical notes to produce chords and chord progressions having a pleasing effect". [2] Consonant chords, such as major chords, are often pleasant, happy, and relaxed. Complex harmonies create instability in a piece of music and activate emotions of excitement, tension, anger, and sadness. [3] [4]

Tempo is "the speed at which a passage of music is or should be played", [2] and is typically measured in BPM. A fast tempo falls into the range of 140 to 200 BPM and a slow tempo could be anywhere between 40 and 80 BPM. Fast tempi are generally considered as lively and exciting, while slow and sustain tempi are majestic and stately. Depending on music, a fast tempo can trigger such emotions as excitement, joy, surprise, or fear. A slow tempo is typical of calmness, dignity, sadness, tenderness, boredom or disgust. [3] [4]

2.3 Feature Extraction of Music

In order to classifying music emotion, two selected features are mapped on each dimension in Russell's model of emotion. Therefore, each feature must be a computable form.

In this paper, high precise BPM information or chords information is not demanded. Therefore, by using the Short Time Fourier Transform, beat and pitch and harmony information can be extracted. [5] [6] BPM information is mapped on arousal/non-arousal and chords value is mapped on pleasant/unpleasant.

3 Music Emotion Classification

Within this section, the music emotion classification framework is presented with the design of an appropriate emotion model.

3.1 Emotion Model Design

The Russell's model of emotion is introduced in section 2.1 provides a basic idea. But, Russell's model of emotion is not perfectly suitable to our music emotion classification system. So, we design a new model of emotion based on Russell's model. A dimensional model is logically divided in its area. We categorized the Russell's dimensional model as 9 areas, shown in Figure 2.

These 9 keywords cannot present all of emotion. Keywords just represent each area. In order to categorize emotional words as keywords, we collect 1,588 comments which include emotional words and music title, from 102 Korean website. In order to categorize 106 collected emotional words as 9 keywords, emotional words in a same comment are treated as a same emotional category.

Figure 3 and Figure 4 show that how to categorize emotional words as keywords .

And also the title of music is tagged with keywords for evaluating music emotion classification method.

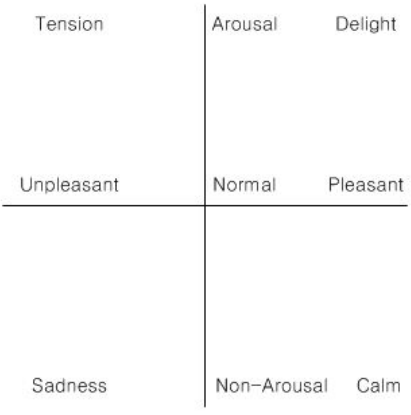


Fig. 2. Valence/Arousal Model with divided areas

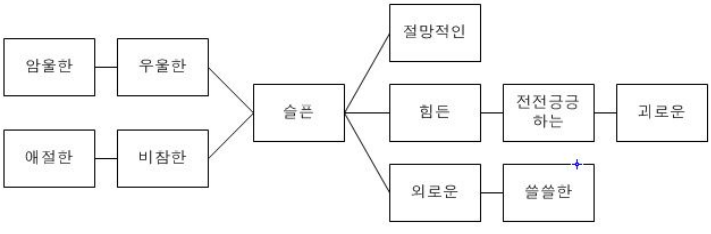


Fig. 3. Korean Emotional Words in a same category

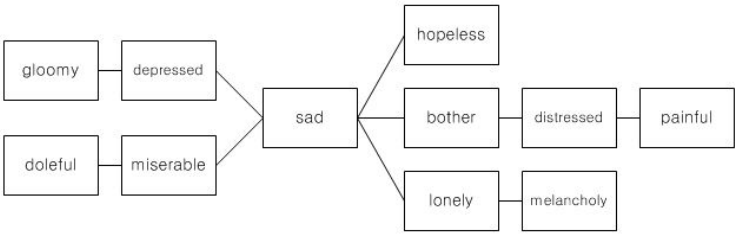


Fig. 4. Translated Emotional Words in a same category

3.2 System Architecture for Music Emotion Classification

Automatic music emotion classification is a relatively common technique. We divided into two features, BPM information and chords. Extraction of two features uses the FFT value of music object. In this system, FFT runs 40 times in a second. Also each processes runs every 40 times in a second.

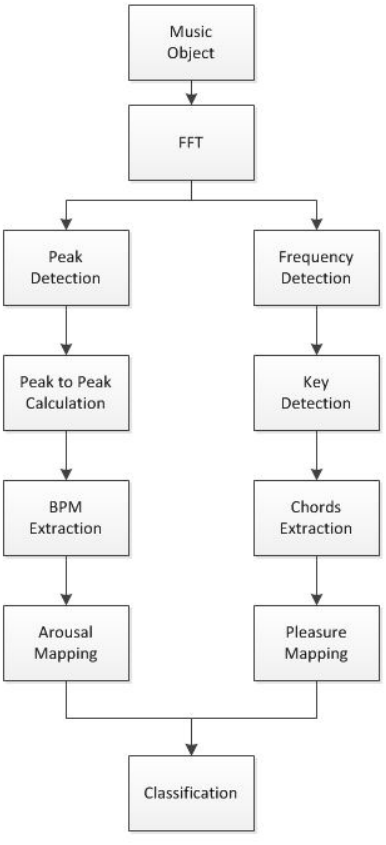


Fig. 5. System Architecture

BPM information is extracted from the value of STFT with peak to peak detection. BPM value can be mapped to Arousal value by rescaling. Criterion of BPM rescaling is in section 2.2.

Chords detection of each process can be extracted from combining of key values. [5] Chords are divided up into major and minor. By calculating the proportion between major and minor chords, chords information can be mapped to valence value.

The emotion of a song is the keyword of area that is decided by valence and arousal value in two-dimensional bipolar space.

3.3 Interface

The emotion classify system interface shown in Figure 6, provides the information such as 'Artist', 'Title', 'Genre' from meta-information of songs and 'Pleasant', 'Arousal', 'Emotion' from features of songs after finishing analysis. And users observe the process of analysis.

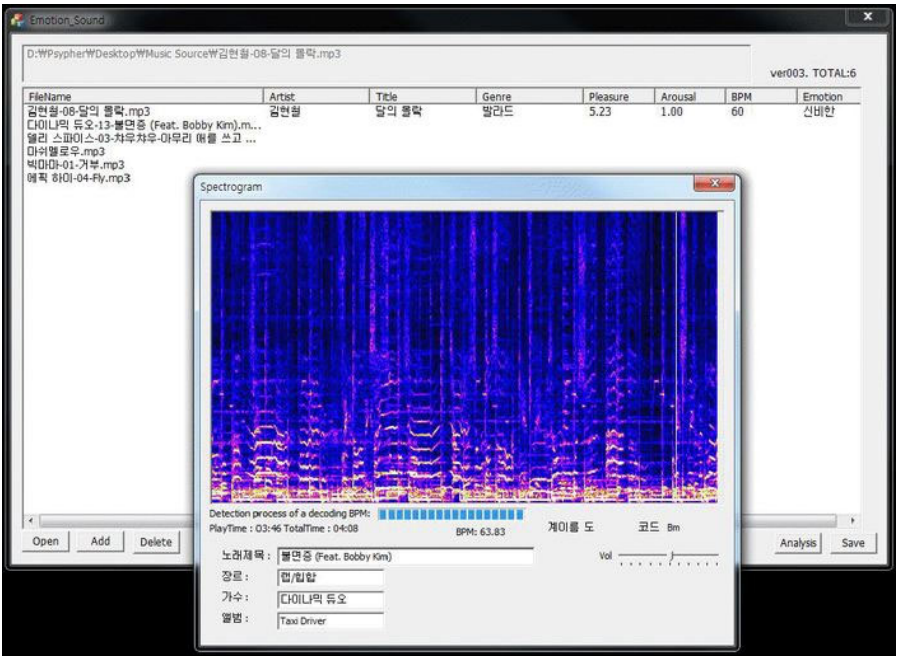


Fig. 6. Interface of the proposed music emotion classification system

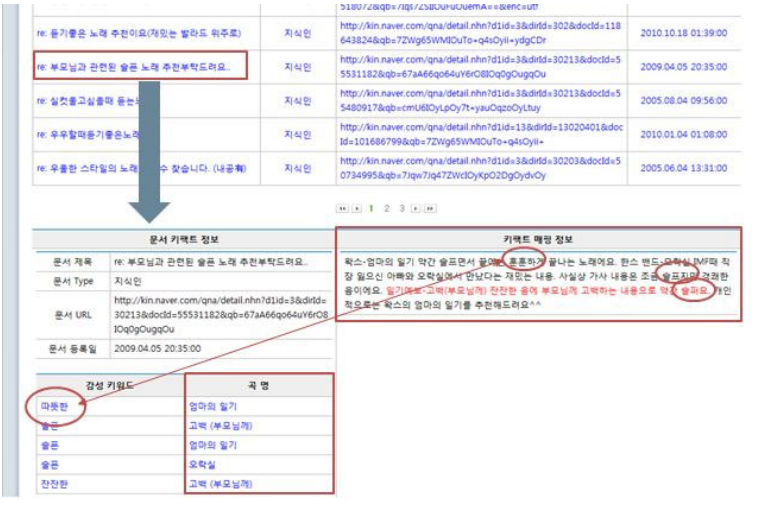


Fig. 7. Reference Set

4 Evaluation

4.1 Reference Set

Generating reference set is important process. In order to get more precise reference set, we collected 6,588 comments which include emotional words and music title, from 102 Korean website. In order to categorize 106 collected emotional words as 9 keywords. Finally, the titles of music are tagged with keywords for evaluating music emotion classification method are shown in Figure. 7.

4.2 Result

In order to measure the classification accuracy, we incorporate the number of successful classification SC, and the number of total sample songs TS. The accuracy of music emotion classification is SC/TS – 47%.

5 Conclusion and Future Work

This music emotion classification system is not perfect and never should be. In spite of that, the demand about music emotion classification is increased consistently. We propose a simple music emotion classification system excluding experimental method for new songs. The accuracy of this system is not competent yet, but this system has enough potential to improve.

The lyric of a song sometimes contain big information connected with emotion. Lyric analysis information could improve this system. However, analyzing Lyrics is tough to expect high accuracy because, same word often has two different emotions, some kind of words has too many meaning in a word. Especially, automatic lyric analysis system is hard to perceive metaphor, sarcasm or irony.

Finally, gathering feedback from users system by providing music emotion information, are demanded. The feedback from users might be helpful to generate more precise reference set.

Acknowledgments

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2011.

References

1. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161–1178 (1980)
2. McKean, E. (ed.): New Oxford American Dictionary, 2nd edn. Oxford University Press, New York (2005)
3. Gabrielsson, A., Lindstrom, E.: The Influence of Musical Structure on Emotional Expression. In: Music and Emotion: Theory and Research, pp. 223–248. Oxford University Press, Oxford (2001)

4. Meyers, O.C.: A Mood-Based Music Classification and Exploration System, Master thesis, Massachesetts Institute of Technology (2005)
5. Jehan, T.: Creating Music by Listening, PhD thesis, Massachesetts Institute of Technology (2005)
6. Catteau, B., Martens, J.P., Leman, M.: A probabilistic framework for audio-based tonal key and chord recognition. In: Advances in Data Analysis, Springer, Berlin (2007)
7. Lu, L., Liu, D., Zhang, H.-J.: Automatic Mood Detection and tracking of music audio signals. IEEE Tras. Audio, Speech, and Language Processing 14(1), 5–18 (2006)
8. Li, T., Ogihara, M., Li, Q.: A Comparative Study on Content-Based Music genre classification. In: Proc. ACM SIGIR 2003, pp. 282–289 (2003)

The Real-Time Stereoscopic 3D Image Conversion of Photo and Video Contents

Jun-Eui Lee, Sang-Won Lee, and Won-Hyung Lee

GSAIM, Chung-Ang University, 221 Heukseok-dong, Dongjak-gu, Seoul, South Korea
reejun@nate.com, adolphlee@naver.com, whlee@cau.ac.kr

Abstract. We propose a method for converting existing 2D image contents to 3D image contents by using hardware based real-time 2D/3D stereoscopic image conversion algorithm. The stereoscopic converted images come from the method that generates the 3D stereoscopic images from 2D images or 2D videos, not videos from stereo camera. The real-time 3D stereoscopic conversion algorithm is produced by hardware that is based on physical board to complement computational complexities and slow down caused by the limitation of processing. The existing interactive video art which is using screen and PC camera, was able to present various kinds of general exhibition, however, it had limited technology to express real-time stereoscopic.

Keywords: Stereoscopic, Real-time 3D Conversion, 3D Converter, 2D/3D Conversion, Media art.

1 Introduction

The 3D converting progress contains auto converting, in-house software converting, specialized software converting. Photos, films or video contents other than 3D stereoscopic are being created in frame by frame. These contents are rebirthed by the process that separates subjects and backgrounds in various layers and sets the depth value. However, these processes need much more human resources and time.

The stereoscopic image above is created by 3D stereoscopic image from original 2D image or video, not from the stereo camera 'Rig'. This method divides the backgrounds and objects from captured 2D original image, then create a 3D stereoscopic video by using depth map extraction algorithm. Moreover, to deal with huge computational complexities and drop down processing speed, it converts the stereoscopic images by hardware based method, not software.

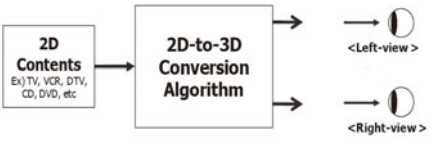


Fig. 1. Basic structure of stereoscopic conversion

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 337–344, 2011.

2 Research Methodology

The convergence between 2D video to 3D video is a method that produces the 3D stereoscopic video which is captured by a camera based on binocular disparity theory above and seemed into both eyes. The convergence of 2D video is processed in sequence-based cut. The case which if convergence is way better than shooting or CG, means it is free to adjust the issue that is hard to view the back when the camera is viewing the front. In case of CG, this can be solve by rendering. The representative case of stereoscopic convergence in hand-by-hand is Depth Mapping. The Depth Mapping is similar with depth map extraction, however, it is difference with the movement of depth map extraction theory (The method that generates similar depth by applying grey level into the frame from the depth map extraction by movement analysis). In this statement, the engineer does the depth extraction.

2.1 2D/3D Conversion Process

In the conversion steps of stereoscopic video, divide the objects and background from existing original 2D video contents, then set the vanish point's pixel distance value from the divided video, and adjust the camera's internal and external parameter value after divide the images above side to side. Thus, generate a stereoscopic video by moving occlusion area and videos according to the camera's location. Proposed method above presents regardless of direction of movement and scene changes. Moreover, it also converts still image to 3D stereoscopic video.

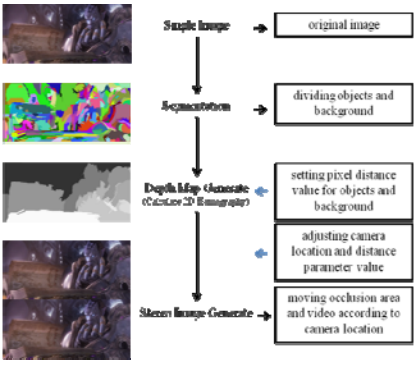


Fig. 2. 3D Stereoscopic Conversion Process

2.2 Automatic Depth Map Generate of 2D Image

The depth map algorithms in this paper require the correct information as an importance. The depth map uses color value from 0 to 255. In case of blockbuster however, it is not enough in case of blockbusters. Most films don't need all 8bit level colors therefore, pixel is integer unit, and the information which is required for generating depth map is bigger than integer.

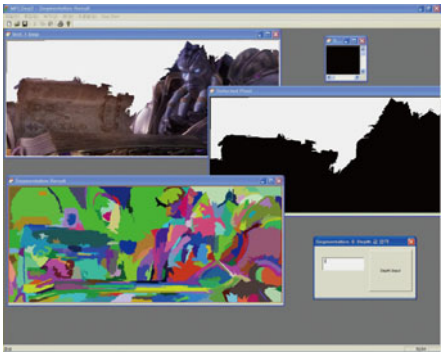


Fig. 3. Generate a depth map automatically by putting depth value for segmented video after inserting original video on the software

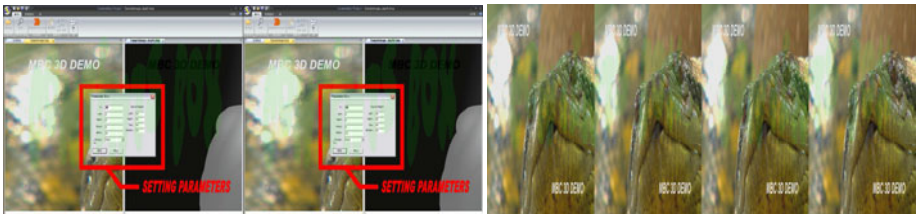


Fig. 4. Automatic Software Setting Parameters and Automatic Software 2D/3D Conversion

Thus, to expand one level more by sub-pixel depth map, it is impossible to represent using 0 to 255. Solve the issue above is most important part. Therefore, depth map has to be different when it sets itself from 0 to 255. Likewise, it has to be different in frame by frame and has to generate new depth map, because the new depth map has to be generated and changed when the objects are moving, blowing and flying. In case of auto generating depth map, it has an issue which is related with generation methodology and it requires higher levels.

2.3 The Epiploar Geometry Algorithm

As stereo convergence sequence in the Fig. 2, divide the objects and background from the 2D original video using depth map generation algorithm in segmentation phase. Then set the any pixel distance value to divided objects and background. In other words, set the Matching Point for original and generated video. Then, seek the 2D homography using such matching point sets. If the Epipolar plane above was a captured video by camera, the generation method is Fig. 5, X means the objects in real space.

The Epipolar Geometry means that the objects in real space have common single plane when it is viewed on the stereo video. Fig. 5 is Fundamental matrix F that explains Epipolar plane. It generates a stereo video using Fundamental matrix of 2D homographic. The numerical formula seeks the value F, then H by using value F. The processing formula Homography is calculated by F.

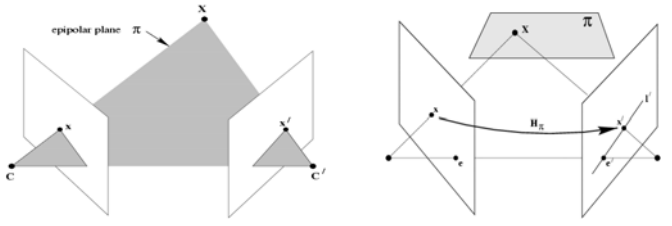


Fig. 5. The Epipolar Geometry and The Fundamental Matrix F

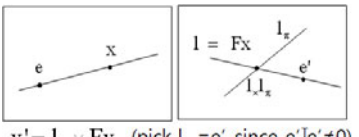
C, C' : camera center / x , x' : 2D points in Image / X : 3D points in real world

Epipolar Plane needs the matching points to seek the 3D stereoscopic video from 2D original video however; the method in this paper creates an optional matching point of 2D original video to express the three-dimensional effects. The objects and background of segmented video is divided from original video, then the divided video has an optional parameter value for moving left and right sides. The two mixed videos would have matching points and Matching Point Set. F can be sought by the matching point set above. The video which is generated by matching points above takes only real matching point value. In this paper, it takes 8 or 9 segmented objects instead point to seek the matching point. By using this value F, it creates a new image. Thus, we can produce stereoscopic images from 2D original video.

$$x' = H_\pi x$$

$$l' = e' \times x' = [e']_\times H_\pi x = Fx$$

mapping from 2-D to 1-D family (rank 2)



$x' = l_\pi \times Fx$ (pick $l_\pi = e'$, since $e'^T e' \neq 0$)

$H_\pi = e' \times F$ (compare with $F = e' \times H$)

The fundamental matrix satisfies the condition that for any pair of corresponding points x↔x' in the two images

$$x'^T Fx = 0 \qquad (x'^T l' = 0)$$

F is the unique 3x3 rank 2 matrix that satisfies $x'^T Fx = 0$ for all x↔x'

① Transpose: If F is fundamental matrix for (P,P'), then FT is fundamental matrix for (P',P)
② Epipolar lines: l'=Fx & l=FTx'
③ Epipoles: On all epipolar lines, thus e'TFx=0, x e'TF=0, similarly Fe=0
④ F has 7 d.o.f. , i.e. 3x3-1(homogeneous)-1(rank2)
⑤ F is a correlation, projective mapping from a point x to a line l' =Fx (not a proper correlation, i.e. not invertible)

Matching points set $(x_1, y_1), (x_2, y_2)$

Homogeneous coordinates

$$x' = \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} x \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}$$

$$x'^{T} Fx = (x^2, y^2, 1) \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = 0$$

Depth map is created through the phase of segmentation processing it is computed by using matching points, not whole depth map

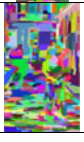
3 Real-Time 2D/3D Convergence System Based on Physical Board

This paper has a purpose to make 3D stereo convergence in real-time possible based on 2D Homogeneous Coordinate which is quite different from existing convergence algorithm. For the real-time 3D stereo convergence of singular original 2D video, I used hardware instead software to solve issues related with huge computational complexity and slow down of processing speed. The 3D stereo convergence algorithm in this paper, is impossible to expressed in real-time if it is based on the software. Therefore, construct the algorithm by hardware is best way to embody the real-time stereoscopic convergence. The processing speed of hardware-based real-time convergence is figured out as below.

Table 1. Hardware-based real-time three-dimensional conversion rate by Resolution (small number of objects

Resolution	Original Image	Segmentation Image	Rate
640X480			75μ/s
800X600			83μ/s
1024X768			107μ/s
1280X720			142μ/s
1920X1080			165μ/s

Table 2. Hardware-based real-time three-dimensional conversion rate by Resolution) (many of the objects)

Resolution	Original Image	Segmentation Image	Rate
640X480			90μ/s
800X600			101μ/s
1024X768			123μ/s
1280X720			157μ/s
1920X1080			180μ/s

S3D TV of SHARP, JVC, LG, SAMSUNG, HYUNDAI IT, ZALMAN, SONY supports the real-time stereo convergence function. Between the such products above, SHARP, LG, HYUNDAI IT, ZALMAN produce the products from same medium and small enterprise. This technology makes heavy distortion to the video by applying equal patterns to whole video instead using the algorithm in this paper. The restriction of watching TV would cause by such video distortion. TV manufactures produces the TV by using the pattern system, and it makes distortion disparity of the video. The proposed stereoscopic video convergence software which is based on Epipolar Geometry algorithm is better than others because of depth map extracted segmentation-based interpolation and distortion correcting. Moreover, such special features is been well received by many galleries from fair. However, the standardization of human factors are impossible to be quantified because of human's various eyesight and emotional information from brain.

4 Conclusion

I expressed the depth map algorithm for dividing objects and background and I also proposed Epipolar Geometry's 2D homogeneous Coordinate to convergence singular 2D original video to 3D stereoscopic video. Expressed depth map algorithm above produces the stereoscopic video in real-time based on hardware instead software in high quality by dividing objects and background. I concluded that it is impossible to solve the issues related with slow down and delay of processing speed if it is based on software. Moreover, I proposed simple system that uses a 3.5 inches HDD sized small hardware and the system that uses S3D TV and projector. Hardware-based real-time convergence system presents efficiency to media artists for producing stereoscopic video, because it gives three-dimensional effects to existing interactive video artworks by connecting hardware.

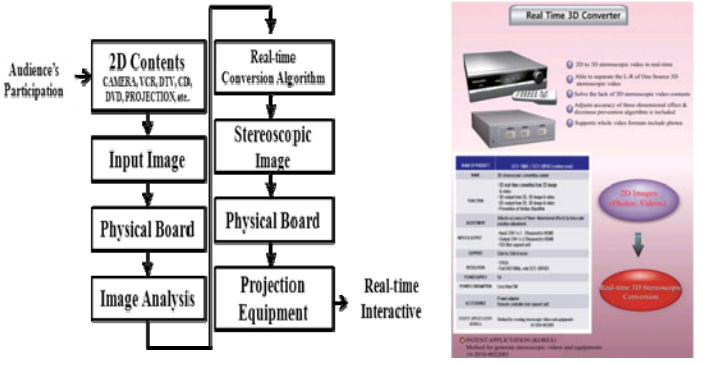


Fig. 6. Real-time Interactive 3D Art Work Process and 2D/3D Converter

This system presents better three-dimensional effect than existing commercial stereoscopic display. However, the 3D stereoscopic video which is produced by 2D video or S3D is quite different from real world, therefore, the visual fatigue may occurred. This visual fatigue should be considered until 3D hologram will be developed. Human's eyes move slightly during watching 3D video to settle the

Discord of convergence distance between the screen's focus and eye's view point. The visual fatigue is occurred caused by such heterogeneous reaction or eye fatigue above. According to statics, 10 to 20% of total population cannot recognize S3D, 3 to 5% of them cannot recognize 3D since when they born, because they are depth blind. Moreover, 15% of total is children or old people. In case of children, this phenomenon could be caused by lack of eyes' balance. In case of old people, it is caused by aging of eye muscle. Moreover, in case of chronic diseases, they could affected by heart attack because of S3D stimuli. In addition, it is difficult to perform the research on condition variables of visual fatigue for such people above, because it is connected directly to their life. The visual fatigue will grown up until the stereoscopic video would be produced under radical measures. These risks that are called human factors may be continued until present broadcast system goes to preproduction system such as film productions. 2D/S3D video convergence has issues such as quality, budget, time and human factor that is caused by visual fatigue. For the higher quality, it needs huge budget and lots of time. To deal with it, there is no way to produce low-quality videos, and it makes the production quality lower. The human factor-based stereoscopic video production technology and real-time 2D/3D convergence technology will reduce the time and costs, moreover, it will improve the quality of such artworks. In addition, the hardware-based real-time stereo convergence system presents efficiency in stereoscopic video production by using screen, computer and pc camera for the media artists. The simplex visual arts can be expressed faster by linking real-time stereo convergence system on the existing visual arts. Media artists prefer screen-based beam projector to generate their art production. The beam projector-based stereoscopic visual formats are suited in Top and Bottom or Side by Side systems. The real-time stereo convergence system in this paper is the best hardware-convergence system that shows good conditions in 3D TV production and 3D expression using beam projector.

For future researches, scientific and systemized studies are needed for stereo convergence system production which is applied visual recognizing human factors.

Fig. 7. Real-time 3D Converter and Real-time 3D Demonstration
Real-time 3D interactive photo mosaic art work, Gallery Now, Seoul, Korea, 2010.

Acknowledgment

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2011.

References

1. Kim, Y.: 2D-TO-3D Conversion of Single-View Video Based on View-Distance Classification. Thesis Master, Information and Communications University. Korea (2009)
2. Ideses, I., Yaroslavsky, L.P., Fishbain, B.: Real-time 2D to 3D video conversion. Journal of Real-Time Image Processiong 2(9-13) (2007)
3. Matusik, W., Pfister, H., Nagan, A., Beardsley, P., Ziegler, R., McMillan, L.: Image-Based 3D Photography using Opacity Hulls. In: ACM SIGGRAPH, pp. 427–437 (July 2002)
4. Iddan, G.J., Yahav, G.: 3D imaging in the studio (and elsewhere..). In: Proc. SPIE, vol. 4298, pp. 48–55 (January 2001)
5. Oike, Y., Ikeda, M., Asada, K.: Smart Access Image Sensors for High-Speed and High-Resolution 3-D Measurement Based on Light-Section Method. Intelligent Automation and Soft Computing 10(2), 105–128 (2004)
6. Je, C., Lee, S.-W., Park, R.-H.: High-Contrast Color-Stripe Pattern for Rapid Structured-Light Range Imaging. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 95–107. Springer, Heidelberg (2004)
7. Park, I., Kim, M., Byun, H.: Robust Zooming Motion Detection for Motionbased stereoscopic conversion. In: International Workshop on Advanced Image Technology (IWAIT 2007), pp. 944–948 (January 2007)
8. Feldman, M.H., Lipton, L.: Interactive 2D to 3D Stereoscopic Image Synthesis. In: Proc of SPIE, Stereoscopic Displays and Virtual Reality Systems XII, vol. 5664, pp. 186–197 (2005)
9. Tam, W.J., Speranza, F., Zhang, L., Renaud, R., Chan, J., Vazquez, C.: Depth image based rendering for multiview stereoscopic displays: Role of information at object boundaries. In: Proc. of the SPIE, vol. 6016, pp. 75–85 (2005)

The Changes in Performance during Stress-Inducing Cognitive Task: Focusing on Processing Difficulty*

Jinyoung Song, Taewon Chung, Jinwon Kang, and Kichun Nam

Department of Psychology, Korea University, Anam Campus, Anam-dong 5-ga, Seongbuk-gu, Seoul, Korea
{hyderay,blurryis,kasterran,kichun}@korea.ac.kr

Abstract. One stress-inducing task and two cognitive tasks were used to investigate how these tasks influence one's cognitive capability. First, the participants performed stress-inducing tasks for 5 minutes. They were then divided into two separate groups, and were given two kinds of cognitive load tasks which differed in difficulty levels. Finally, a cognitive task for evaluating cognitive capabilities of the two groups was conducted. This procedure was repeated 4 times, 2 minutes for each task, which showed that there was a tendency of improved performance over time. We found no reaction time and performance rate changes with respect to processing difficulty. However, the interaction between difficulty and repetition was significant, which meant participants who were assigned to easier tasks showed unstable performance compared to those who were assigned to high-difficulty tasks. These results suggest that imposing a proper amount of stress and cognitive load could promote cognitive performance in general.

Keywords: stress-inducing task, cognitive load, processing difficulty.

1 Introduction

We already know that stress have influences on people's cognition and behavioral performances. The cognitive load theory predicts that when proper amount of cognitive load is given, it can lead to a better performance [1]. We investigated how these two factors affect one's cognitive capability. Traditionally, experimenters used industrial noises, mental arithmetic, laboratory speech, public speech, cold-water to induce stress. In previous cognitive load studies, researchers classified the types of experimental manipulations into 5 demands; attention, response, familiarity, memory, processing change, processing difficulty [2]. In the present study, four specific computer programs were developed to induce stress, give cognitive load in processing difficulty demand and measure the cognitive ability.

* This research is supported by the KOCCA (Korea Creative content Agency) funded by the Korean Government [2-10-7602-001-10748-07-003, Technology of the Emotion Measurement and Standardization on Functional Contents].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 345–347, 2011.

2 Method

First we measured participants' cognitive capability for 5 minutes in order to use these data as a baseline. Then, participants completed stress-inducing task, cognitive load task, and cognitive capability evaluation task. Each task lasted for 5 minutes, 2 minutes, 2 minutes in each. This procedure was repeated 4 times. At the end of the fourth task, participants had to complete another task but we will not discuss it in details for that is included in a separate article.

2.1 Participants

27 paid Koreans whose age was 20 to 30 participated in this experiment. (16 in low-difficulty group and 11 in high-difficulty group).

2.2 Stimulus

Stress-inducing task. 100 English proverbs were selected.
Cognitive load task. 200 target words and 520 distractor words were used for both low and high-difficulty groups. All the words used in the task were Korean bisyllabic nouns. The target words for low-difficulty group shared the same first syllable, and target words for high-difficulty words were included in the same category. The distractor words were not semantically related to target words.
Cognitive capability evaluation task. 120 pairs of Korean sentences which have complicated syntactic structures were chosen. 60 pairs of sentences had the same meanings and 60 pairs of sentences had different meanings.

2.3 Procedure

Stress-inducing task. Participants typed English proverbs in a noisy environment for 5 minutes. They were frequently disrupted during the session, which may have caused some degree of irritation.
Cognitive load task. The task in low-difficulty group was to count the number of words in 4*4 matrixes which first syllable was identical to the target word. The high-difficulty group did similar task, but they counted words that were in the same semantic category as the target word. 3 to 5 target words were randomly positioned within the matrix. As soon as participants filled in the blank at the bottom of the screen by typing the correct number, the next stimulus was presented. Each task was continued for 2 minutes.
Cognitive capability evaluation task. A pair of sentences appeared on the computer screen, and participants made decisions whether two sentences had the same meaning or not. Depending on one's performance ability, maximum of 30 sentence-pairs were presented in each trial. This task also lasted for 2 minutes

3 Results

Performance rate and reaction time for cognitive capability evaluation was analyzed by 2(group)*5(trial) mixed design ANOVA. On the whole, performance gradually

improved over trial [$F_1(4, 100)=53.788^{***}$, $p<.001$], but neither performance rate changes [$F_1(1, 25)=1.866$, $p=0.184$] nor reaction time differences [$F_1(1, 15)=0.352$, $=0.553$] were observed between two groups. The high-difficulty group showed rapid changes in performance rate and reaction time at trial 2 and trial 4. [$F_1(4, 100)=5.470^{**}$, $p<.01$, $F_1(2.50, 62.49)=4.141^{*}$, $p<.05$] There was no error rate difference in all conditions.

Table 1. Performance Rate (%), Error Rate (%) and Reaction Time (ms) Changes for Repetition in low-difficulty (Low) and high-difficulty (High) groups

	Performance rate (%)		Error rate (%)		RT (ms)	
	Low	High	Low	High	Low	High
Trail 1	45	66	11	11	8374	6528
Trial 2	48	59	13	6	7493	6667
Trial 3	63	79	7	6	6010	4922
Trial 4	72	74	13	10	4966	5030
Trial 5	90	89	12	10	3165	3663

4 Conclusion

In the present study, we could infer that cognitive capability did not change as a result of processing difficulty. We used linguistic materials such as words and sentences to adjust the difficulty of two tasks, but it turned out to be ineffective as opposed to previous research [3]. If we had simply compared the results of two cognitive load tasks in difficulty, the similar results would have been obtained. However, our experimental analysis focused on the cognitive capability evaluation task which is followed by the cognitive load task, and rather complex results were revealed. The results revealed that proper amount of stress and cognitive load clearly boosts one's performance regardless of the difficulty of the task. Furthermore people who took part in higher level task showed unstable performance.

References

1. Pass, F., Tuovinen, J.E., Tabbers, H., Van Gerven, P.W.M.: Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. Edocational Psychologist 38(1), 63–71 (2003)
2. Block, R.A., Hancock, P.A., Zakay, D.: How cognitive load affects duration judgments: A meta-analytic review. Acta Psychologica 134, 33–343 (2010)
3. Block, R.A., Reed, M.A.: Remembered duration: Evidence for a contextual change hypothesis. Journal of Experimental Psychology: Human Learning and Memory 4, 656–665 (1978)

The Changes in Performance during Cognitive Load Inducing Tasks*

Heejin Park, Yoonkyung Choe, Yoonhye Na, and Kichun Nam

Department of Psychology, Korea University, Anam Campus, Anam-dong 5-ga, Seongbuk-gu, Seoul, Korea
{littlegirl-5,yoonkc86,kikongza,kichun}@korea.ac.kr

Abstract. This study examined the performance of one's cognitive abilities during various different mental cognitive load tasks. A stress-inducing task was administered during the course of performing three different kinds of cognitive load tasks. The three types of cognitive load demands included attention, memory and processing change. Participants in this study performed two tasks for each demand. Error rates, number of completed trials, and response times of all three types of cognitive load demand tasks were measured. The results suggest that for attention demand task, the response time decreased as time increased whereas no difference was found in memory demand task. In processing change demand task, the number of completed trials increased as the response time increased. This study shows a change in the performance of cognitive tasks depending on the type of cognitive load, which in turn suggests the importance of implementing different kind of methods and measuring tools when assessing cognitive abilities in individuals.

Keywords: stress-inducing task, cognitive load.

1 Introduction

We are often exposed to work environments which we get stressed easily. Several cognitive abilities are needed in order to perform the tasks and these abilities can work as cognitive loads as well. According to Sweller, cognitive loads occur when the information to-be-learned is unable to process in the working memory because of its limits, and results in a loss or overload of information [1]. Among many existing cognitive loads, the cognitive loads in this study were used by Richard [6]. In his previous study, he suggested six dimensions of cognitive load. Among the six, three cognitive load demands related to work environments were selected in this study. The study examined the performance of one's cognitive abilities during various different mental cognitive load tasks in stressful environment as well as in environments which require one's cognitive load on their performance.

* This research is supported by the KOCCA(Korea Creative content Agency) funded by the Korean Government [2-10-7602-001-10748-07-003, Technology of the Emotion Measurement and Standardization on Functional Contents].

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 348–351, 2011.

2 Experiment 1

2.1 Method

Participant. 24 people who are ages between 20 to 30 year olds participated in the attentional demand of cognitive load task after performing stress-inducing task.
Procedure. After performing stress-inducing task, the participants were given cognitive load tasks. They then were given tasks that were used to measure cognitive abilities of the participants. The participants were instructed to detect letter "Z" among 6 by 25 arrangements of alphabet in the cognitive loads task. They were then instructed to find blue triangles in conditions which blue or red triangles and squares were presented to distract the target. The number of distractor has increased as the number of trials increased.

2.2 Result

The dependent measures, error rate, number of completed trials and response times were measured and analyzed. The effect of exposure time to stress-inducing task was not significant in error rates ($F(2.216,50.968)=.748$, $p>0.05$), but significant in completed number of trials ($F(2.928,67.353)=9.590$, $p<0.0005$) and response time ($F(2.889,66.443)=9.657$, $p<0.0005$). Therefore, as the exposure time of stress inducing task increased the number of trials also increased but reaction time of cognitive ability has decreased.

3 Experiment 2

3.1 Method

Participant. 24 people who are ages between 20 to 30 year olds and did not participated in experiment 1 performed the cognitive load of memory demand task after participating stress-inducing task.

Procedure. The same procedure is used as in experiment 1. The task which gives participants cognitive loads included 2-back task in this dimension. The numbers ranging from 1 to 4 were presented on the screen in random order. The task required people to match with the number presented in second order from the numbers presented in random. The cognitive ability was measured with backward-digit span task.

3.2 Result

The error rate, number of completed trials, response times and memory span were measured and analyzed. The effect of increase in exposure time to stress was not significant in all four measurements; error rates ($F(2.289, 52.648)=.358$, $p>0.05$), number of trials ($F(2.497, 57.427)=1.827$, $p>0.05$), response time ($F(1.473,33.890)=3.579$, $p>0.05$) and memory span ($F(2.491, 57. 288)=2.174$, $p>0.05$). Therefore, memory demands are not affected by the times exposed to stress.

4 Experiment 3

4.1 Method

Participants. 25 participants who are ages between 20 to 30 year olds and did not participated in experiment 1and 2 completed the task of processing change demand of cognitive loads after performing stress inducing task.

Procedure. The same procedure is used as in experiment 1 and 2. This dimension of load task required people to perform correctly according to each different instruction in the tasks. In the task A, the participants had to make a choice of "up" or "down". If the presented number was bigger than "5", they pressed "up" and pressed "down" when it is smaller than "5". The paradigm of task B was the same but they had to choose "up" and "down" according to number of presented numbers. After performing this cognitive loads task, another task was given to measure their cognitive ability. In this task, participants pressed the same directional key as presented on the screen when the arrow was in red. On the other hand, they pressed the opposite directional key when the blue arrow was presented.

4.2 Results

Same as in two previous experiments, error rates, number of completed trials and response time were analyzed. In result, the effect of increase in exposure time to stress was not significant in error rates ($F(4, 96) = 1.001$, $p > 0.05$) and response times ($F(1.054, 25.292) = 3.870$, $p > 0.05$) but statistically significant in number of completed trials ($F(2.018, 48.4343) = 13.099$, $p < 0.0005$). According to these results the number of completed trials increased as people were exposed longer to stress inducing task in processing change demand.

5 Conclusion

The study examined the performance of one's cognitive abilities during various different mental cognitive load tasks in stressful environment as well as in environments which require one's cognitive load on their performance. On that account three categories of demand were introduced, attentional demands, memory demand and processing demands. As a result of this study, the number of trials also increased but reaction time of cognitive ability decreased as people were exposed longer to stress inducing task in attentional demands. However, no significant change appeared in memory demands. The results in processing change demand showed that the number of completed trials had increased as people were exposed longer to stress inducing task. From present study we can infer that the effect of stress influences differently in different demands of our cognition. This study used only one type of task in accordance of each demand but it is recommended to use different types of task for further study.

References

1. Sweller, J.: Cognitive Load Theory, learning difficulty, and instructional design. Learning and Instruction 4, 295–312 (1994)
2. Schoofs, D., Preuß, D., Wolf, O.T.: Psychosocial stress induces working memory impairments in an n-back paradigm. Psychoneuroendocrinology, 33, 643–653 (2008)
3. Neisser, U.: Visual Search. Scientific American 210(6), 94–102 (1964)
4. Richard Clarka, C., Veltmeyera, M.D., Hamiltonb, R.J., Simmsb, E., Paul, R., Hermens, D., Gordon, E.: Spontaneous alpha peak frequency predicts working memory performance across the age span. International Journal of Psychophysiology 53, 1–9 (2004)
5. Armony, L., Greenshpan, Y.: Switching tasks and attention policies Daniel Gopher. Journal of Experimental Psychology: Oeneral 129(3), 308–339 (2000)
6. Block, R.A., Hancock, P.A., Zakay, D.: How cognitive load affects duration judgments; A meta-analytic review. Acta Psychologica 134, 330–343 (2010)

Emotion Based Gesture Animation for Mobile Phone

Haewon Byun

303 Medial Infirmation buil, 169-1
Don-sun dong 2nd street, SeongBukGu,
Seoul, Republic of Korea
hyewon@sungshin.ac.kr

Abstract. Recently people do not only use mobile phone to call, but also use it to send a SMS. However, it is difficult to express own complicated emotion with text and emoticon of exited SMS service. We pay attention to express user's emotion interesting and correct, we use character animation. This paper suggests emotion based gesture animation generation system that uses character's facial expression and gesture to delivery emotion excitably and clearly. Michael[1] investigated interview of two people who has stylized gesture. They suggested gesture generation graph for stylized gesture animation. In this paper, we focus to analyze and extract emotional gestures of Disney animation characters and create 3D models of extracted emotional gestures. To express emotion of person, we use emotion gesture generation graph that import the emotion flow graph that expresses emotion flow for probability. We investigated user reaction for research proprieties of suggested system and the alternation propriety.

Keywords: Mobile phone 3D animation, Emotional gesture generation, Emotion flow, Gesture profile, Mobile phone 3D engine.

1 Introduction

Now mobile phone is more than a creature of high-technology for people. Mobile phone makes people connecting to communicate with other person. People do not only use mobile phone to call, but also use it to send a SMS. However, SMS is formed of only text and emoticon, so it has limitation to express atmosphere and nuance. In modern societies, expressing individual character is very important; therefore the desire to express emotion became stronger. But, current SMS services standard and limited tools to express detail of emotion perfectly. In this paper, we suggest using character animation for SMS service to communicate user's idea dynamical and correctly. We tried to create character animation that express emotion by facial expression and body gesture. Michael et al suggest character animation recreation system that annotated interview video to built stylized gesture animation such as standard gesture of an announcer.

Michael et al focused on characteristic gesture; on the other hand we focused on that SMS service is more used to expressing emotion than to send informative contents, we designed system that generates expressing emotion gesture animation to notify user's emotion at a glance. We annotated and analysis character's gesture of

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 352–356, 2011.

Disney animation to use exaggerated gesture. Based on this analysis, we recreated gesture animation that represents SMS message. Experimental results demonstrate the effectives of our approaches.

2 System Architecture

We suggest the gesture generation system to express emotion of SMS messages as character gesture and facial expression animation and to service on a mobile phone. The system is constructed of emotion recognition, gesture profile, emotion gesture generation and mobile phone engine (see Figure 1). When user put the text to mobile phone, the system recognizes emotion of text at server and generates final gesture and facial expression based on recognized emotion. Generated data from server system is transmitted to mobile phone 3D engine and play animation on the mobile phone.

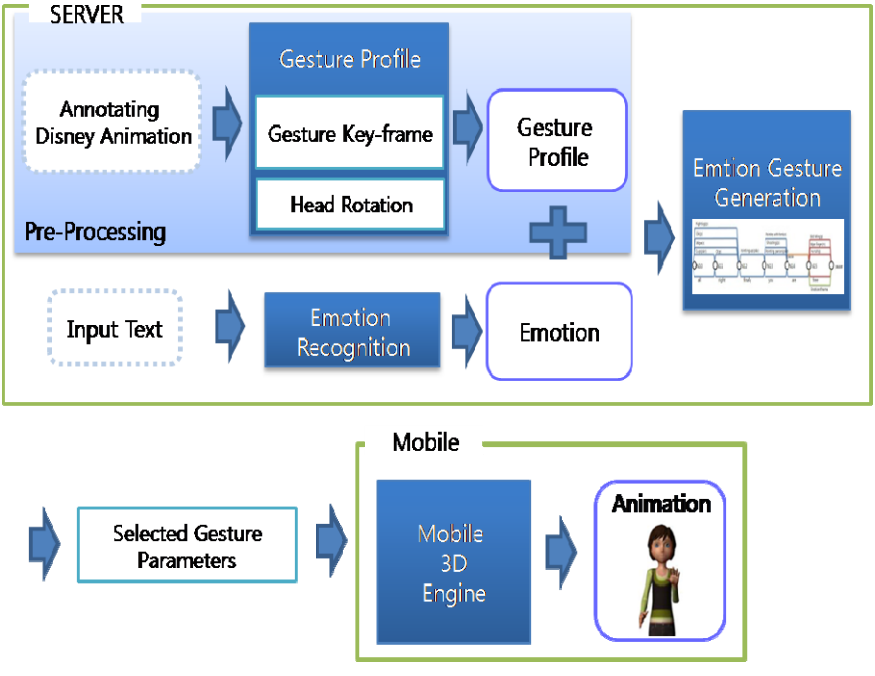


Fig. 1. System architecture

Emotion recognition system: This system recognizes 8 emotions (neutral, happy, sad, angry, surprise, afraid, disgust) from English text. We used open source Synesketch that recognize emotion from text and animate colors and shapes represented the recognized emotion. We defined emotion words that decided emotion of entire sentence from Disney animation and add emotion recognition system's database.

Gesture profile. Gesture profile is gestures data classified by noun and emotion to generate animation that express emotion of text. We analysis and annotate Disney animation's gesture and extract gestures that has each emotion. Also head rotation, feature of animation and time information are saved to gesture profile for animation generation.

Emotion gesture generation. Emotion gesture generation proceed by emotion gesture generation graph that selects final gestures to animate from gestures that transmitted form gesture profile. Create node with inputted text's noun and make link with final gesture classified by emotion. For the last, linked gestures are priority sorted by emotion that emotion recognition system recognized.

Mobile phone 3D engine. The role of mobile phone 3D engine is to create 3D character animation based on final gesture that selected by emotion gesture generation graph. Final selected gesture parameters are transmitted to mobile phone 3D engine and do the blending with facial animation. Finally, creates emotion based character animation and service on the mobile phone.

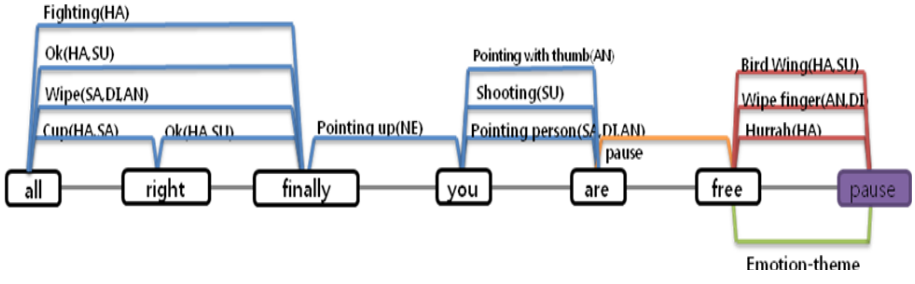


Fig. 2. Emotion Gesture Generation Graph

3 Mobile Phone 3D Engine

The role of mobile phone 3D engine is to create the final animation by transmitting final gesture data from emotion gesture generation system at server and composing with facial animation. To generate 3D character animation, do modeling gesture key-frames that transmitted from gesture profile by MAYA to OBJ files. For real-time playing at the mobile phone, we do modeling low polygon models.

During the models, we load the models chosen by emotion gesture graph to mobile phone engine and compose with facial animation. Arrange a low with key-frame and blending key-frame to generate animation. We use Google Android as emulator and our system is manufactured by java and OpenglES.

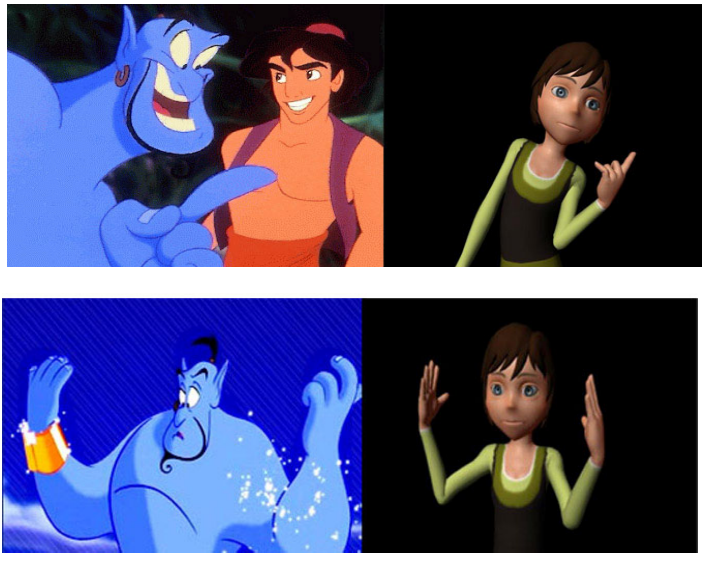


Fig. 3. Disney animation 3D modeling

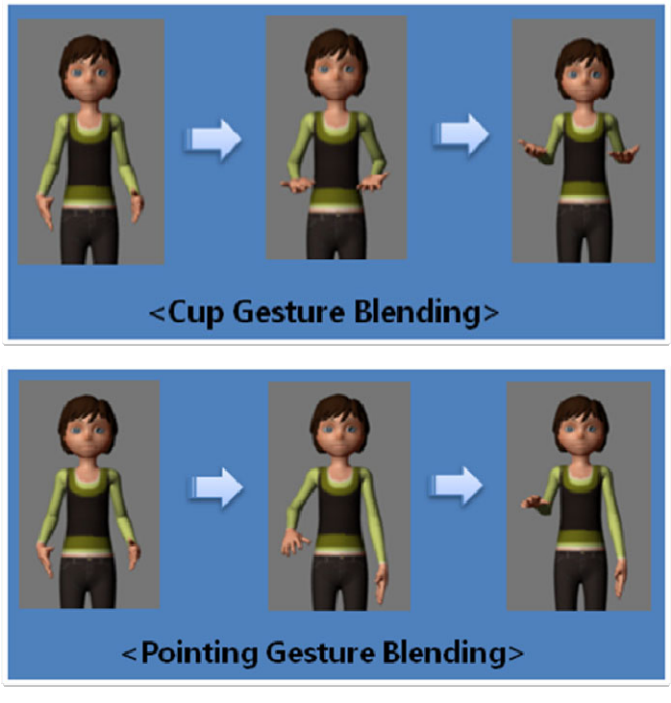


Fig. 4. Blending the Gestures

Fig. 5. Emulate at Google Andriod

4 Conclusion

This work presents a system for generating gesture expressing emotion for mobile phone SMS service. We use Disney animation to extract exaggerate gesture for noun, emotion. We organized data as gesture profile and use graph structure to select gesture based on recognized emotion.

To select the best suitable gesture for emotion, we define and import 'Emotion Flow Graph' through Disney animation character's emotion. Selected gesture's parameters are sanded to mobile phone and are blended with facial animation to generate animation. We concentrate to enhance user's satisfaction and interest by providing more realistic and emotional conversation animation.

Acknowledgments. This work was supported by the Korea Research Foundation (KRF) grant funded by the Korea government(MEST) (No.2010-0017246)

References

1. Michel, N., Michel, K., Irene, A., Hans-Peter, S.: Gesture Modeling and Animation Based on a Probabilistic Re-Creation of Speaker Style. ACM Transactions on Graphics 27 (2008)
2. Computer generated animation of faces. In: Proceedings of the ACM Annual Conference. ACM Press, New York
3. Massaro, D.W., Cohen, M.M.: Visible Speech and Its Potential Value for Speech Training for Hearing-Impaired Perceivers. Bibliographic Reference, In Still, 171–174 (1998)
4. Kalberer, G.A., Luc, J., Van, G.: Realistic face animation for speech. Journal of Visualization and Computer Animation (2002)
5. Garcia-rojas, A., Vexo, F., Thalmann, D., Raouzaiou, A., Karpouzis, K., Kollias, S.: Emotional Body Expression Parameter In Virtual Human Ontology. IEEE org (2005)
6. Cassell, J., Vilhj Almsson, H., Bickmore, T.: BEAT: The Behavior Expression Animation Toolkit. In: Proceedings of SIGGRAPH, pp. 477–486
7. Lisetti, C.L., Gmytrasiewicz, P.: Can a Rational Agent Afford to Be Affectless? Applied Artificial Intelligence 16(7-8) (2002)

Interaction for Interactive Popup Book

Haewon Byun

303 Medial Infirmation buil, 169-1
Don-sun dong 2nd street, SeongBukGu,
Seoul, Republic of Korea
hyewon@sungshin.ac.kr

Abstract. Recently popular media, book, has problems to satisfy a child who wants to be a character in a story, because it just transmits fixed story. Therefore, many people are interested in interactive contents that are changed variously by user's acting in multimedia division. This paper introduces an interactive pop-up book system that can control a story of book by interesting interactivity. This system suggests a special story graph that transforms a linear architecture story to a interactive story by traversing diverse node. We focus to make a special pop-up book interface and wind blowing interface for physical environments to have interesting and familiar user interaction, and to conclude by user studies. We also investigated user preference for various interactive story paths, so we can exclude unnecessary stories and surveyed to compare before the excluding and after. Finally, story paths that user prefers are introduced in the Interactive Popup Book System.

Keywords: Mobile phone 3D animation, Emotional gesture generation, Emotion flow, Gesture profile, Mobile phone 3D engine.

1 Introduction

The generation, new media appears and combines with each other, begins with rapid growth of media. One of the media, book, is improved to new media by importing pop up contents. Pop-up book expresses characters or objects for drawing instead words for story. People who are looking a pop-up book feel interesting and fresh, so they can concentrate on the story and imagine the story easily.

However, there are problems to maintain children's attention by only pop-up contents, recently online fairy tale animations are produced a lot. At online, it is possible to make animation contents fantastic by using variable colors and cute characters, so it is really effective for children education. However, online fairy tale animation has linear architecture of story, so most of children do not watch it over one time.

This paper suggests interactive pop-up book system that combines pop-up book, which has interesting pop-up contents, and online animation contents, which can express various stories. This system offers special experiences that even user can see pop-up contents through the book; also user can see the animation contents by pressing switch on the pop-up contents or turning the page. For this system, we suggest a method to translate a linear architecture story that pop-up book has, to an interactive architecture story. We also suggest a special poop-up book interface for

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 357–359, 2011.

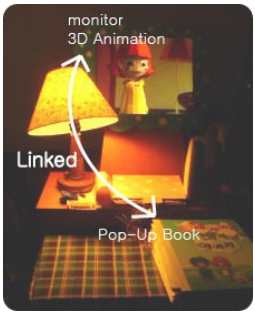


Fig. 1. Interactive Popup Book System

physical environment to have familiar and interesting interaction with children. We had exhibition at 2006 graduation exhibition of Sungshin Women University and researched user study.

2 System Architecture

Purpose of interactive pop-up book system is that suggests new media like fig.2 for combining existing offline book and online fairy tale animation. Scenes and objects compose pop-up book for physical interactive environment. Fairy tale animation connected with pop-up book has user interactions and has multiple architecture of a story that has same 3D scenes and characters with pop-up book.

Interactive pop-up book system is consisted of interactive storytelling, pop-up book interface and network module.

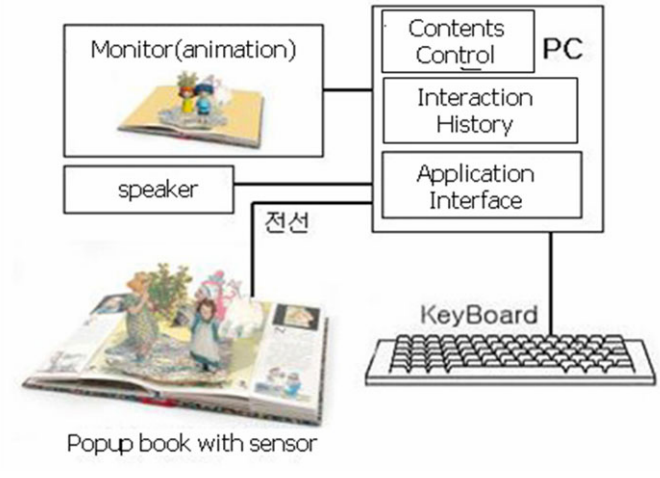


Fig. 2. System Overview

Interactive story telling: we do modeling interactive story contents that are connected with linear story of pop-up book. At this part, system makes a story graph that has nodes, minimum unit story bit. Finally traversing graph nodes completes interactive story.

Pop-up book interface: interactive pop-up book as a special madden interface transmits interaction to the computer such as turning pages or pressing buttons. It is easy to manufacture an interface, because we apply the theory of keyboard and use it. Everyone who has a computer can connect a new pop-up book interface and enjoy it.

Network module: Making an addition of network module that makes it possible to send new stories whenever, wherever embody ubiquitous media. Adding new stories periodically can suffer evolutional contents. Network module also can be a method to share interactive contents that individually has.

Acknowledgments. This work was supported by the Korea Research Foundation (KRF) grant funded by the Korea government(MEST) (No.2010-0017246)

References

1. Michael Young, R.: An Overview of the Mimesis Architectures: Integrating Narrative Control into a Gaming Environment. In: AAAI (2001)
2. Cavazza, M., Charles, F., Mead, S.J.: Character-Based Interactive Storytelling. IEEE Intelligent Systems 17(4) (2002)
3. Blom, K.J., Beckhaus, S.: Emotional Storytelling. In: IEEE Virtual Reality 2005 Conference, pp. 23–27 (March 2005)
4. Riedl, M.O., Michael Young, R.: Form Linear Story Generation to Branching Story Graphs. IEEE Computer Graphics and Applications 26(3) (2006)
5. Billinghurst, M., Kato, H., Poupyrev, I.: The Magic Book- Moving Seamlessly between Reality and Virtuality. IEEE Computer Graphics and Applications 21(3) (2001)
6. Kilbourn, K., Sitorus, L., Zupan, K., Hey, J., Gauziskas, A., Stiensta, M., Andresen, M.: Exploring narratives for physical play: A pop-up guide to an interactive playground. In: Göbel, S., Spierling, U., Hoffmann, A., Iurgel, I., Schneider, O., Dechau, J., Feix, A. (eds.) TIDSE 2004. LNCS, vol. 3105, pp. 240–245. Springer, Heidelberg (2004)
7. Camarata, K., Do, E.Y.-L., Gross, M.D.: Computationally-Enhanced Books, CMU, CoDe Lab working paper (2005)

The Mobile ID Based Biometric Undeniable Signature Scheme

SungHyun Yun

Division of Information & Communication Engineering,
Baekseok University, Cheonan, Chungnam, Korea
shcrpt@gmail.com

Abstract. The wide use of mobile devices such as smart phones makes the mobile commerce industry be growing-up rapidly. In mobile commerce security, how to secure a copyright of mobile contents and how to distribute it are of major concern. The user can carry the smart phone regardless of the places. The utilization of a smart phone is very high than that of personal computers. The USIM(Universal Subscriber Information Module) inserted in the smart phone binds the user with the device. This means that the smart phone can be used to represent the owner's identity. In this paper, the mobile ID based biometric undeniable signature scheme is proposed. The mobile ID is created with USIM and the user's biometric template. In addition, undeniable property of our signature scheme can make ID based applications such as mobile voting be possible with smart phones.

Keywords: Mobile Security, Biometric Signature, Undeniable Signature, Mobile ID, ID based Signature Scheme.

1 Introduction

The wide use of mobile devices such as smart phones makes the mobile commerce be of major concern today. In mobile commerce, the business transaction is carried out with users who own mobile devices such as smart-phones operable with wireless network. Since the user can carry smart phone regardless of places, the usage rate of the smart phone is very high compared to personal computers. The mobile commerce market is expected to be growing-up larger than that of PC based e-commerce [2].

User authentication means proving one's identity to others. In general, ID/Password based methods are widely used in user authentication. This approach is based on the user's ability to remember things only available for himself/herself. However, it has drawback that the proxy authentication is possible. The authentication scheme based on the user's memory does not physically binding user's identity. Thus, the scheme that binds biometric data to the user's identity is required. The public key and the private key of the user should be generated from identity information including biometric data. To sign a document, the signer (the key owner) should present his/her own biometric data every time to get the private key.

Identity based cryptography has been introduced by Shamir in 1984 to simplify key management and avoid the use of digital certificates [3]. The goal of this approach is

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 360–366, 2011.

to let a public key be publicly and uniquely derivable from a user's identifiable things such as e-mail address, IP address, social security number, etc. The associated private key can only be computed by a PKGA(Private Key Generating Authority). The PKGA server serves a similar role to the CA in a PKI system. This approach allows bypassing the trust problems that arise in traditional CA based PKIs. In ID based signature scheme, since the public key represents owner's identity, it becomes useless to bind it by the digital certificate. Although a PKG's public key still has to be certified, the need of digital certificates is really reduced as reasonably many users may depend on the same PKG [4, 5].

The advantage of using ID based scheme is that we don't need a copy of someone's certificate in order to encrypt or verify a message. We can compute their public keys knowing only their identity and which PKGA they use. This means that there's no need for a certificate directory, which has historically been one of the inconvenient parts of PKI [5].

The drawback of existing ID based scheme is that the privacy related information such as social security number, e-mail address, … are easily accessible from all sources since the public key is generated from ID itself. Thus, we need privacy protected ID such as i-Pin(internet Personal identification number) managed by trusted authority. It can be used to represent user's identity in various authentication services.

In this paper, we consider what properties should be satisfied to make the mobile ID. The mobile ID should be available in commercial services such as e-voting, digital content distribution, … Biometric data source is needed to bind the user to the ID to minimize the illegal use of the mobile ID and related digital keys. The trusted ID management center is also needed. Its role is registration of mobile ID, revocation of it and management of ID related services. The mobile ID should be registered with legal process. How can we make the legal ID ? To use the smart phone, people must register their USIMs by submitting social security number off-line in trustful registration center. Therefore, the mobile ID created with USIM can represent the user's legal identity. For revocation and renewal of mobile ID, a random secret value of the user is needed.

There are many business transaction models on how to distribute products between customers and sellers. The proposed undeniable signature scheme can be applied to on-line contents business model where the digital contents are distributed among authors, sellers and customers. The contents such as music, video, animation, apps, … are available in mobile devices to play by downloading them via wire/wireless network [1, 2].

In this paper, the biometric undeniable signature scheme based on mobile ID is proposed. The mobile ID is created with smart phone user's USIM, a random secret value and the biometric template. The USIM inserted in the smart phone binds the user with the device. It requires registration stage where the user must submit one's identity information such as social security number card to the trustful registration authority by off line. Therefore, the smart phone can be used to represent the owner's identity. The wide use of smart phones and growing number of users make ID based cryptographic schemes useful in real life applications. Thus, it's important to study methods of how to make mobile ID and cryptographic digital keys.

The biometric data is unique to the individual. To minimize the risk of signing by a proxy agent, a biometric based digital key is needed to sign the digital data. However, the digitized biometric data such as a fingerprint template is stolen, the biometric data cannot be reused for authentication purposes. For privacy and practical use, the biometric template should then be cancelled and regenerated from the same biometric data [8]. Thus, the random secret value is added to make the mobile ID in this study. By changing the random value, the mobile ID can be updated on-line.

The proposed scheme is consisted of four stages. Those are biometric key generation, signature generation, signature verification, and disavowal stages. It has an undeniable property in which the signature cannot be verified without help of the signer. Undeniable property of our signature scheme can make ID based applications such as mobile voting be possible with smart phone users.

In section 2, mobile ID based biometric key generation steps are presented. In section 3, the proposed undeniable signature scheme is described. In section 4, undeniable property is analyzed with disavowal protocol.

2 Digital Key Generation

In this section, we show how to make mobile ID and related digital signature keys. We assume that the trustful MC(Mobile ID Authentication Center) exists. Its role is to manage ID registration, authentication and revocation of it.

Definition 1. *p is a large prime number and GF(p) is a cryptographically secure finite field. g is a generator defined in GF(p) having order p-1 [6]. If these properties hold, it's computationally infeasible to solve discrete logarithms over GF(p).*

The followings are public key and private key pair of MC.

$$sk_{MC} \in Z_{p-1},\ pk_{MC} \equiv g^{sk_{MC}} \pmod{p} \tag{1}$$

A smart phone user makes his/her own mobile ID as follows.

Step 1: The user authenticates himself/herself to the MC off-line and registers smart phone's USIM.

Step 2: The mobile ID is created with the user's USIM, a random secret value and the biometric template.

$$\begin{aligned} &ID_U : The\ user\ U's\ ID, \quad H : Hash\ Function \\ &USIM_U : ID_U's\ USIM, \quad pc_U : ID_U's\ passcode, \quad BT_U : ID_U's\ Biometric\ Template \\ &ID_U = H(USIM_U \| pc_U \| BT_U) \in Z_{p-1} \end{aligned} \tag{2}$$

Step 3: The user makes his/her own private key and public key as follows. The private key should not be stored in smart phone for minimizing the risk of disclosure. Thus, the user should always input password to make the private key.

$$sk_U = H(ID_U \| pw_U \| BT_U) \in Z_{p-1},\ pk_U \equiv g^{sk_U} \pmod{p} \tag{3}$$

Step 4: The user signs mobile ID using the private key and encrypts it with the MC's public key. The user sends it to the MC.

Step 5: MC decrypts the encrypted mobile ID and verifies the user's signature on it. If verification succeeds, the MC signs the user's mobile ID. Then, the MC keeps a copy of the user's signed ID and sends another copy to the user.
Step 6: The user receives the signed ID using smart phone and verifies the signature of the MC.

3 The Undeniable Signature Scheme

In this section, the undeniable signature scheme is presented. The proposed scheme consists of signature generation and confirmation protocols. The undeniable signature cannot be verified without help of the signer. The signer must participate in the process of signature verification.

3.1 Signature Generation Protocol

In the proposed scheme, the following signature equation is used. Cryptographically secure GF(p) is defined in definition 1 and the security of the proposed scheme depends on discrete logarithms problem [6].

$$k_U \cdot (H_U + S_U) \equiv sk_U \cdot R_U \pmod{p-1} \tag{4}$$

Step 1: The signer generates random number k_U relatively prime to p-1 and keeps it secret. H_U is made by hashing the message M with padding bits PD. PD is randomly generated to make the H_U as primitive root over GF(p).

$$\gcd(k_U, p-1) = 1, \quad H_U = H(M \| PD) \tag{5}$$

Step 2: The signer creates the R_U with the random number k_U and the hash result H_U.

$$R_U \equiv H_U^{K_U} \pmod{p} \tag{6}$$

Step 3: The signer generates the undeniable signature S_U as follows. To compute the private key sk_U, the signer should submit his/her own biometric data and password.

$$k_U \cdot S_U \equiv sk_U \cdot R_U - k_U \cdot H_U \pmod{p-1} \tag{7}$$

Step 4: The signer sends (M, PD, R_U, S_U) to the verifier.

3.2 Signature Confirmation Protocol

Step 1: The verifier receives (M, PD, R_U, S_U) from the signer. Then the verifier calculates hash value H_U as follows.

$$H_U = H(M \| PD) \tag{8}$$

Step 2: The verifier selects random number (r_1, r_2) and makes the challenge value CH_1 as follows.

$$CH_1 \equiv R_U^{r_1 \cdot (H_U + S_U)} \cdot pk_U^{R_U \cdot r_2} \pmod{p}, \quad (r_1, r_2) \in Z_{p-1} \tag{9}$$

Step 3: The verifier sends challenge CH_1 to the signer.
Step 4: The signer generates the response RP_1 as follows and sends it to the verifier. To compute the private key sk_U, the signer should submit his/her own biometric data and password.

$$RP_1 \equiv CH_1^{sk_U^{-1}} \pmod{p},\ sk_U \cdot sk_U^{-1} \equiv 1 \pmod{p-1} \tag{10}$$

Step 5: The verifier verifies the response RP_1 as follows. If the signature (R_U, S_U) on the message M is valid, the equation (11) holds.

$$\begin{aligned} RP_1 &\equiv H_U^{R_U \cdot r_1} \cdot g^{R_U \cdot r_2} \pmod{p} \\ RP_1 \equiv CH_1^{sk_U^{-1}} \pmod{p} &\equiv (H_U^{k_U \cdot (H_U + S_U) \cdot r_1} \cdot g^{sk_U \cdot R_U \cdot r_2})^{sk_U^{-1}} \pmod{p} \\ \equiv (H_U^{sk_U \cdot R_U \cdot r_1} \cdot g^{sk_U \cdot R_U \cdot r_2})^{sk_U^{-1}} \pmod{p} &\equiv (H_U^{R_U \cdot r_1} \cdot g^{R_U \cdot r_2}) \pmod{p} \\ (\because k_U \cdot (H_U + S_U) &\equiv sk_U \cdot R_U \pmod{p-1}) \end{aligned} \tag{11}$$

4 Disavowal Protocol

In this section, undeniable properties of the proposed scheme are analyzed. In section 4.1 the disavowal protocol is described and in section 4.2 the conversion method that makes the undeniable signature to the ordinary one is shown.

4.1 Disavowal Protocol

If equation (11) holds, the verifier ensures that the signature is valid. Otherwise, the verifier launches disavowal protocol to identify whether the signature is invalid or the signer has cheated as follows.

The second challenge and response protocol launched by the verifier is as same as that of the signature confirmation protocol in section 3. The verifier chooses two random numbers (r_3, r_4) and computes the second challenge CH_2 as follows.

$$CH_2 \equiv R_U^{r_3 \cdot (H_U + S_U)} \cdot pk_U^{R_U \cdot r_4} \pmod{p},\ (r_3, r_4) \in Z_{p-1},\ r_1 \cdot r_4 \neq r_2 \cdot r_3 \pmod{p-1} \tag{12}$$

Step 4: The signer sends the response RP_2 to the verifier.

$$RP_2 \equiv CH_2^{sk_U^{-1}} \pmod{p} \tag{13}$$

Step 5: The verifier makes the following discrimination equations and compares R_1 with R_2. If R_1 equals to R_2, the signature is invalid. Otherwise, the signer has cheated on the valid signature.

$$R_1 \equiv (RP_1 \cdot g^{-R_U \cdot r_2})^{r_3} \pmod{p},\ R_2 \equiv (RP_2 \cdot g^{-R_U \cdot r_4})^{r_1} \pmod{p} \tag{14}$$

4.2 Conversion Protocol

The signer can make undeniable signature to the ordinary one that has self-verification property. The signer releases the partial secret value G and the verifier can verify the converted signature by himself/herself as follows.

$$G^{H_U+S_U} \equiv g^{sk_U \cdot R_U} \equiv pk_U^{R_U} \pmod{p},\ G = g^{k_U},\ H_U = H(M \| PD) \tag{15}$$

In equation (15), if k_U is disclosed, the private key sk_U is easily computed. However, to get the k_U from G, we must solve discrete logarithms over GF(p) as follows.

$$k_U \equiv \log_g G \pmod{p} \tag{16}$$

From the definition 1, it's computationally infeasible to solve discrete logarithms over cryptographically secure GF(p). Therefore, the security of the converted signature is as same as ordinary signature schemes based on discrete logarithms [6].

5 Conclusion

In this paper, the biometric undeniable signature scheme based on mobile ID is proposed. The mobile ID is created with the user's USIM, biometric template and a random secret value. MC is assumed to be trustful center to register, authenticate, revoke and renew mobile ID. The proposed ID can represent the user's legal identity since the USIM is registered with legal off-line process by trustful registration authority. The use of biometric template to generate and get the private key can minimize the risk of proxy signing. If the public key generated directly from the user's ID such as existing ID based signature scheme, it can give rise to privacy related problems. Therefore, renewable mobile ID is needed to replace the legal ID for commercial applications. A random secret value is used to revoke and renew the mobile ID.

In the proposed scheme, the private and public keys are generated from the mobile ID. Then, the signature generation and confirmation protocols are described. It also describes disavowal protocol and a method to convert undeniable signature to the ordinary one. In future work, the study on applications of the proposed scheme is needed. The proposed undeniable signature scheme is best suited to applications such as e-voting where the user's ID should be legal and not allowing proxy voter and participating candidates should have confidence of the voting results. The undeniable property of the proposed scheme can satisfy these requirements.

References

1. Tepandi, et al.: Wireless PKI Security and Mobile Voting. Computer. IEEE Computer 43(6), 54–60 (2010)
2. Want: iPhone: Smarter Than the Average Phone. IEEE Pervasive Computing 9(3), 6–9 (2010)
3. Shamir, A.: Identity-Based Cryptosystems and Signature Schemes. In: Williams, H.C. (ed.) CRYPTO 1985. LNCS, vol. 218, pp. 47–53. Springer, Heidelberg (1986)
4. Boneh, D., Franklin, M.: Identity-based encryption from the Weil pairing. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 213–229. Springer, Heidelberg (2001)
5. http://en.wikipedia.org/wiki/ID-based_cryptography

6. Elgamal, T.: A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. IEEE Transactions on Information Theory IT-31(4), 469–472 (1985)
7. Chaum, D.: Undeniable Signatures. In: Brassard, G. (ed.) CRYPTO 1989. LNCS, vol. 435, pp. 212–216. Springer, Heidelberg (1990)
8. Ratha, N., Ratha, J.N., Connell, J., Bolle, R.: Enhancing security and privacy in biometric-based authentication systems. IBM Systems Journal 40(3), 614–634 (2001)

Automatic Metadata Conversion System with Ontology

Min-Ho Lee[1], Hwa-Mook Yoon[1], Won-Kyung Sung[1], and Jae-Cheol Ryou[2]

[1] KISTI, 245, Daehakro, Yuseong-gu,
305806 Daejeon, Korea
[2] Chungnam National University, 99, Daehakro, Yuseong-gu,
305764 Daejeon, Korea
{cokeman,hmyoon,wksung}@kisti.re.kr, jcryou@home.cnu.ac.kr

Abstract. Some information providers or information circulation centers are collecting various metadata to serve users. The collected information has various structures and expression types of contents as well as file types. For information service, these various types of metadata should be processed and converted into one unified form. Converting metadata has lots of difficulties and mapping between elements of metadata schema is most important and difficult. We are making an automatic metadata conversion system with ontology. This paper describes the conversion method with ontology and schema mapping tool in the system.

Keywords: Ontology, Bibliography Metadata, Conversion.

1 Introduction

Information circulation centers or Information providers are harvesting and serving bibliography metadata from various information sources to users. In the case of our organization, The metadata that we harvest have various file types, elements of metadata and issue number expression as well. Some of the elements follows standard such as DC (Dublin Core) or MARC and the others are independently defined for the need of the information sources.

We should convert expression types and elements of input data into one unified form for service. This paper describes the easy method of making convert rules with ontology mapping tool.

2 Making Bibliography Schema of Our Own

We constructed two bibliography schema of our own and called them KBS (Kisti Bibliography Schema). We reused the conventional bibliographic schema and modified it a little bit for our needs.

For journal metadata schema standard, MODS which is based on XML so that it is excellent for readability, extensibility and interoperability but it can be lost information for its non-capacity of some metadata elements[2]. The existing data that constructed by MARC are very large amount and have indispensable elements. We decided to use MARC-XML which has low readability, though.

For article metadata schema standard, we slightly modified the original PMC-XML schema for some omitted elements.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 367–368, 2011.

3 Ontology Mapping and Data Conversion

Figure 1 shows the concept of metadata conversion using ontology mapping suggested by this paper.

The manager analyzes metadata schemas or document forms from each information sources, makes ontologies of them and called as input ontologies. Our own schema is also converted into ontology as same way.

The manager compares the input ontology to our own ontology and conducts mapping same meaning class or hierarchical class with a mapping tool. The manager makes data value transforming rules in ontology mapping step as well.

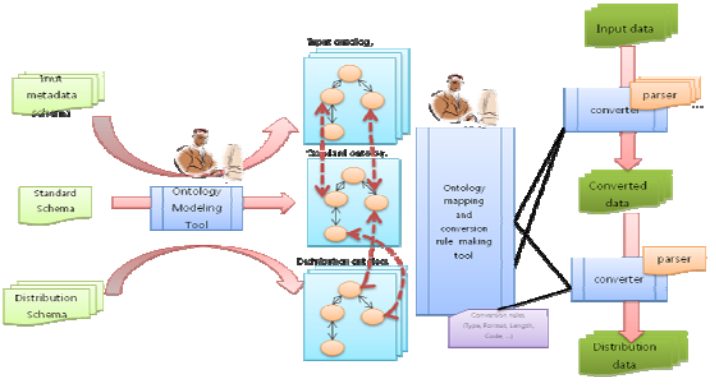


Fig. 1. Metadata conversion based on ontology mapping

Then the system conducts ontology reasoning and transferring original metadata into suitable elements of our standard metadata and converts expression and form of the values.

4 Conclusion

In this paper, we explained difficulties in converting bibliography metadata and described metadata conversion method with ontology. It contributes that users can metadata conversion without any system modification when new bibliography metadata schema input or conventional schema are amended. The system has limitation for using general-purpose metadata conversion except bibliography metadata, because data value forms are too various to transform with several rules.

We are trying to adapt MODS to our schema instead of MARC-XML, because it is more excellent for readability, extensibility and interoperability than MARC-XML.

References

1. Seo, B.D.: A Study on the Metadata standard of Digital Library applied MODS/METS. Thesis, University of Yonsei, Korea (2008)
2. Kim, E.K., Nam, Y.J.: The Comparative Study on the Methodologies of Building Ontology toward Semantic Web. Journal of Information Management 35(2), 57–85 (2004)

A Research Model for Evaluating the Success of the DCMS: A Digital Content Management System

Sung-Ho Shin, Wha-Mook Yoon, and Won-Kyung Sung

335 Gwahangno, Yuseong, Daejeon, Korea
{maximus74,hmyoon,wksung}@kisti.re.kr

Abstract. The digital content management system, called DCMS, is similar to early information system ideas. For evaluating success of the DCMS, some success factors except success factors by Delone and Mclean are considerable because it is a system for digital content management and with web-based technologies. Based on the information system success model by Delone and Mclean, we added two success factors presented by research of Quaddus and Lau; user participation, user characteristics. Also three factors are taken directly from Delone and Mclean, but modified in appropriate ways for the DCMS. The main contribution to the literature from this study is to theoretically articulate conceptual model for evaluating the success of the DCMS.

Keywords: digital content management system, DCMS, information systems success model, web(-based) information system.

1 Introduction

CMS(Content Management System) is also a class of information system[1]. CMS is an information management system that preserves, organizes, disseminates and locally developed documents and external documents with associated metadata.

Information system evaluation is recognized as one of the problematic issues that can be interpreted in many different ways. For evaluating success of CMS, some success factors except success factors by Delone and Mclean are considerable because it is a system for content management and with web-based technologies. In this study, we propose a research model for evaluating success of the DCMS, a digital content management system.

2 Literature Review and Research Motivation

Delone and Mclean's IS success model has received much attention among IS researchers, as it provides a foundation for research in the CMS domain[2]. This is model identifies six interrelated dimensions of IS success; system quality, the output of information quality, consumption (use) of the output, the user's response(user satisfaction), the effect of the IS on the behavior of the user (individual impact), and the effect of the IS on organizational performance (organizational impact). Adeyinka and Mutula[2] adapted and extended further the updated Delone and Mclean IS

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 369–370, 2011.

success model. Based on the foregoing and the updated Delone and Mclean's information system success model, their model is developed as shown in Fig. 1.

Quaddus and Lau have pointed that user participation and user satisfaction have the potential to influence organizational effectiveness and that system complexity can moderate the relationship between user participation and user satisfaction[3]. In addition, their model also shows four independent characteristics of the user as factors contributing to user satisfaction.(Fig. 2)

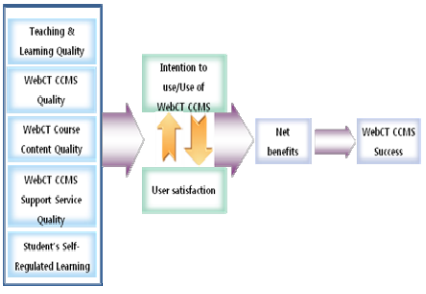


Fig. 1. Adeyinka and Mutula's model

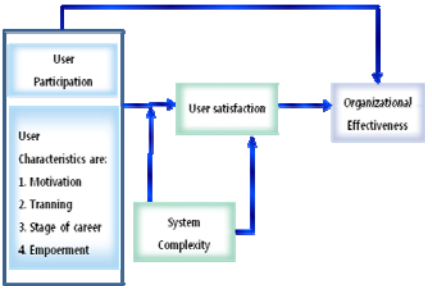


Fig. 2. Quaddus and Lau's model

3 Research Model for the DCMS Evaluation

Based on the Information system success model by Delone and Mclean, we added two success factors presented by Quaddus and Lau; user participation, user characteristics. Also three factors are taken directly from Delone and Mclean, but modified in appropriate ways for the DCMS. Therefore, a research model for this study is developed as shown in Fig. 3.

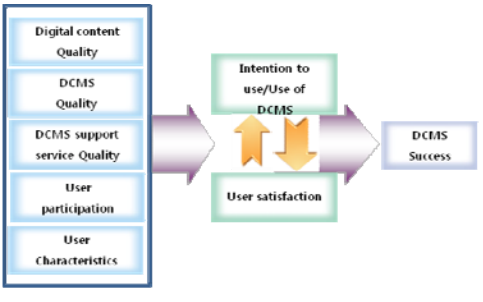


Fig. 3. Research model for the DCMS success

References

1. Han, Yan: Digital content management: the search for a content management system, vol. 22(4), pp. 355–365 (2004), http://LibraryHi.Tech
2. Adeyinka, T., Mutula, S.: A proposed model for evaluating the success of WebCT course content management system. Computers in Human Behavior 26, 1795–1805 (2010)
3. Quaddus, M., Lau, A.: Impact of user participation on Web-based information system: The Hong Kong experience. In: 10th International Conference on Computer and Information Technology, Dhaka, pp. 1–6 (2007)

User's Location Prediction System Using the Filtering with Correlation Coefficients Weight

Kyoung-Mi Im[1], Chan-Yung Choi[2], and Jae-Hyun Lim[2]

[1] Green Home Energy Technology Research Center, Kongju National University
[2] Computer Science Department, Kongju National University
Budeadong 275, Cheonan, Chungnam, Korea
{omnibus,chlcks0,defacto}@kongju.ac.kr

Abstract. This paper proposes User's Location Prediction System using the Filtering with Correlation Coefficients Weight. This system heterogeneous data occurred during the process of collecting context information into homogeneous data, and improves the accuracy of clustering needed for location-awareness. Also, it applies correlation coefficients weight to extract features and predicts user's location through ARIMA time-series analysis. For the evaluation of the proposed method a test bed is constructed, tested and compared with the method without weighting.

Keywords: location prediction, location awareness, context awareness, ARIMA, time series analysis, correlation coefficients.

1 Introduction

Recently, Public Safety Services, Location Tracking Services, Information Services and more that recognize minimum user input information and real time user context information to provide personalized services are becoming a reality, and Energy Saving Services using location information is being actively researched[1][10].

For context-aware service's location awareness, collection of context information and analysis of collected information must be done. There are many cases where data collected during the process of collecting context information is not homogeneous, but heterogeneous data. The closest current examples are the cases of AP(Access Point)'s MAC address or wireless LAN's SSID(Service Set Identifier). If data collected in these ways are of heterogeneous values, they will be expressed in different data formats, decreasing the accuracy of clustering needed for location detection. Therefore the data needs to be converted into homogeneous data through feature extraction.

The numerous data that is used during context information analysis does not necessarily have the same level of contribution. Context awareness can be said to be determined by data with high level of contribution and cases where data does not have a high level of contribution, where weighting following the level of contribution was not applied, where incorrect data is collected, etc. effect the result hugely. Therefore weighting following the level of contribution must be applied, with the context relationship between the data in mind.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 371–378, 2011.

However, in existing services, when user's location prediction or context prediction is done using user's context information, characteristics are extracted from the data collected without considering the correlation coefficient[3],[11],[12]. Moreover, in the current state of things, feature-extraction is being carried out through calculation using solely identical data from two context information[6].

In this paper, data in heterogeneous form is converted into homogeneous data, and context information's feature extraction method using filtering based on correlation coefficients weight is proposed in order to apply weighting within context information. The system of this paper is separated into three stages, Feature Extractor, Context Classifier and Predictor. Feature extraction is done using filtering with correlation coefficients weight and user's location is recognized through feature classification. Also the user's location is predicted through ARIMA (Auto-Regressive Integrated Moving Average) time-series analysis based on location information. The efficiency of the method proposed in this paper is shown through its comparison with a method without correlation coefficients weight, which shows the accuracy of the proposed method.

2 Location-Prediction System

The structure of the location prediction system proposed in this paper consists of three stages, like in Figure 1. In the first feature extraction stage, user's context information is collected and feature is extracted through filtering based on correlation coefficients weight. This is stored in the database. In the second Context Classifier Stage, feature data stored in history database is studied through context information classification algorithm. In the final third Predictor stage cluster member variables are taken from the time-series database and the user's location is predicted through time-series analysis's prediction process.

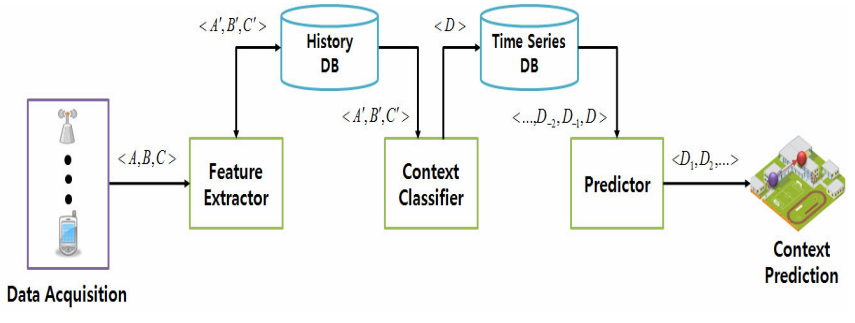


Fig. 1. Structure of location prediction system

2.1 Feature Extractor

Feature Extractor is a stage that processes user's context information collected, extracting data in heterogeneous form as homogeneous form using filtering with

correlation coefficients weight. Data used in general feature extraction can be classified into four forms(nominal, ordinal, numerical, interval) and measurement of similarity is applied according to the form of data[5][7].

The result of context awareness is different depending on the level of contribution of data within context information data and so data with high level of contribution must have a greater proportion, but calculation used for general feature extraction cannot express the importance of the data because it ignores the level of contribution of data. Therefore, this paper proposes a method that applies great proportion in case of direct proportion or inverse proportion and small proportion if there is no relationship, by calculating the change within the context information data.

Feature Extraction's $f(x, y)$ is equal to equation 1 and weighting has been applied to it by getting the correlation coefficient between the two variables.

$$f(x, y) = x' \times W_{x'y'} \tag{1}$$

Here x' is the value of changing variable x to $0 \le x \le 1$ and the changed value is placed in order in same data to get the minimum value and maximum value, like in equation 2.

$$X = \{\min, ..., x_n, ..., \max\} \tag{2}$$

Also to convert x into a value between 0 and 1, equation 3 is applied which is a standardization of data method through controlling the boundaries[2]. In this case, to set the boundary for value of $x_n^{'}$ to $0 \le x_n^{'} \le 1$, 1 is applied to $w\max$ 1 and 0 is applied to $w\min$.

$$x_n' = \frac{x_n - Min(X)}{Max(X) - Min(X)}(w\max - w\min) \tag{3}$$

$W_{x'y'}$ of Equation 1 uses the Pearson correlation coefficient and expresses the relationship between the two variables and depending on their degree of relation, the value of $-1 \le W_{x'y'} \le 1$ is found. If the correlation $W_{x'y'}$ is close to the straight line with a positive gradient on the coordinates it will have positive(+) correlation, where y increases as x increases. If it is close to a straight line with a negative gradient it will have a negative (-) correlation. Also as the variables become less related to each other, the correlation coefficient will be closer to 0[4]. In this paper, correlation coefficient is used to only express the correlation between two variables so correlation coefficient weighting is calculated using the absolute value gained from the Pearson correlation coefficient, like in equation 4. Here W_{xy} is correlation coefficient weighting and x, y are two data within context information. Feature values extracted through filtering are stored in the database after calculation.

$$W_{xy} = \left| \frac{\sum xy - \frac{\sum x \sum y}{N}}{\left(\sum x^2 - \frac{(\sum x)^2}{N} \right) \left(\sum y^2 - \frac{(\sum y)^2}{N} \right)} \right| \tag{4}$$

2.2 Context Classifier

The separation algorithm used in this paper uses SOFM (Self Organizing Feature Map) which was proposed by Kohonen[8]. Kohonen's SOFM is a nerve network model for competitive learning, an unsupervised algorithm that separates high dimensional n number of data into c number of clusters according to similar features[13].

2.3 Predictor

Predictor predicts the value of group member by extracting time-series data stored in time-series database as group member value, and through ARIMA time-series analysis. The user's location is predicted using the predicted group member value attained in this way. This paper uses ARIMA (Auto-Regressive Integrated Moving Average) model's difference method for time-series analysis and predicts user's next location[9].

3 Experiments and Evaluation

In this chapter, feature extraction method proposed in chapter 2 is used to undergo location prediction and its accuracy is proved. Also it will be compared with feature extraction method without weighting to show the proposed structure's superiority. For the experiment, the user's location data was collected in 30 second intervals for three weeks and the number of AP within the boundary and AP's MAC address and signal strength by using the user's Smart Phone in areas where Wi-Fi is available. The boundary of the moving area was set as the laboratory, the office room, and the lecture room of K University's engineering department building's 7th, 8th, and 9th floor.

Total number of data collected was 16,000 and out of this, 11,000 were used as data for study and the remaining 5,000 was used as experiment data.

3.1 Feature Extractor

Data collected from AP and time-series data read from the history database are converted in the feature extractor through standardization of data and application of correlation coefficients weighting, and are stored in the history database.

3.1.1 Collection

Context information is collected as time, MAC address of closet AP, signal strength of closest AP, and number of AP within the boundary, in approximately 30 second intervals.

3.1.2 Standardization of Data

The context information collected is represented as the result in Table 1 by using equation 3, and its values are the value of data before weighting was applied. In the case of MAC address, the collected data is in nominal form and is converted into integer form which is able to be calculated before standardization of data. The boundary was set as $0 \le x_n^{'} \le 1$.

Table 1. Feature extraction of data before applying weighting

Time	MAC address	Signal strength	Number of AP
...	...	...	...
09:35:05	0.071	0.706	0.222
09:35:35	0.214	0.686	0.333
09:36:05	0.796	0.510	0.555
09:36:35	0.296	0.588	1.000
...	...	...	...

3.1.3 Correlation Coefficients Weight and Feature Data Save

The correlation is calculated using feature extracted data with value between 0~1 with equation 4.

Table 2. Feature extraction with weighting applied

Time	MAC address	Signal strength	Number of AP
...	...	...	...
09:35:05	0.071	0.019	0.035
09:35:35	0.214	0.019	0.053
09:36:05	0.796	0.014	0.088
09:36:35	0.296	0.016	0.158
...	...	...	...

The calculated correlation coefficients weighting is the correlation within context information. In the case of MAC address, the weighting is 1, 0.158 and 0.027 each depending on the number of AP within the boundary and the signal strength of closest AP. If this is applied to equation 1, feature extraction as shown in Table 2 is done. The signal strength of closest AP and number of AP within the boundary is feature extracted in the same way. The feature extracted data is stored in the history database.

3.2 Context Classifier

Feature extracted data is classified in the context information classifier using the SOFM algorithm and the member value of the classified group is stored in the time-series database.

3.2.1 Clustering

Feature extracted data is classified into 750~3830, 11230~17720, 18860~22000 for the lecture room, the office room and the laboratory each, using SOFM and the remaining section was processed as noise.

3.3 Predictor

The predictor brings out the data stored in time-series database and goes through time-series analysis stage and predicts the user's location.

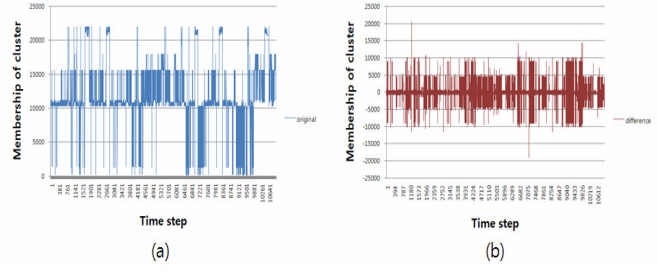


Fig. 2. Time-series chart. (a) original data (b) after 1st difference

3.3.1 Time Series Data Load

The data loaded from time-series database is analysed using a statistics program SPSS 14.0 and as the first stage of analysis, the time-series data is drawn as a time-series graph like in Figure 2 (a).

3.3.2 Time Series Analysis

In Figure 2 (a), it can be seen that it is a non-stationary time-series where the average is not constant. Therefore it needs to be made stationary through difference. Figure 2 (b) shows a converted form using stage 1 difference, into a stationary time-series. Then, the time-series can be determined with ARIMA(0, 1, 5).

3.3.3 Prediction

The result of time-series ARIMA(0, 1, 5) is graphed as seen in Figure 3. The graph's x-axis show the flow of time in 30 second intervals and the y-axis shows cluster sectional value. If data with correlation coefficient's weighting applied and data without correlation coefficient's weighting applied is compared with the original data graphically, the accuracy of the weighting-applied data is similar to the raw data and the proposed method's accuracy is 94.38%. According to the classified 3 sections, accurate and inaccurate is separated and so accuracy is determined.

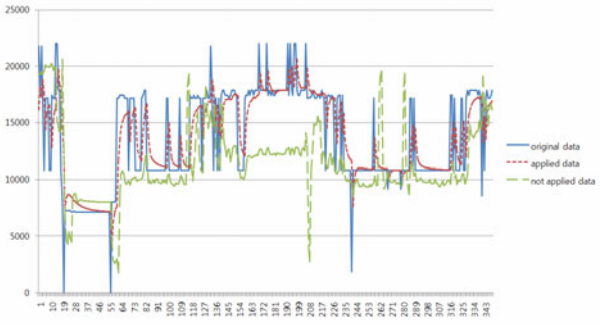


Fig. 3. Comparison of prediction value of feature extraction

4 Conclusions and Further Studies

This paper proposed the User's Location Prediction System using the Filtering with Correlation Coefficients Weight method which has a better accuracy than the general similarity determination method which is the standard method for user's location prediction in a Smart environment. Unlike the general model that does not consider the correlation between the data in context information, the proposed User's Location Prediction System using the Filtering with Correlation Coefficients Weight method converts each situation's heterogeneous data into homogeneous data. Also filtering depending on the level of contribution within the context information data with the correlation coefficients weighting in consideration is proposed. Location was predicted using ARIMA time-series analysis and to evaluate the proposed method's performance, user's context information was collected and the accuracy between filtering method based on correlation coefficients weighting and filtering method not based on correlation coefficients weighting was compared. The accuracy of location-prediction with the application of the proposed method is 94.38%, which is a 88.41% improvement from the method on which the proposed method was not applied.

In this paper the performance of the prediction method is evaluated based on data needed for location prediction out of numerous context information data, and research for the improvement of accuracy where the field is widened to not just location prediction, but to context prediction, should be attempted in the future.

Acknowledgments

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0028289).

"This work was supported by the Grant of the Korean Ministry of Education, Science and Technology" (The Regional Core Research Program / Zero Energy Green Village Technology Center).

References

1. Dev, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing Journal 5(1), 4–7 (2001)
2. You, C.W., Chen, Y.C., Chu, H.H., Huang, P., Chiang, J.R., Lau, S.Y.: Sensor-enhanced mobility prediction for energy efficient localization. In: Proc. Sensor and Ad Hoc Communications and Networks (SECON 2006), vol. 2, pp. 565–574 (2006)
3. Mayrhofer, R.: An Architecture for Context Prediction. PhD thesis, Johannes Kepeler University of Linz, Altenbergstrasse 69, 4040 Linz, Austria (2004)
4. Sigg, S.: Development of a novel context prediction algorithm and analysis. Kassel University Press, GmbH (2008)
5. Choi, C.-Y., Lim, J.-H.: Location Prediction of Time Series Analysis using Correlation Coefficient Weight. Korean Society for Internet Information 12(2), 73–76 (2009)
6. Negri, S., Belanche, L.A.: Heterogeneous kohonen networks. In: Mira, J., Prieto, A.G. (eds.) IWANN 2001. LNCS, vol. 2084, pp. 243–252. Springer, Heidelberg (2001)
7. Mayrhofer, R., Radi, H., Ferscha, A.: Feature extraction in wireless personal and local area networks. In: Agha, K.A., Omidyar, C.G. (eds.) TheProceedings of The Fifth IFIP-TC6 International Conference on Mobile and Wireless Communications Networks (MWCN 2003), pp. 195–198 (2003)
8. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publisher, San Francisco (2001)
9. Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to data mining. Pearson Addison-Wesley, London (2006)
10. Everitt, B., Landau, S., Leese, M.: Cluster Analysis. Edward Arnold, London (2001)
11. Kohonen, T.: Self-Organizing Maps. Springer, Berlin (1997)
12. Lee, J.-S., Kang, M.-K.: A Clustering Algorithm Using the Ordered Weight of Self-Organizing Feature Maps. In: Korean Operations Research and Management Science Society, pp. 41–51 (2006)
13. ki, Y.-H.: Time Series Prediction. In: Hyung Seul Publish (2001)

An Efficient String Searching Algorithm Based on Vowel Occurrence Pattern

Kwang Sik Chung[1,*], Heon-Chang Yu[2], and Sung Ho Jin[3]

[1] Dept. of Computer Science, Korea National Open University
kchung0825@knou.ac.kr
[2] Dept. of Computer Science Education, Korea University
yuhc@korea.ac.kr
[3] Dept. of Computer Science Education, Korea University
wintop@korea.ac.kr

Abstract. String searching algorithm is used to determine whether a word, a statement and a phrase exists in newspaper article, books, web pages or not. Especially as the volume of web pages and documentation increases, time and cost of searching a pattern increases. In this paper, we consider the characteristics of English alphabets combination in a pattern and a text(frequency and position of vowels in a pattern and a text), and propose new efficient string searching algorithm based on vowel composition structure of pattern and text. If a pattern has two or more vowels, proposed vowel-based string searching algorithm is more efficient at the both case of best matching case and worst matching case. Especially, a pattern's length does not affect performance of proposed vowel-based string searching algorithm. Only a pattern's location in a text can affect running time of algorithm. Proposed vowel-based string searching algorithm can avoid un-matching vowels in a text and un-necessary comparison between a pattern and a text. Therefore in reality, vowel-based string searching algorithm is usefully and effectively applicable to string searching function of a pattern and a text.

Keywords: string searching algorithm, vowel-based string search, pattern, text.

1 Introduction

As volume of text information to be processed by computer increases, importance of a text information processing time has also increased. At first, misspelling finding was the main goal of string searching algorithm. String searching algorithm was developed for those goals. In general word processor and editor, a string searching and misspelling finding became basic function. But now, searching string or some information in the web and digitized text becomes the most important goal of string searching algorithms. String searching decides whether given pattern exists in given file or text or not, and shows the position of given pattern in given file or text. Brute-Force algorithm[4,5] is naïve. KMP(Knuth-Morris-Pratt) algorithm [2,3,4,5,6] is

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 379–386, 2011.

based on finite automata, and Boyer-Moore algorithm [l,4.5] and Rabin-Karp algorithm [4,5] are very effective.

Since previous methods did not consider characteristics of English word composition, previous algorithms made unnecessary overload and non-effective performance if they were applied to real text files. An English word is made by alphabets' combination of vowels and consonants. We consider frequency of vowels and consonants, and same vowel's occurrence repetition at two or more position in a pattern in order to search a pattern in text file. And in order to use these characteristics of a pattern, we analyze a pattern structure of vowels and consonants, extract vowels and repetition position of vowels and after that, search and compare the pattern in text file. In other words, we decide a pivot vowel that occurs most frequently in a given pattern, and the pivot vowel is firstly compared with an alphabet of text after mismatch happens. We compare the pivot vowel firstly and then compare the other vowels.

In this paper, we propose vowel-based string searching algorithm and evaluate the proposed algorithm. The proposed vowel-based string searching algorithm is compared with other string searching algorithm in the view of time complexity. Lastly we show the other study plan.

2 Pattern Search Methods

2.1 Brute-Force Algorithm

Brute-Force algorithm is the simplest method among string search algorithms. Brute-Force algorithm compares first one alphabet of pattern with one alphabet of text that is a second alphabet on previous failed search. If one alphabet of pattern matches with one alphabet of text, then the next alphabet of pattern is compared with the next alphabet of text. And these comparisons continue till all alphabets of pattern orderly match with consecutive alphabets in a text and a pattern is consumed. If a mismatch occurs between one alphabet of a pattern and one alphabet of a text, then first alphabet of a pattern is compared with a next alphabet of a mismatched alphabet in text. If a pattern and a text consist of binary codes, the search performance worsens. If a pattern is the first word of a text file, then number of comparison between pattern and text is m(length of pattern). It is the best case of Brute-Force algorithm. But at the worst case, number of comparison between a pattern and a text is m•n(n is length of text) and time complexity is O(n).

Performance of Brute-Force method is subject to length of pattern or text file since mismatching makes repeated comparisons of same alphabets in text and pattern. But Brute-Force method is simple and legible.

2.2 KMP (Knuth-Morris-Pratt) Method

Knuth-Morris-Pratt string searching algorithm[2,6] can find out 'incorrect start position' in a text when a pattern does not match with a text. Therefore, after mismatching, Knuth-Morris-Pratt string searching algorithm can decide more advanced comparison restart position in a pattern. And it is very effective to search a binary pattern in a binary string. Since Knuth-Morris-Pratt string searching algorithm needs repetition of some alphabets for deciding an advanced comparison restart position. But since there are only few patterns that have repetitions of several

alphabets, Knuth-Morris-Pratt string searching algorithm can rarely be useful to decide advanced comparison restart position in real text.

2.3 Boyer-Moore Method

Boyer-Moore string searching algorithm[4] basically has efficiency derived from the fact that it avoids each unsuccessful comparison to find a match between a pattern and a text. Boyer-Moore string searching algorithm uses the information gained from that comparison in order to rule out as many positions of a text as possible where a pattern cannot match. To make it more efficient, Boyer-Moore string searching algorithm makes two kinds of tables, one table calculates how many positions ahead to start the next search based on the identity of the character that caused the match attempt to fail and the other makes a similar calculation based on how many characters were matched successfully before the match attempt failed. But Boyer-Moore string searching algorithm needs consecutive repetition of a alphabet unit with vowels and consonants. But those kinds of repetition of a alphabet unit are few cases in a real pattern and a real text.

3 Vowel-Based String Searching Method

Vowel-based string searching algorithm has pattern preprocessing steps that consist of a vowels searching step and a vowels analyzing step. At a vowels searching step, the most frequent vowel and other vowels are extracted from a pattern. And we examine the number of vowel occurrence in pattern. The most frequent vowel becomes pivot vowel for vowel-based string searching algorithm, and searching strategy is set based on the pivot vowel. There are six vowels in English alphabet, and the most frequently used vowels are less than six. If vowels' positions and alphabets in a pattern and a text match together, it means the high probability of remains searching success. Especially a pivot vowel(most frequent vowel in a pattern) that is the most significant vowel in a pattern is initially compared with a text vowel at every searching step. Pivot vowel based searching strategy can decrease the number and probability of consonants mismatch and other vowels in a pattern and a text, and increase performance of string searching algorithm.

For example of pattern preprocessing, let a pattern example be 'control'. In this pattern, vowel 'o' occurs two times in the pattern and there is no vowel except for 'o'. Then we initially match the first 'o' of the pattern with vowels in a text. If the first 'o' of the pattern matches with a vowel in the text, then the second 'o' of the pattern is compared with forward-fourth alphabet from matched 'o' in the text. The second example is 'documentation'. In that case, there are five vowels('o', 'u', 'e', 'i', 'a') in the pattern and 'o' is a pivot vowel and the most significant vowel in a pattern. We initially compare 'o' of the pattern with vowels in a text. If the first 'o' of the pattern matches with a vowel in a text, then the next 'o' of the pattern is compared with forward fourth alphabet from matched 'o' in a text. After that, if two 'o's comparisons success, the next vowel 'u' is compared with forward first alphabet from the first matched 'o' in text with pattern. As like that, vowel comparison processing does not stop until there is no un-compared vowel in a pattern. After all vowels' comparisons of the pattern success, then consonants of the pattern are compared with consonants within the range of vowel matched text.

3.1 Vowel-Based String Searching Algorithm

In vowel-based string searching algorithm, the matching sequence can be changed according to the occurrence frequency of vowels in a pattern. After success comparison of a pivot vowel, the next significant vowel in a pattern is compared with a vowel of a text. After successful comparison of all vowels in a pattern, consonants of the pattern can be compared with consonants in a text. It means that the number of consonants in a text to be compared with a pattern decreases since only consonants to be compared are within the range of vowel matched text.

Vowel-based string searching algorithm runs a pattern preprocessing step. A Pattern preprocessing step returns *vowels and their positions in a pattern*, 2-dimensional array of vowels of a pattern. Vowels-positions of a pattern consist of index as vowels of a pattern and values as their position in a pattern. And sequence of vowel in Vowel-position of a pattern is determined by vowel frequency in a pattern. Pattern_Preprocessing() is as follows.

```
Pattern_Preprocessing()
{
 i = 0;
 j = 0;
while(end of pattern)
{
if(pattern[i] == vowel)
    {
     pattern_matrix[j][0] = pattern[i];
     pattern_matrix[j][1]=i;
        j = 0;
    }
    i = i+1;
}
Calculate_Frequency_of_Vowel()
}
```

Calculate_Frequency_of_Vowel() determines vowel matching order strategy of a pattern and save it into vowels and their positions in a pattern(pattern_matrix).

```
Calculate_Frequency_of_Vowel()
// calculate frequency of vowel and decide pivot vowel of pattern
{
 while(end of pattern_matrix)
  {
    j=0;
    for(i=j;i++;i < number of pattern_matrix)
    {
      if(pattern_matrix[j][0] == pattern_matrix[i][0])
    pattern_matrix[j][2] = pattern_matrix[i][1]-
                                  pattern_matrix[j][1];
     }
  for(end of pattern_matrix)
    temp[][1] = temp[][1]+pattern[];

  for(number of temp)
   {
```

```
   if(frequency_of_vowel < temp[][1])
     pivot_vowel = temp[][0];
//comparison unit length of the most frequent vowel
 }

  for(number of temp)
   {
   if(frequency_of_vowel == pattern_matrix[i][0])
     {
      length_of_pivot_vowel[j] = [i][1];
      j++;
      }

  jump_length_of_pivot_vowel
 =length_of_pivot_vowel[j-1]-length_of_pivot_vowel[0];
  // comparison unit as length between vowel and vowel
   }
 }
```

find_most_frequent_vowel_in_pattern() finds out and returns *the most significant vowel in a pattern*.

```
Find_Pattern_in_Text( )
{
while(end of text && FOUND != T)
{
if(pivot_vowel == text[i])
 {
  // when matching vowel of text with the last position's
  //  vowel of the most frequent vowel
if(pivot_vowel == text[i-jump_length_of_
                            pivot_vowel])
   {
      // comparison of pivot vowel
     while(pivot_vowel  == pattern_matrix[j][0]
         &&  MATCH_pivot_vowel == T)
      {
     if(pivot_vowel!=text[i-jump_length_of_pivot_vowel
                                  +pattern_matrix[j][2]])
         MATCH_vowel == F;
       j++;
      }

    // check the position of other vowel except for pivot vowel
     while(MATCH_pivot_vowel == T)
      {
        if(pattern_matrix !=
         text[i-
ump_length_of_pivot_vowel+pattern_matrix[j][1]
                                        -
pattern_matrix[0][1])
       MATCH_vowel == F;
       j++;
```

```
      }

// after comparison of all vowels in pattern successes, //matching
consonant
     j = 0;
     while(end of pattern && FOUND == T)
      {
        if(pattern[j]              !=              text[i-
jump_length_of_pivot_vowel
                                    +pattern_matrix[j][1])
          FOUND == F;
        j++;
      }
   }

// when matching vowel of text with the last position's
//  vowel of the most frequent vowel

 if(pivot_vowel==text[i+jump_length_of_pivot_vowel])
   {
   // comparison of pivot vowel
     while(pivot_vowel  == pattern_matrix[j][0]
                          && MATCH_pivot_vowel == T)
      {
     if(pivot_vowel!=text[i+jump_length_of_pivot_vowel
                                - pattern_matrix[j][2]])
         MATCH_pivot_vowel = F;
       j++;
      }
// check the position of other vowel except for pivot vowel
     while(MATCH_pivot_vowel == T)
      {
        if(pattern_matrix[j][0] !=
           text[i+ump_length_of_pivot_vowel
             - attern_matrix[j][1]+pattern_matrix[0][1])
            MATCH_pivot_vowel = F;
        j++;
      }

// after comparison of all vowels in pattern
// successes, matching consonant
     j = 0;
    while(MATCH_pivot_vowel&&end_of_pattern&&FOUND==T)
      {
       if(pattern[j]!=text[i+jump_length_of_pivot_vowel
                                  -pattern_matrix[j][1])
          FOUND == F;
        j++;
      }
   }
i++;
}
}
```

4 Time Complexity Evaluation of Vowel-Based String Searching Algorithm

In this chapter, we analyze time complexity of vowel-based string searching algorithm and compare it with traditional other string searching algorithms that are Brute Force algorithm, Boyer-Moore algorithm, and KMP algorithm. Brute-Force algorithm has O(n·m) time complexity. Boyer-Moore algorithm has O(n) time complexity. KMP algorithm has O(m+n) time complexity. Vowel-based string searching algorithm has O(n·m) time complexity.

But since text, newspapers, and web pages do not consist of binary codes and consecutive occurrence repetition of same alphabets rarely happens in real world, KMP algorithm and Boyer-Moore algorithm are not suitable for real world's text. Vowel-based string searching algorithm makes use of frequency of vowels, position of vowels, and the most frequent vowel in a pattern in order to avoid unnecessary comparisons of vowels and consonants. If there are two or more vowels in pattern and one or more vowels repeatedly occurs in a pattern, then unnecessary repeated vowel comparison of same vowel in a text does not happen and always forward comparison of pattern and text is possible. At that point, vowel-based string searching algorithm is executed like Brute-Force algorithm and has O(m) time complexity. Only by vowel position comparison we can increase probability of searching success of a pattern in a text. And from mismatch position analyzing, vowel-based searching algorithm can proceed forward and restart comparison at advanced position of a text. Therefore if a pattern has two or more repeated vowels, then it has O(m) time complexity. Especially in the real world, vowel-based string searching algorithm can filter off un-matched vowels of a text and avoid unnecessary comparison.

5 Conclusion and Future Works

In this paper, we consider the characteristics of English alphabets and text structures, and propose efficient new string searching algorithm based on vowel structure of a pattern and a text. We compare the time complexity of proposed vowel-based string searching algorithm with others. Proposed vowel-based string searching algorithm considers the frequency of vowels in a pattern and compares a pattern with a text mainly by the most frequent vowel in the pattern. For frequency of vowels and selection of a pivot vowel(the most frequent vowel in the pattern), we define pattern_matrix as data structure. Proposed vowel-based string searching can avoid un-matching vowels in a text and un-necessary comparison between a pattern and a text. Therefore in reality, vowel-based string searching algorithm is usefully and effectively applicable to string searching function of a pattern and a text.

In the future, we will apply automata theory to vowel-based string searching algorithm. It will reduce the formalization overhead of pattern preprocessing step. And we consider text preprocessing step that will extract vowels of text and formalize vowels structure of a text.

References

1. Boyer, R.S., Moore, J.S.: A Fast String Searching Algorithm. Communication of the ACM 20 (1977)
2. Knuth, D., Moris, J.H., Pratt, V.: Fast String searching in Strings. SIAM Journal on Computing 6(2) (1977)
3. Baeza-Yates, R., Gonnet, G.H.: A New Approach to Text Searching. Communications of the ACM (1992)
4. Baase, S., Van Gelder, A.: Computer algorithms: introduction to design and analysis. AddisEdisson-Wesley Pub. (1999)
5. Sedgewick, R.: Algorithms in C++. Addison-Wesley Pub., Reading (1998)
6. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn. Section 32.4: The Knuth-Morris-Pratt algorithm. MIT Press and McGraw-Hill (2001)

A Flexible Cost-Based Privacy-Aware Data Integration System in Cloud

Yuan Tian, Biao Song, and Eui-Nam Huh

Department of Computer Engineering,
Internet Computing and Network Security Lab,
KyungHee University Global Campus, South Korea
{ytian,bsong,johnhuh}@khu.ac.kr

Abstract. When all the services provided in cloud computing are on-demand and highly commoditized, consumer would like to pay different prices for those privacy services with different protection assurances according to the importance of their data. We recently present a privacy-aware inter-cloud data integration system considering tradeoff between the privacy requirements from users and the charging for those data protection and processing. In this paper, we extend our previous work by providing four query execution schemes which assures a more sophisticated and flexible privacy protection. Another contribution of our paper is that we proposed an approach for allocating those schemes. As the schemes applied in user's query are relevant and will affect the scheme selection for the next data query, our method assures that those schemes can be implemented properly.

Keywords: Cloud computing, Privacy, Noise injection, Service pricing.

1 Introduction

Cloud computing, in which services are carried out on behalf of customers on hardware and software, becomes increasingly pervasive these days. More and more people are starting to take the advantage of the power of cloud. The provided services range from processing vast datasets for complicated scientific problems to using clouds to manage and provide access to medical records [1]. Similar to the real world utilities, nearly all the services provided in cloud computing are on-demand and highly commoditized, based on consumers' usage and quality of service expectations, they just need to pay for those services like water, electricity, gas and telephony [6], rather than investing heavily to maintain their own computing infrastructures.

However, whenever information is shared in the cloud, privacy and security questions may arise. When a user stores his data in a third party like cloud computing provider, those data may have fewer or weaker privacy protections than when the data just remains in the possession of this user. As a vital building block in many fields, privacy is possible to be a new service in the cloud environment. Consumers do not have to concern about how to protect their privacy by which technology, instead, according to consumers' requirements, the privacy services can be offered by providers to execute their applications only if consumer pay for that [3]. The pricing

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 387–396, 2011.

mechanism decides how service requests are charged. Pricing serves as a basis fro managing the supply and demand of computing resources within the data center and facilitates in prioritizing resource allocations effectively [11].

Currently, there are few privacy laws apply to restrict the disclosure of customers or employees' personal information from a business to the cloud provider. Even privacy laws apply to particular categories of customer or employee information, disclosure to a cloud provider may not be restricted, as current laws that protect electronic communications may apply differently to different aspects of cloud computing [10]. We can only enjoy the full benefits of Cloud computing if we can address the very real privacy and security concerns that come along with storing sensitive personal information in databases and software scattered around the Internet [10].

In order to meet the growing requirement for the security and privacy management, many related technologies have been proposed. In [4], Yau et al present a method which safeguards the privacy of data in a very secure way. They provided a privacy preserving repository to accept integration requirements from users, help data sharing services share data, collect and integrate the required data from data sharing services, and return the integration results to users. The main contribution of their work is that the processing of data is kept securely in both data sharing services and repository: data is randomized before sending to repository and encryption/decryption are used from information releasing in the repository. Thus, the compromise of the repository can only reveal the results of the specified data integration request. Compare with the method mentioned in [4] which fully supports the privacy protection, Shaozhi Ye et al [5], proposed a user-side privacy protection model for search users, which is able to prevent part or entire privacy breaches according to the number of noise injection. They gave the lower bound for the amount of noise queries required by a perfect privacy protection and provided the optimal protection given the number of noise queries.

Those technologies meet users' requirements from privacy concern, however, at the same time, with the increasing number of processed data, the cost for privacy protection also increases dramatically. Thus, based on the above analysis, in [2], we proposed a privacy-aware inter-cloud data integration system to safeguard users' privacy with a cost concern. In general, it is up to customers to decide a strategy about how to get a service fulfilled on the basis of their personal feeling of the importance of their data, and the cost for processing data. Our previous work has, however, some limitations. The main problem is that we only provided two options for users: fully protect their requested data or, not protect it at all. While in the real scenario, when all the services provided in cloud computing are on-demand and highly commoditized, consumer would like to pay different prices for those privacy services with different protection assurances according to the importance of their data [8, 9]. The degree of protection should vary based on user's privacy preference and the cost of privacy protection.

Thus, in this paper, we proposed a more sophisticated privacy protection approach. We present four schemes from different privacy concerns, which ranges from no protection, controllable protection to fully protection. Such scheme is the first contribution of this paper. However, we meet a challenge when we apply this approach to our proposed data integration system, that is, the schemes applied in user's query are relevant and will affect the scheme selection for the next data query. So we present a method which used to allocate those schemes properly considering users privacy and cost for protecting privacy. Such method is the second contribution of our paper.

The structure of the rest of the paper is as follows. Section 2 introduces a scenario that is used as a running example throughout the paper. A component-based view of our system is presented in Section 3 and the design approach is given in Section 4, and the last section of this paper presents our conclusions.

2 A Motivating Example

This section presents a scenario used throughout the paper and we use this motivating example to show how our system works. The scenario is a revised version of the case study proposed in [4, 7]. In a healthcare system there are multiple collaborated clouds which participate in processing, sharing and integrating data. We assume for the purpose of getting the menu which is proposed to ulcer suffering, the clouds may contains the data from medical research institutes, hospitals and pharmacies.

Table 1. The databases which used in the motivating example

Disease	Patient
Tuberculosis	Bree
Aids	Bob
Ulcer	Ada
Diabetes	Alice

(a)

ID	Patient
10001	Alice
10002	Tom
10003	Susan
10004	Bree

(b)

ID	Drug	Menu
10001	D1	M3
10003	D2	M3
10004	D3	M2
10005	D1	M1

(c)

Disease	Drug
Ulcer	D4
Cancer	D2
Flu	D1
Diabetes	D3

(d)

Disease records T1 (b)Identification Information T2 (c)Hospital T3 (d)Pharmacy T4.

For simplicity, only four databases are considered in Table 1: a disease record database *T1(Disease, Patient)* which storing patients' names and corresponding diseases which they are suffering, an identification information database *T2(ID, Patient)* which stores patients' names and their IDs, a hospital database *T3(ID, Drug, Menu)* storing hospitalization information and a pharmacy database *T4(Disease, Drug)* which stores popular drugs for each disease.

Table 2. The integrated query

Q1 -> Tmp1 Select *T1.Patient* From *T1* where *T1.Disease*= "Ulcer"	Q2 -> Tmp2 Select T2.ID From Tmp1, T2 where Tmp1.Patient=T2.Patient
Q3 -> Tmp3 Select T4.Drug From T4 where T4.Disease= "Ulcer"	Q4 Select T3.Menu From Tmp2, Tmp3 and T3 where T3.ID=Tmp2.ID and T3.Drug=Tmp3.Drug

Now we are going to express our motivating example by the following four SQL queries which shown in Table 2. Query Q1, Q2 and Q3 generates three temporary tables Tmp1, Tmp2 and Tmp3 respectively, and the final results are from the last query Q4.

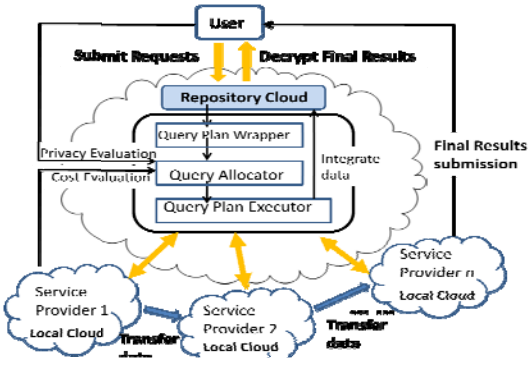


Fig. 1. System Architecture

3 The Proposed System

Our proposed system aims to mediate between users' privacy preferences and the cost for privacy protection. The overall architecture of our system is illustrated in Fig.1. The user sends his integration requirements to the repository cloud and only the required data for users' integration request is collected. The query plan wrapper in the repository cloud will correctly construct a query plan for users' integrated query, decompose the query into a set of sub-queries, and discover corresponding service providers. The data which was conducted or collected in the query plan wrapper are then sent to the query allocator.

Each sub-query could be executed in local cloud or repository cloud. For the former case, the sub-query will be executed in the service provider which stores the corresponding database. Thus, only the results for the sub-query are returned to the repository cloud or transferred to the successive service provider. The price for processing data in local cloud is relative cheaper, however, it may not be secure as both the data storing and processing are done in the same service provider. While the later one is to fetch data in local cloud, randomize all of those data and then send to the query plan executor for further processing. Executing query in query plan executor guarantees secure protection as the data is stored and processed separately. Therefore, the cost for processing data in repository cloud is high as it requires randomizing all the data in the local service provider and transfers it back to the repository cloud.

In order to help users find a balance between privacy insurance and the cost for privacy protection, the query plan executor decides at where the sub-query should be executed. Before sending those sub-queries to the query plan executor, three separating phases are proposed:

- ✓ Privacy evaluation: based on user's query, allows user to express their privacy preferences by setting risk values to decide the importance of data and the relationship between those data.
- ✓ Cost evaluation: here, the pricing mechanism is that the service provider in repository cloud should estimate a price for processing data, whereas other

providers which located outside the repository cloud provides the costs for both processing data and uploading data to the repository cloud.

- ✓ Query Allocation: the query allocator chooses an optimal strategy according to user's preferences in the above two steps, and sends it to the query plan executor to enforce that method. At last, the requested data is sent back to user from the query plan executor.

4 System Design

In this section, we presented the process of our system implementation. Before discussing the functionality of our system, some basic definitions are introduced in order to establish a common ground of concepts.

4.1 Query Plan Graph

Firstly, the query plan wrapper converts user's integrated query to a query plan graph *G*. For a integrated query, the query plan graph $G=\{V, E, C\}$ is a labeled directed acyclic graph. Among which, $V=\{v_1, v_2, \ldots, v_m\}$ is a set of nodes where each v_i represents an intermediate search result; $E=(e_1, e_2, ..., e_l)$ is a set of edges where each edge $e_{ij}=(v_i, v_j)$ represents a data integration relation between v_i an v_j. $C=(c_1, c_2, ..., c_l)$ is a set of labels attached to each $e_{ij} \in E$ and each label $c_{i,j}$=*(op, attr1, attr2)*$\in C$ specifies that the data of v_i and v_j is integrated by the data integration operator *op* between v_i's attribute attr1 and v_j's attribute attr2. Generally, the operator *op* can be any binary comparison operator chosen from {=, ≠, >, <} or any aggregate operator chosen from {*SUM, AVG, MAX, MIN*}.

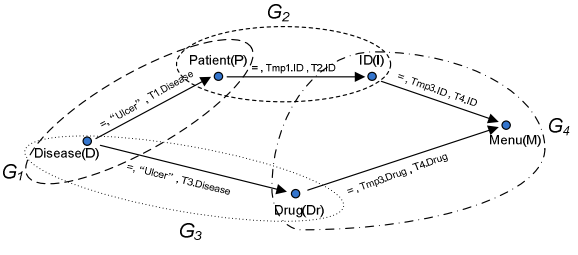


Fig. 2. The query plan graph from the motivating example

The motivating example in Section 2 can be represented by the query plan graph shown in Fig. 2. The query plan graph is decomposed to several sub-graphs $\{G_1, G_2, G_3, ..., G_n\}$ where each sub-graph represents a sub query in Table 2. The overlapping of sub-graphs shows that the intermediate results from one sub-query is used as the input for another sub-query.

4.2 Query Execution Schemes

For each sub-query, we adopt four alternative execution schemes that are presented in Table 3. Scheme 1 applies the privacy protection technology which mentioned in [4]. In this scheme, repository cloud takes the responsibility for executing sub-query. It first collects hashed data table from local cloud, and then executes searching. Thus, the query execution requires hashed input data, and produces hashed output data. Scheme 1 provides 100% privacy protection since the processing of data is kept securely in both data owner (local cloud) and query executor (repository cloud).

Table 3. Query Execution Schemes

Query Execution Scheme	Method	Executor	Input Data Form	Output data Form	Privacy Protection
Scheme 1	Hash	Repository Cloud	Hashed	Hashed	100%
Scheme 2	Hash	Local Cloud	Hashed	Normal	0%
Scheme 3	None	Local Cloud	Normal	Normal	0%
Scheme 4	Noise	Local Cloud	Noise injected	Noise injected	Controllable

Scheme 2 and 3 are similar as none of them adopts privacy protection approach. Repository cloud simply sends the sub-query and input data to local cloud without any pre-protection. Thus, the input data and sub-query processing are entirely disclosed to local cloud. However, the input data forms in scheme 2 and 3 differ, where hashed input data is processed by scheme 2, and normal input data is processed by scheme 3. Since the hash function used in scheme 1 is one way function, it is impossible to convert hashed data to normal data directly. Scheme 2, as an alternative solution, can convert hashed input data to normal output data. In sub-query allocation phase, we utilize scheme 2 to bridge the scheme providing hashed output and the schemes requiring normal input data.

Scheme 4 represents the noise injection technique presented in [5]. In this scheme, the repository cloud first injects extra data to the input data set, and then sends to local cloud. Since local cloud receives noisy data for processing, the privacy breaches can be reduced. After the output data set is returned to the repository cloud, the true result can be obtained by removing the noisy part. This scheme provides controllable privacy protection as the amount of noise can be decided based on user's privacy protection requirement. An essential condition for injecting noise is that the repository cloud has to view the original data rather than hashed one. In other word, this scheme also requires normal data as the input.

4.3 Cost Evaluation

Based on each sub-query, the corresponding local cloud should fix several data-processing prices, which will help the repository cloud to estimate the price for different query execution schemes. Following costs should be provided by local cloud:

1) Data randomization cost

In scheme 1, the sub-query is assigned to be executed in the repository cloud in order to assure the data can be stored and processed separately. Thus, all the data in the

local service provider needs to be hashed before transmitting to the repository cloud. Service provider evaluates this price according to the amount of data. We define *Cr(i)* as the data randomization cost for sub-query *i*. Meanwhile, scheme 2 requires local service provider to pre-process their data before executing the sub-query. Since the input data in sub-query has been hashed, local cloud provider needs to hash the data in corresponding field. We use *Cp(i)* to denote such partial data randomization cost for sub-query *i*.

2) Cost for uploading all data

As scheme 1 requires repository cloud to do searching, local cloud should upload all of the randomized data to the repository cloud. *Cu(i)* denotes the cost for uploading all data of sub-query *i*.

3) Query execution cost

This cost is for executing sub-query either in local cloud or in repository cloud. For each sub-query *i*, the corresponding local cloud provides a cost function $f_i(a)$, where variable *a* denotes the amount of input data in sub-query.

Using above estimations, repository cloud is able to calculate the cost for each scheme. As presented below, the cost for executing sub-query *i* can be calculated by formula (1)-(4):

$$C_1(i) = Cr(i) + Cu(i) + f_i(a) \tag{1}$$

$$C_2(i) = Cp(i) + f_i(a) \tag{2}$$

$$C_3(i) = f_i(a) \tag{3}$$

$$C_4(i) = f_i(a') \tag{4}$$

where $C_1(i)$, $C_2(i)$, $C_3(i)$ and $C_4(i)$ represent the costs of applying scheme 1, 2, 3 and 4 respectively. In formula (4), we use *a'* to denote the input which contains the number of input data *a* and the noise.

4.4 Risk Evaluation

Before submitting user's sub-queries to the query plan executor, repository cloud allows user to express his privacy preference by setting risk values on the query plan graph. Given a query plan graph *G*, let r_{ii} be the risk value on intermediate searching result v_i, and r_{ij} be the risk value on the relationship between the intermediate searching results v_i and v_j.

Each risk value is an integer ranging from 0 to 9, or $+\infty$. The risk value on a node represents how sensitive the corresponding data is, and the risk value on a link denotes how sensitive the corresponding relationship is. Let $R(G_k)$ be the risk value of sub-query *k*, we define $R(G_k)=max(r_{ii}, r_{ij})$ where $v_i \in G_k$ and $e_{ij} \in G_k$. In brief, the risk value of a sub-query equals to the maximum risk value in it.

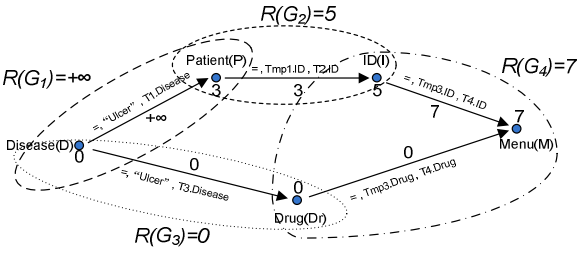


Fig. 3. An instance of risk evaluation

Based on the motivating example, we provide an instance of user risk evaluation in Fig. 3. We can see from Fig. 3 that the value "0" is assigned to several nodes and links. Such low values indicate that the user does not concern whether those information can be protected or not. Whereas that user cares which patients suffer from which diseases, as this kind of data disclosure may offend patients' privacy. Thus, the risk value on the relationship between "Disease" and "Patient" is set to be infinity.

4.5 Query Allocation

Based on the above analysis, the query allocator provides several strategies for user considering both price and privacy. User could select a service with lowest cost allocation which subjects to privacy risk constraint, or a service with lowest privacy protection which subjects to cost constraint, or a service considering trade-off between cost and privacy penalty. Based on user's choice, every possible allocation is examined and the best allocation will be selected.

As the sub-queries are relevant to each other, we need to consider the interdependency of the schemes applied on the sub-queries. In detail, scheme 3 and 4 cannot be allocated as the successor of scheme 1 as both of those schemes require normal data be the input. Since scheme 4 can be the predecessor of any scheme, and also provide nearly perfect privacy protection by adding large amount of noise, it is feasible to use scheme 4 as a substitution of scheme 1. Let $A(G_i) \in \{1,2,3,4\}$ be the allocation decision for sub-query G_i, where the value $A(G_i)$ denotes the scheme applied to G_i. Let G_{i-1} be the predecessor of sub-query of G_i, and G_{i+1} be the successor sub-query of G_i. Following rules are used to enforce query allocation:

- ✓ If $A(G_{i-1}) = 1$, $A(G_i) \in \{1,2\}$; otherwise, $A(G_i) \in \{1,2,3,4\}$.
- ✓ Given $R(G_{i+1}) = 0$, if $C_1(i) + C_2(i+1) > C_4(i) + C_3(i+1)$, set $A(G_i) = 1$; otherwise, set $A(G_i) = 4$.
- ✓ Given $R(G_i) = +\infty$ and $R(G_{i+1}) \neq 0$, set $A(G_i) = 4$.
- ✓ Given $R(G_i) = 0$, if $A(G_{i-1}) = 1$, set $A(G_i) = 2$; otherwise, set $A(G_i) = 3$.
- ✓ Given $R(G_i) \neq 0 \,\&\, R(G_i) \neq +\infty \,\&\, R(G_{i+1}) \neq 0$, set $A(G_i) = 4$.

The first rule represents the scheme sequence constraint for query allocation. Other rules indicate how the scheme will be chosen in different situations.

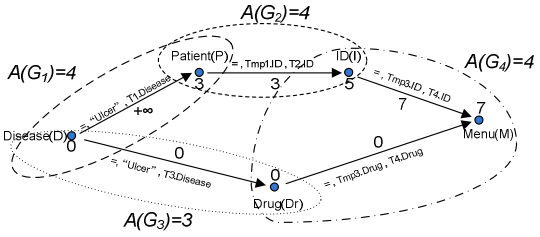


Fig. 4. Allocation Results

Risk and cost have been considered in the allocation process. Above rules guarantees that user's privacy requirements are always fulfilled, while the cost is also optimized. By adopting above rules, the sub-queries in motivating example can be allocated, where the results are presented in Fig.4. The final allocation result in Fig.4 illustrates that user's sub query 1, 2, and 4 are applied scheme 4, while query 3 is applied scheme 3. Checking the query execution schemes in Table 3, we can see that all the queries should be executed in local cloud in which guarantees user's privacy can be protected while the cost is also optimized.

6 Conclusions and Future Work

In this paper we presented a more flexible and complicated approach for privacy protection considering user's privacy preference and the cost for protecting data. We also provided an efficient method for allocating the proposed schemes. On the implemented system, performance comparisons are considered and evaluated. The experimental results could not be presented in this paper for the limitation of page number. In the future work, we will apply the interface for more conflicted situation like how the system works under the trustable service providers. Another possible future work is to provide the interface for our system implementation. On the implemented system, performance and storage efficiency will be considered and evaluated.

Acknowledgement

"This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2011-(C1090-1121-0003). "This work was supported by the IT R&D program of MKE/KEIT [K10035321, Terminal Independent Personal Cloud System].

References

1. Hand, E.: Head in the clouds. Nature 449, 963 (2007)
2. Tian, Y., Song, B., Huh, E.-n.: Inter-cloud Data Integration System Considering Privacy and Cost. In: ICCCI 2010 (2010) (in press)

3. Yeo, C.S., Venugopal, S., Chu, X., Buyya, R.: Autonomic Metered Pricing for a Utility Computing Service (2008)
4. Yau, S.S., et al.: a privacy preserving repository for data integration across data sharing services. IEEE Transactions on Services Computing
5. Ye, S., Wu, F., Pandey, R., Chen, H.: Noise Injection for Search Privacy Protection. In: International Conference on Computational Science and Engineering, vol. 3, pp. 1–8 (2009)
6. Foster, I., Kesselman, C. (eds.): The Grid 2: Blueprint for a new computing infrastructure. Morgan Kaufmann, San Francisco (2003)
7. Canfora, G., Costante, E., Pennino, I., Visaggio, C.A.: A three-layered model to implement data privacy policies. In: Computer Standards & Interfaces, pp. 398–409. Elsevier Science Publishers B.V., Amsterdam (2008)
8. Yeoa, C.S., Venugopalb, S., Chua, X., Buyya, R.: Autonomic metered pricing for a utility computing service. Future Generation Computer Systems (2009)
9. Singh, T., Vara, P.K.: Smart Metering the Clouds. In: 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, pp. 66–71. IEEE Computer Society, Washington, DC
10. Privacy in the clouds: risks to privacy and confidentiality from cloud computing, `http://www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf` (accessed on April 1, 2010)
11. Buyya, R., Yeo, C.S., Venugopal, S.: Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities. In: HPCC 2008, pp. 5–13 (2008)

An Applicable Model Enabling IPTV Service Delivery over Virtual Network

Biao Song, Mohammad Mehedi Hassan, and Eui-Nam Huh

Department of Computer Engineering,
Internet Computing and Network Security Lab,
KyungHee University Global Campus, South Korea
{bsong,hassan,johnhuh}@khu.ac.kr

Abstract. In this work, an applicable model enabling IPTV service providers to use a virtual network (VN) for IPTV service delivery is designed. The current topologies used in backbone or IP overlay-based content networks to deliver IPTV services cannot be directly applied as VNTs since virtual node, link, or video source failure on those topologies will cause service interruptions for customers. Therefore, in this work, a more reliable virtual network topology for solving a single virtual node, virtual link, or video server failure problem is designed. A novel optimization objective and an efficient VN construction algorithm have been developed for building the proposed topology. Various simulations were conducted to verify the effectiveness of the proposed VNT as well as the associated construction algorithms in terms of reliability and efficiency.

Keywords: IPTV service delivery, virtual network topology, VN allocation, VN node or link failure.

1 Introduction

Internet protocol television (IPTV) services are becoming more and more popular among communication and entertainment companies since they support the delivery of today's most sought-after video services while establishing a path for future next-generation applications [1]. As such, IPTV providers require the deployment of on-demand, reliable, and cost-effective service delivery networks to deliver their services. Current approaches [2-3] utilize a backbone or IP overlay-based content network for IPTV service delivery. Meanwhile, the virtual network (VN) [4] has been proposed as a promising way to overcome the limitations of backbone or IP overlay networks in terms of guaranteed service delivery, cost effectiveness, flexible control, and scalable network infrastructure [5]. From the perspective of an IPTV service provider, the delivery of IPTV services over VNs is beneficial for the following reasons:

- The use of a VN can potentially provide a desirable level of flexibility and control over the required content delivery network. Without the need to maintain an infrastructure, IPTV service providers are released from the task of managing infrastructure networks. Thus, the providers can save substantial amounts of

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 397–406, 2011.

money and focus on methods to deliver content. The use of a VN also effectively lowers the threshold for IPTV market entry.

- While multiple VNs may coexist on a shared substrate network, all of them are isolated from each other. Thus, a VN is a secure and reliable solution for IPTV content delivery service since each IPTV provider is allowed to run its own protocol without interrupting other providers.

The major challenge involved in utilizing a VN-based IPTV service delivery network is to design an efficient, cost effective, and reliable virtual network topology (VNT) for IPTV service delivery. In current solution [5] there is usually a lack of precise and comprehensive consideration toward topology design for VN-based IPTV service delivery. In addition, topologies used in backbone-based IPTV content delivery networks [2-3], such as star or tree topologies, cannot be directly applied as a VNT for delivering IPTV service because any virtual node or link failure within these topologies may cause all VN topologies to be disconnected. The virtual node or link may fail when the infrastructure providers (InPs) fall short of meeting the Quality of Service (QoS) requirement of a VN due to the failure of the physical or virtual router at runtime. Such a failure cannot be resolved by a rerouting algorithm because of insufficient physical network resources for the VN.

In this paper, the two aforementioned challenges to deliver IPTV services over a virtual network are discussed and candidate solutions are presented. The main contributions of this paper are as follows:

- We propose a novel and realistic VNT called Multi-Depots Ring-Star with 2-Depots per Ring (MDRS-2DR) for supporting IPTV content delivery. In MDRS-2DR, multiple rings are built for connecting depot and customer nodes. This setting can solve a single virtual link or node failure problem. Moreover, two depots (local video servers) are put in each ring so that the failure of a single server will not cause the loss of a video source. However, we still face the challenges such as the need for a new optimization objective for the MDRS-2DR topology and a new construction algorithm to meet the new optimization goal.
- To construct the MDRS-2DR topology, we use a new optimization metric that includes i) the distance between each customer and two depots, ii) the node, link and server failure probability and iii) the penalty of service interruption. Based on the proposed metric, a cluster-first route-second construction algorithm is constructed to build the MDRS-2DR topology.

The remainder of this paper is organized as follows. The system architecture is described in Section 2 and the proposed VNT associated with the new optimization metric and the VNT construction algorithm are presented in Section 3. Section 4 details the simulation environment and results. Finally, we present some concluding remarks in Section 5.

2 VN-Based IPTV Service Delivery Network

As shown in Fig. 1, various types of servers are connected through a VN. The global video server is considered as the original video source of IPTV content. In a real scenario, the global video server may be a TV station or another type of content

producer. The content is stored in its content server and distributed through its streaming server. The global video server distributes contents to local video servers through a virtual network which adopts star topology. Since it is not cost-efficient to store all of the replicated content on each local video server, the replica allocation algorithm in [2] should be adopted to decide which local video servers should obtain the content replicas. Connecting global video server with local replica servers can be viewed as stage 1.

Each large area, such as a city, may need one or more local video servers to store the content, accomplish trans-coding tasks, and forward the content to the local cache server. Each local video server should be able to obtain the content from either the global server or from other local video servers that have the replicated content (video source). The MDRS-2DR topology will be adopted as the VNT to guarantee that each local video server is connected to two video sources. Thus, when any single node (except the Global video server) or link fails, the VN still provides a path for the local video server to receive the content. Connecting local video servers with video sources can be viewed as stage 2.

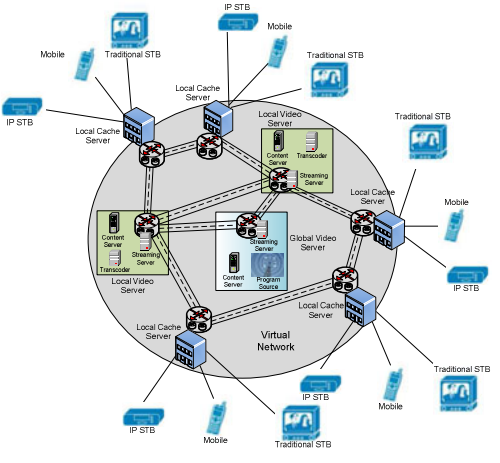


Fig. 1. VN-based network for delivering IPTV services

In order to reduce the load on the local video servers and shorten the response time, many local cache servers are allocated to various small local areas close to the end-users. Through a mobile, IP, or traditional set-top box (STB), users receive the content from the local cache server and play it on their own devices. Each local cache server should be able to obtain the content from a minimum of two local video servers. Thus, we again build the MDRS-2DR topology for distributing the content from local video servers to local cache servers. This stage can be viewed as stage 3.

3 MDRS-2DR Topology

This section introduces a novel VNT that can be used in stage 2 and stage 3. We first provide the mathematical description of proposed MDRS-2DR topology, and then present our optimization objective for the MDRS-2DR construction algorithm. The MDRS-2DR construction algorithm is also described in this section.

3.1 Mathematical Description of MDRS-2DR Topology

We begin by defining the problem with a graph theory model. Analogous to the terminology used in the routing literature, we will call the video sources depots. Thus, we call our problem the Multi-Depot Ring-Star Problem with two Depots per Ring (MDRSP-2DR). The MDRSP-2DR consists of designing a set of node-disjoint simple cycles, each one passing through two and only two depots and through some other nodes representing customers. The total number of customer nodes visited by a cycle is bounded by an upper capacity limit Q and the total number of cycles emanating from each depot is limited by an upper limit. The objective is to shorten the total distance between customers and depots. More precisely, the MDRSP-2DR is as follows. Consider a graph $G = (V, E)$, in which $V = D \cup V'$ is the set of nodes and $E = \{\{V_i, V_j\} : V_i, V_j \in V, i \neq j\}$ is the set of edges. Node set D represents the depots, while node set V' denotes the customers. Each edge $e\{V_i, V_j\} \in E$ is associated with a non-negative connection cost d_{ij} which is related to the distance between node V_i and node V_j.

Given a subset $E' \in E$, let $G' = (V(E'), E')$ denote the induced sub-graph of graph G. A 2-depot ring is a pair (D', E') where: (i) $D' \subseteq D$ and $|D'| = 2$ denotes that D' always contains two depots and (ii) $E' \subseteq E$ is a set of routing edges that form a simple cycle containing D'. Given that graph G consists of several 2-depot rings, if each customer in V' is assigned to exactly one 2-depot ring, the MDRS-2DR topology is said to be complete.

3.2 Optimization Objective

To define the optimization objective function for the MDRSP-2DR, we need to clarify how the content is sent to a customer in a 2-depot ring. Since there are two possible routes in a ring, a customer node normally receives content from the route, which is comparatively shorter than the reverse route. However, each single node, edge, and depot on that route may fail. If a failure causes the termination of content delivery, the customer will shift to the reverse route and receive content from the other depot. We assume that the probabilities of a single node, edge, and depot failure are known as p_n, p_e, and p_d. Suppose that single node, edge, and depot failures are inter-independent random events. Given a customer node V_i' and two depots $\{D_{k_1}, D_{k_2}\}$ in a 2-depot ring, let d_{ik_1} and d_{ik_2} ($d_{ik_1} \leq d_{ik_2}$) be the routing distance between V_i' and D_{k_1}, and V_i' and D_{k_2}, respectively. Also, let h_{ik_1} and h_{ik_2} be the number of hops between V_i' and D_{k_1}, and V_i' and D_{k_2}, respectively. For customer V_i', the probability of receiving content from D_{k_1} is denoted as:

$$p_{ik_1} = (1-p_n)^{h_{ik_1}} \times (1-p_e)^{h_{ik_1}} \times (1-p_d) \quad (1)$$

where $(1-p_n)^{h_{ik_1}}$, $(1-p_e)^{h_{ik_1}}$, and $(1-p_d)$ represent the probabilities that no failure happens to the nodes on the route, the edges on the route, and the depot D_{k_1}, respectively. The probability of receiving content from D_{k_2} is denoted as:

$$p_{ik_2} = (1-p_{ik_1}) \times (1-p_n)^{h_{ik_2}} \times (1-p_e)^{h_{ik_2}} \times (1-p_d) \quad (2)$$

where $(1-p_{ik_1})$ is the probability that D_{k_1} cannot serve V_i', while $(1-p_n)^{h_{ik_2}}$, $(1-p_e)^{h_{ik_2}}$, and $(1-p_d)$ represent the probabilities that no failure happens to the nodes on the route, the edges on the route, and the depot D_{k_2}, respectively. In a 2-depot ring, the distance between customer node V_i' and the depots can be denoted as $d_{iD'}$, where:

$$d_{iD'} = p_{ik_1} d_{ik_1} + p_{ik_2} d_{ik_2} + (1-p_{ik_1}-p_{ik_2}) penalty,\ D' = \{D_{k_1}, D_{k_2}\} \subseteq D,\ d_{ik_1} \le d_{ik_2} \quad (3)$$

For $S \subseteq V$, let $\delta(S) \subseteq E$ be the cut of the set S defined as $\delta(S) \subseteq \{\{V_i, V_j\} \in E : V_i \in S, V_j \notin S\}$. If $S = \{V_i\}$, we simply write $\delta\{i\}$ instead of $\delta(\{V_i\})$. For each edge $e \in E$ and each depot $D_k \in D$, let x_e^k be a binary variable with a value of 1 if and only if e belongs to a ring emanating from depot D_k. Henceforth, if e connects the two nodes V_i and V_j, then $\{V_i, V_j\}$ and e will be used interchangeably to denote the same edge. If a customer V_i' is assigned to a ring containing depot set D', let $z_i^{D'} = 1$; otherwise, let $z_i^{D'} = 0$. If a customer i is assigned to a ring containing depot D_k, let $z_i^k = 1$; otherwise, let $z_i^k = 0$. The optimization of the MDRSP-2DR topology can be formulated as follows:

$$\min \sum_{D' \subseteq D} \sum_{i \in V'} d_{iD'} \times z_i^{D'} \quad (4)$$

3.3 MDRS-2DR Construction Algorithm

In this section, new heuristic methods for solving the MDRS-2DR construction problem are proposed. The approach consists of two major stages: cluster-first and route-second. In the clustering stage, the k-means clustering algorithm is used to classify all customers. For each customer group, two depots are assigned to provide the content. After grouping, several rings are constructed to connect the customers and depots in the routing stage.

3.3.1 Clustering Stage

The clustering stage partitions the customers and depots into many groups. The customers in each group are served by two depots. The main goal of this stage is the minimization of an objective function taken as the distance between all customers and depots from their respective cluster centers. We adopt a constrained k-means

algorithm to efficiently cluster customer sets. Two depots are then selected and assigned for each customer group based on the probability function of the customer density. As the constrained k-means algorithm is a well-known algorithm, we do not explain this part. The details of constrained k-means algorithm can be found in [8]. The remaining of this sub-section introduces a novel two-depot assignment algorithm.

Assume that we have already separated our customers using the constrained k-means algorithm. Thus, we have k sets, $S_1,\ldots,S_k$, with the customers in S_j having been drawn independently according to the probability law $p(\mathbf{x}|\omega_j)$. We say that such samples are independent identically distributed random variables. We also assume that $p(\mathbf{V}'|\omega_j)$ has a known parametric form and is therefore determined uniquely by the value of a parameter vector $\boldsymbol{\theta}_j$. The customer set S_j contains n samples, $\mathbf{V_1'},\ldots,\mathbf{V_n'}$. Since the samples were drawn independently, we have

$$p(S_j|\boldsymbol{\theta}_\mathbf{j}) = \prod_{i=1}^{n} p(\mathbf{V_i'}\,|\,\boldsymbol{\theta}_\mathbf{j}) \tag{5}$$

In (5), $p(S_j|\boldsymbol{\theta}_\mathbf{j})$ is called the likelihood of $\boldsymbol{\theta}_\mathbf{j}$ with respect to the set of samples. The maximum likelihood estimate of $\boldsymbol{\theta}_\mathbf{j}$ is, by definition, the value $\boldsymbol{\theta}_\mathbf{j}$ that maximizes $p(S_j|\boldsymbol{\theta}_\mathbf{j})$. In each group, it is assumed that the customer density function can be expressed as a 2-D Gaussian distribution function, i.e., $p(S_j\,|\,\boldsymbol{\theta}_\mathbf{j}) \sim N(\boldsymbol{\mu}_\mathbf{j}, \boldsymbol{\Sigma}_\mathbf{j})$, where $\boldsymbol{\theta}_\mathbf{j}$ consists of the components of $\boldsymbol{\mu}_\mathbf{j}$ and $\boldsymbol{\Sigma}_\mathbf{j}$. Thus, the problem is to estimate the unknowns $\boldsymbol{\mu}_\mathbf{j}$ and $\boldsymbol{\Sigma}_\mathbf{j}$ using the customer coordinate values in S_j. According to the theory of maximum likelihood estimation, the maximum likelihood estimates for $\boldsymbol{\mu}_\mathbf{j}$ and $\boldsymbol{\Sigma}_\mathbf{j}$ are given by:

$$\boldsymbol{\mu}_\mathbf{j} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{V_i'} \tag{6}$$

$$\boldsymbol{\Sigma}_\mathbf{j} = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{V_i'} - \boldsymbol{\mu}_\mathbf{j})(\mathbf{V_i'} - \boldsymbol{\mu}_\mathbf{j})^t \tag{7}$$

Apparently, the maximum likelihood estimate for $\boldsymbol{\mu}_\mathbf{j}$ is $\mathbf{Z_j}$, which we found with the constrained k-means algorithm. Based on $\boldsymbol{\mu}_\mathbf{j}$, $\boldsymbol{\Sigma}_\mathbf{j}$ can be calculated from (7). Since the density function $N(\boldsymbol{\mu}_\mathbf{j}, \boldsymbol{\Sigma}_\mathbf{j})$ for customer cluster S_j can be estimated, we can also evaluate which two depots are appropriate for serving S_j. Given a depot D_x, we have

$$p(\mathbf{D_x}|\boldsymbol{\theta}_\mathbf{j}) = \frac{1}{(2\pi)\,|\,\boldsymbol{\Sigma}_\mathbf{j}\,|^{1/2}} \exp\left(-\frac{(\mathbf{D_x} - \boldsymbol{\mu}_\mathbf{j})^t\,\boldsymbol{\Sigma}^{-1}(\mathbf{D_x} - \boldsymbol{\mu}_\mathbf{j})}{2}\right) \tag{8}$$

Considering (8) as an individual objective function for choosing a single depot, another objective function should be applied to measure the relationship between two depots. In our case, an optimal situation is attained when each depot perfectly covers half of the customers in each cluster. Thus, a scenario where two selected depots are

symmetric about the cluster center is ideal. Given two depots D_x, D_y and one cluster center node Z_j, we can calculate the angle between two vectors $\overrightarrow{\mathbf{Z_jD_x}}$ and $\overrightarrow{\mathbf{Z_jD_y}}$ using the following formula:

$$angle(\overrightarrow{\mathbf{Z_jD_x}}, \overrightarrow{\mathbf{Z_jD_y}}) = \arccos(\mathbf{V}_1 \bullet \mathbf{V}_2) \tag{9}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are the normalized vectors of $\overrightarrow{\mathbf{Z_jD_x}}$ and $\overrightarrow{\mathbf{Z_jD_y}}$, respectively, such that $|\mathbf{V}_1| = |\mathbf{V}_2| = 1$. The range of $angle(\overrightarrow{\mathbf{Z_jD_x}}, \overrightarrow{\mathbf{Z_jD_y}})$ is between 0°, which is the worst case, and 180°, which is the best case.

Combining the probability density function and the co-relation function, we can obtain an overall optimization objective:

$$O_{jxy} = \alpha \times (p(\mathbf{D_x}|\boldsymbol{\theta}_\mathbf{j}) + p(\mathbf{D}_y|\boldsymbol{\theta}_\mathbf{j}))/2 + \beta \times angle(\overrightarrow{\mathbf{Z_jD_x}}, \overrightarrow{\mathbf{Z_jD_y}})/180 \tag{10}$$

where α and β represent the weight coefficients of the probability density function and co-relation function, respectively.

Let (Z_j, D_x, D_y, O_{jxy}) be a possible assignment. The set of possible assignments R is defined as:

$$R = \{(S_j, D_x, D_y, O_{jxy}) \mid x \neq y\} \quad \forall j, x, y \tag{11}$$

Based on the value of O_{jxy}, R is sorted in descending order. For each cluster of customers, we choose one assignment result from the sorted R using the greedy method. The constraints are: i) one and only one assignment for each cluster and ii) the number of clusters associated with a certain depot D_x cannot exceed m_x.

3.3.2 Routing Stage

The main goal of the routing stage is to find a closed cycle for each cluster of customers associated with two depots. The customer-based and failure-aware distance in Eq. (4) is suggested as a more proper metric for the scenario in this study. In our distance metric, the cost of the connections, which are adjacent to the depots, will be counted several times. To attain better optimization results, these connections should be selected and optimized prior to other connections.

Based on the above statements, a Nearest Customer Extension (NCE) algorithm is proposed to solve the routing problem for the MDRSP-2DR. The NCE is a heuristic algorithm that produces semi-optimal results within polynomial time. The NCE algorithm can be expressed as follows.

Using D_1 and D_2 as starting points, we create four chains C_1, C_2, C_3, and C_4 by connecting the nearest customer to the end of the chains. Initially, $H_1 = D_1$ and $H_2 = D_1$ are defined as two ends of the chain emanating from D_1, while $H_3 = D_2$ and $H_4 = D_2$ are defined as two ends of the chain emanating from D_2. Meanwhile, let $\gamma_1 = 0$ and $\gamma_2 = 0$ be the lengths of the chain emanating from D_1, and $\gamma_3 = 0$ and $\gamma_4 = 0$ be the lengths of the chain emanating from D_2.

The NCE algorithm contains n rounds. In each round, it calculates the distances from every $V_i \in V^{\pi}$ to D_1 and D_2 through different chains. By connecting V_i with H_1, H_2, H_3, and H_4, we obtain the distances $\kappa_{i1} = d_{iH_1} + \gamma_1$, $\kappa_{i2} = d_{iH_2} + \gamma_2$, $\kappa_{i3} = d_{iH_3} + \gamma_3$, and $\kappa_{i4} = d_{iH_4} + \gamma_4$, respectively.

$$\kappa_{kl} = \min\{\kappa_{ij} \mid V_i \in V^{\pi}, 1 \le j \le 4\} \tag{12}$$

Let κ_{kl} in (12) be the minimum distance among all of the distances. It is then decided that customer V_k is to be served through chain C_l. The NEC algorithm updates $H_l = V_k$, $\gamma_l = \kappa_{kl}$, $V^{\pi} = V^{\pi} - \{V_k\}$, and $V^{\mu} = V^{\mu} \cap \{V_k\}$. After all of the customers are included in the chains, the end nodes of two chains must be connected. If $d_{H_1H_3} + d_{H_2H_4} < d_{H_1H_4} + d_{H_2H_3}$, H_1 is connected to H_3 and H_2 is connected to H_4; otherwise, H_1 is connected to H_4 and H_2 is connected to H_3.

4 Simulations

We first compare the reliability of the proposed topology with existing topologies. The optimization capability of the proposed VNT construction algorithm is then compared with an existing VNT construction algorithm that uses a TSP solver [6-7].

4.1 Reliability

We compare our proposed VNT with four types of existing topologies. Same amount of customer nodes are severed by using each topology. Let p_l, p_n and p_d be the hourly failure probability of each virtual link, virtual node and depot respectively. We used p_l =0.5%, p_n =0.5% and p_d =0.05% in our simulation.

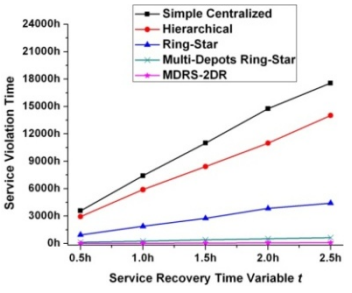


Fig. 2. Service recovery time vs. service violation time

Any failed virtual node or link needs *t* hours to be repaired, while any failed depot node needs 5*t* hours to be recovered. Thus, when redirecting is not possible, the customer node must wait *t* hours, 5*t* hours, or longer before the service can be restored. By varying the value of *t*, we simulated five rounds over a total of 50,000

hours. In each round, we counted the total service violation time for all customers caused by the use of different topologies.

The comparison results are shown in Fig. 2. As evident in the figure, the simple centralized topology is the worst among all of the topologies. Since customers using this topology suffer from the failure of a single node, link, and depot, the total service violation time is around 18,000 hours. The hierarchical, ring-star, and multi-depot ring star topologies partially solve the failure problem and thus, a lower service violation time is obtained. Due to its capability of handling single node, link, and depot failure problems, the MDRS-2DR topology provides the highest reliability among all of the topologies.

4.2 Optimization Capability of the Construction Algorithm

In this simulation, we took the optimization metric presented in Eq. (4) as the comparison objective. Our proposed VNT construction algorithm and the algorithm proposed in [6] were simulated for constructing MDRS-2DR topology. The main difference is that [6] use TSP solver to solve routing problem while our construction algorithm contains a heuristic to achieve the same goal.

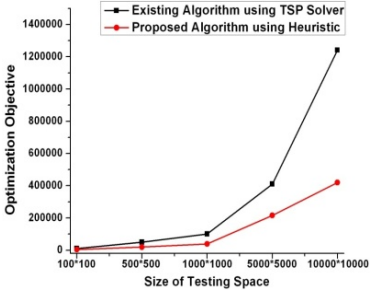


Fig. 3. Size of testing space vs. optimization objective

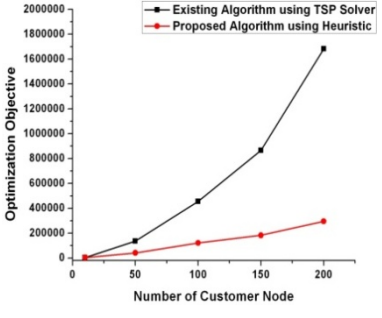


Fig. 4. Number of customer node vs. optimization objective

By varying the size of the testing 2-D space, we conducted our first simulation. The number of customer nodes and depot nodes were fixed as 50 and 5 respectively. They were uniformly generated in the testing space. The size of the testing space was

varied from 100*100 to 10,000*10,000. The comparison results of the first simulation can be found in Fig. 3. As we can see, the optimization capability of our proposed algorithm is much better than existing one. The main reason is that our proposed algorithm is specially designed for two-depot ring topology, thus performs better than existing approach.

In our second simulation, we varied the number of customer nodes from 10 to 200. The size of testing space was 1,000*1,000. We also fixed the number of depot as 5. The comparison results are presented in Fig. 4. We can observe that existing construction algorithm is less efficient than our proposed algorithm. With the number of customer node increases, the gap between our approach and existing one becomes lager and lager. Our proposed algorithm greatly shortens the distance between customer nodes and depot nodes for MDRS-2DR.

5 Conclusions

In this work, a VN-based IPTV service delivery network containing a novel VNT was designed. The main contribution of this paper is a novel and reliable VNT called Multi-Depots Ring-Star with 2-Depots per Ring. This topology provides a proactive solution to solve a single virtual node, virtual link, or single server failure problem that may occur during runtime. Based on the MDRS-2DR topology, we proposed a novel optimization objective that considered the failure probability and QoS for all customers. According to the proposed metric, an efficient topology construction algorithm was developed.

Acknowledgement

“This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)” (NIPA-2011-(C1090-1111-0001)).

References

[1] Yarali, A., Cherry, A.: Internet protocol television (IPTV). IEEE Tencon, 1–6 (2005)
[2] Li, M., Wu, C.-H.: A cost-effective resource allocation and management scheme for content networks supporting IPTV services. Computer Communication 33, 83–91 (2010)
[3] Zhang, X., Liu, J., Li, B., Yum, T.: CoolStreaming/DONet: A data-driven overlay network for efficient live media streaming. In: Proceedings of IEEE INFOCOM, vol. 3, pp. 13–17 (2005)
[4] Feamster, N., Gao, L., Rexford, J.: How to lease the internet in your spare time. SIGCOMM Comput. Commum. Rev. 37, 61–64 (2007)
[5] Virtualization in the Core of the Network, Juniper networks, White Paper (2010), http://www.juniper.net/us/en/local/pdf/whitepapers/2000299-en.pdf
[6] Baldacci, R., Dell’Amico, M.: Heuristic algorithms for the multi-depot ring-star problem. European Journal of Operational Research 203, 270–281 (2010)
[7] Naji-Azimi, Z., Salari, M., Toth, P.: A heuristic procedure for the Capacitated m-Ring-Star problem. European Journal of Operational Research 207(3), 1227–1234 (2010)
[8] Kennington, J., Helgason, R.: Algorithms for Network Programming. Wiley, New York (1980)

Virtualization of Access Network for Multihomed Wireless Thin Client

Md. Motaharul Islam, Jun-Hyung Lee, and Eui-Nam Huh

Department of Computer Engineering, College of Electronics and Information,
Kyung Hee University, 1 Seocheon-dong, Giheung-gu, Yongin-si,
Gyeonggi-do, 446-701, Republic of Korea
{motahar,junhyung,johnhuh}@khu.ac.kr

Abstract. An access network is the part of a communications network which connects the multihomed thin client to the immediate service provider. In a thin client computing architecture, application processing is delegated to a remote server rather than running it locally. User input is forwarded to the server, and the rendered images are relayed through a dedicated remote display protocol to the thin client. For this reason availability and seamless guaranteed connectivity to the internet are the key issues in multihomed wireless thin client. Traditional multi homed thin client architecture does not provide seamless service provisioning due to the unavailability of suitable handoff mechanism among the access network. Virtualization of wireless access network can be a promising solution for seamless connectivity to the multihomed thin client. In this paper we focus on the virtualization of access network for the multihomed wireless thin client architecture and the selection of appropriate access network with the special attention to user satisfaction.

Keywords: Access network, multihomed wireless thin client, virtualization, User satisfaction.

1 Introduction

Recently, wireless networking technologies are in widespread use and different types of wireless network service exists. These wireless services vary greatly in their coverage area and quality of service. The heterogeneity of radio access networks and almost every wireless thin client and other mobile host have multiple radio interfaces, which are introducing many challenges for the selection of the most appropriate access network [1].

Portable information appliances come in every size and shape. Today, users enjoy a wide variety of computing devices that promise to satisfy every imaginable need. These devices include different types of wireless thin client such as PDAs and handheld computes, smart phone, MP3 players, portables video players, and light notebooks. Wireless access networks offer wireless thin client users the benefits of mobile computing and global connectivity to the internet [2] [3].

Wireless thin client computing refers to the paradigm in which the user device relies on a remote server to perform a significant fraction of its computational tasks. It

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 407–414, 2011.

will provide the user with a GUI but with considerably lower total cost of ownership compared to traditional fat client. Especially in the mobile context, the wireless thin client concept is very promising. Users are able to access demanding applications from mobile devices, which often lack the required processing resources to execute the application locally. Because only basic functionality and processing power is required at the terminal, wireless thin client devices can be made lightweight and potentially energy efficient [3][4].

Instead of different promising applications of multihomed wireless thin client, if it can not communicate seamlessly with the server then it becomes worthless. Now a day, internet connectivity is available almost everywhere. A multihomed wireless thin client can use different types of network such as 3G, Wi-Fi, WiBro/WiMAX. But the important issue is how to select one of the best access networks without user intervention and following different types of transparency. To solve this challenge we focus on the virtualization of access network (VAN) in the network protocol stack. It ensures all form of transparency specially the access transparency. If a mobile multihomed wireless thin client user move from one place to another then the access network selection module will decide the best network for the user based on different decision factor such as received signal strength indicator, bandwidth cost, network condition, quality of service etc.

The rest of the paper is organized as follows. Section 2 reviews background related wireless thin client. The proposed VAN architecture, along with its system overview and scenarios are depicted in section 3. Section 4 describes evaluation, and finally section 5 concludes the paper.

2 Backgrounds

The typical wireless thin-client platform consists of a client application that executes on a user's local desktop machine and a server application that executes on a remote system. The end user's machine can be a hardware device designed specifically to run the client application or simply a low-end personal computer. The remote server machine typically runs a standard server operating system, and the client and server communicate across a network connection between the desktop and server. The client sends input data across the network to the server, and the server returns display updates.

Several wireless thin-client computing platforms have been developed so far. In general, these existing wireless thin-client systems can be classified into two categories according to the mechanisms of representing the display information [3] [4]. The first category of thin-client systems utilizes high-level commands to represent the screen update, such as the Citrix Metaframe [5], and Microsoft Remote Desktop Protocol (RDP) [6]. This method opts for a higher level encoding that is more closely tied to the operating system's windowing and display commands. In this case, graphics commands are transmitted from the server to the client, which is responsible for processing the updates. So the interpretation of high-level commands heavily depends on the operating systems, it is hard to develop the servers and clients on different operating systems with different rendering mechanisms. The second category of thin-client systems utilizes low-level approaches to represent the screen of

remote servers, including VNC [2] and Sun Ray systems [7]. They process updates to the display on the server and transmit only compressed pixel data representing the new display to the client. The most important is that they are platform-impendent. Now a day, multihomed thin client is increasingly getting popularity since it can render service to different types of network and wide geographical areas. Since a multihomed thin client has different network interface, it can access different network based on the situation and demand.

3 Proposed Architecture

3.1 VAN Architecture

We propose architecture called Virtualization of Access Network (VAN) which is depicted in figure 1. VAN architecture provides a virtual layer that controls multiple access network interfaces in order to optimize the multihomed wireless thin client resources. In our proposed architecture we introduced a virtual access layer which will provide a transparent access to individual access network. Virtual access layer is software based. The intent of software-based radios is to shift the hardware-oriented application to flexible software applications performing communication functions on a common computing platform. The communication functions of a transceiver are completely realized as programs operating on a suitable processor, this transceiver is referred to as software radio.

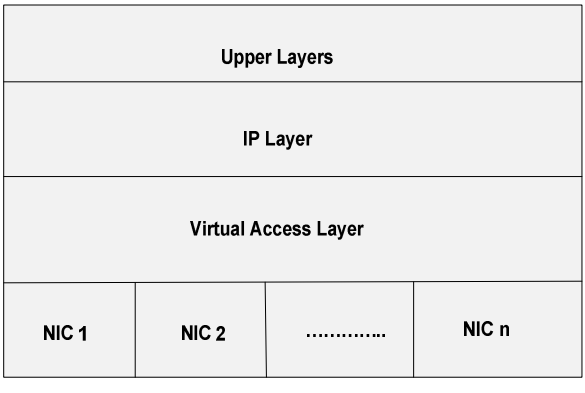


Fig. 1. VAN architecture

The digitization is done directly after radio wave reception at the antenna and all the signal processing is done by software. Thus virtual access layer provide many opportunities to access the network transparently. The wireless thin client can choose any better access network.

The switching among different access network will not hamper the ongoing communication. The scheduling mechanism in the virtual access layer will decide the switching among different network. Switching or handoff may be of two types: horizontal and vertical. In case of horizontal handoff, receive signal strength, cost

and QoS will be the decision criteria. For vertical handoff availability of the network condition and customer satisfaction will be the decision making criterion.

3.2 VAN System Overview

In figure 2 we briefly describe the VAN selection manager architecture. Its major components are explained as follows.

i. The Access Networks module provides the basic functionalities required to communicate to the Internet. I consist of different type of network such as 3G, Wi-Fi and WiBro etc. Selection of the access network by the multihomed wireless thin client device is based on the VAN selection module which gets some network status by access network monitor.

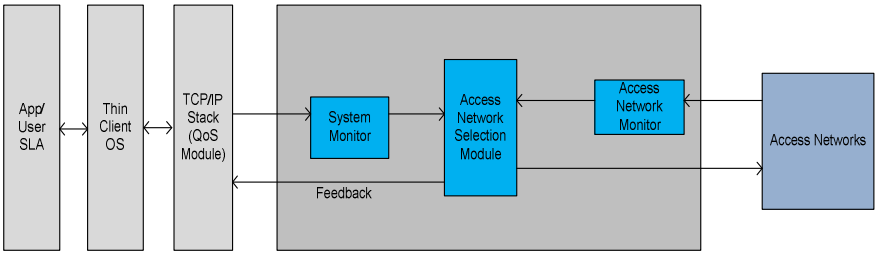


Fig. 2. Virtual Access Network Selection Manager

ii. The Access Network Monitor component monitors the access network status spatially the RSSI information and other networking information. It finally gives this information to the access network selection module.
iii. The Access Network Selection Module takes network selection and handoff decision. This is the most important component in VAN architecture. It takes decision based on the access network monitor component and system monitor component. It also gives some feedback to the QoS module of the TCP/IP stack. The feedback will enrich the user centric information for customer satisfactions.
iv. The QoS Module of the TCP/IP stack is a special component of the virtual access layer of the VAN architecture. It bridges up the communication between the upper layer and lower in the stack. It helps the access network selection module. Since access network selection module can not use the upper layer information directly, this module gives some user satisfaction information to the selection module. Thus it facilitates decision making process.
v. The Wireless thin Client Operating System (OS) represents the system software that helps it to connect to the remote server as well as to manage its own local resources. Traditionally, a thin client ran a full operating system for the purposes of connecting to other computers. A newer trend is sometimes called an ultra-thin client or a zero client, which no longer runs a full operating system: the kernel instead merely initializes the network, begins the networking protocol, and handles display of the server's output.

vi. The User application or user service level agreement (SLA) represents the service quality that will be available to the application users. Through this interface user can avail different types of services with predefined service quality.

3.3 Scenario

We imagine a scenario where several wireless networks are operating and they may overlap each other. A user may roam in these networks. While roaming several situations may occur. A user may change the cell of same type of network e.g. Wi-FI to Wi-Fi, which is called horizontal handoff. Horizontal handoff is easily achievable as the protocol itself provides the handoff mechanism. However, when the user moves to a different network e.g. Wi-Fi to 3G, technical as well as economic criteria come into play. We consider seamless network switching in this environment considering the contexts: mobility, heterogeneous networks, and cost of network connectivity. Figure 3 represents the scenario.

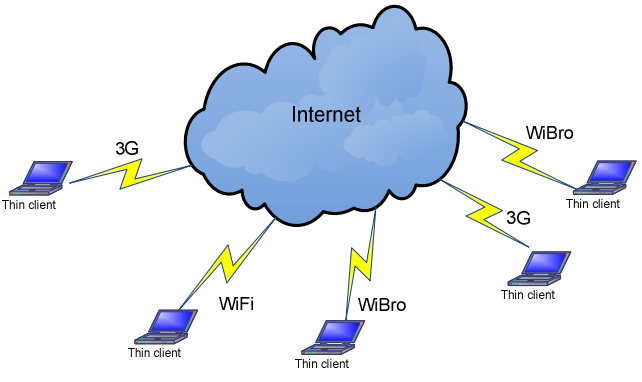


Fig. 3. Mobile multihomed thin client scenario

Suppose a thin client user is watching an online video stream on the Wi-Fi network. There is also a 3G network. However, due to high cost of 3G, the user did not use it. But suddenly the Wi-Fi becomes congested or packet error rate increases, so the user switches to 3G. Virtual access layer deals it in a transparent manner.

4 Evaluation

Multihomed thin client architecture consists of more than one network interfaces. It facilitates the multihomed wireless thin client for seamless connectivity with the external world. This omnipresent connectivity gives the end user a great flexibility and foredoom of choice. The network connectivity will ensure access and location transparency. In our proposed VAN architecture, virtual access layer is introduced that it will help providing the optimal network connectivity among the available wireless network. The network will be selected based on the RSSI value, Bandwidth cost, user satisfaction, QoS provision, network condition, cost and receive signal

strength etc [1] [8] [9]. To get the best access network we can use an objective function which may be stated as follows.

Access Network selection = f (RSSI, BW_Cost, Customer Satisfaction)

Where RSSI is the received signal strength indicator, BW_Cost indicate bandwidth cost, Customer satisfaction indicates users attitude towards the network. In the subsequent section we will discuss the individual decision parameter.

1. Received Signal Strength Indicator (RSSI) is a measurement of the power present in a received radio signal. RSSI is generic radio receiver technology metric, which is usually invisible to the user of the device containing the receiver, but is directly known to users of wireless networking. RSSI can be represented by Heaviside step function which is represented by following equations.

$$H(r) = \int \delta(t)\, dt$$

This can be written as

$$H(r) = \begin{cases} 0 & for\ r < R_{min} \\ 1 & otherwise \end{cases}$$

Here H(r) represents RSSI. RSSI is measured by the relative metric which ranges from 0 to 10. We assumed the minimum value R_{min} is 5 which is regarded as the threshold value. If RSSI value is less than the threshold value R_{min} then H(r) will return zero value which indicates coverage of the selected network is very poor. And it is better to switch another access network. And if the value is bigger than the thresh hold value then the coverage is better. Figure 4 represents the RSSI utility curve.

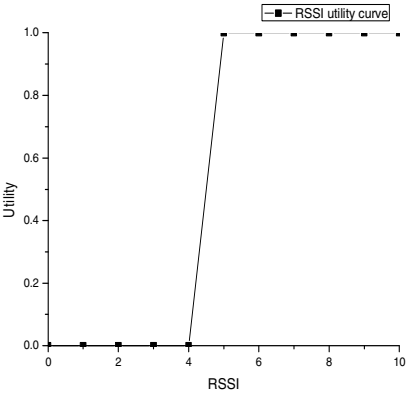


Fig. 4. RSSI utility curve

2. In computing bandwidth refers to the rate of data transfer, bit rate or throughput, measured in bits per second. Bandwidth cost plays an important role in selecting the access network because it varies on different network such as 3G, Wi-Fi, and WiBro/WiMax etc. Due to the unavailability of Wi-Fi coverage, sometimes expensive

bandwidth of 3G may be selected for the sake of seamless connectivity. Bandwidth cost may be represented by the following utility function.

$$B(c) = 1 - e^{\frac{c^2}{k+c}}$$

Here B(c) indicates the bandwidth cost function where k is a constant. For different types of applications value of k differs. For adaptive application we consider k=0.5. Figure shows the utility function of the bandwidth cost of an adaptive application is convex but not concave in neighborhood around zero. The point where the curve is convex depends on k. The larger the value of k, the larger the value of bandwidth where the curve is convex. Figure 5 depicts bandwidth utility curve.

3. Customer Satisfaction obtained from a network service depends on the performance of the network, which in turn depends on the capability of the access network to satisfy the bandwidth requirements of user's request. When a mobile wireless thin client user stays in a single network through out call duration, user satisfaction is equal to the bandwidth utility function of that network. However, if the user while holding a call moves from one network to other network its overall satisfaction is the function of bandwidth utility of each network and handoff latency.

$$S_u(t_h) = e^{-\frac{t_h^2}{2\sigma_h^2}}$$

Here σ_h is a constant that depends on application type; larger for non-real time (adaptive and elastic) and smaller for real time (rigid) applications. Figure shows the customer satisfaction versus switching delay. Here we consider three different values for σ_h for different application such as rigid, elastic and adaptive. Figure 6 shows the customer satisfaction versus Switching delay curve.

By evaluating the individual sub-functions such as RSSI, bandwidth cost and customer satisfaction, the main function *f* select the access network for a wireless thin client user. This decision is performed by the access network selection module of the VAN architecture.

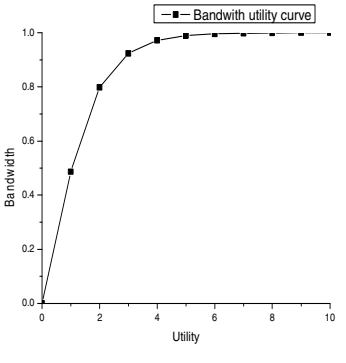


Fig. 5. Bandwidth utility curve

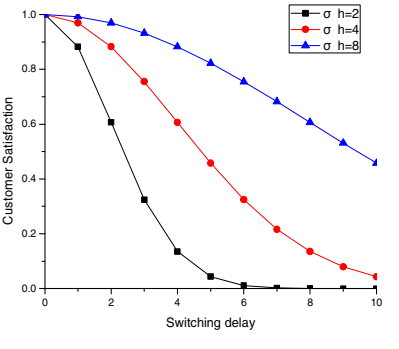


Fig. 6. Customer satisfaction versus Switching delay

5 Conclusion

Virtualization of access network for availability and seamless guaranteed services are the key concern in multihomed wireless thin client computing. The wireless thin client computing model puts all tasks on the shoulders of a powerful server that resides in the network. The client is responsible for receiving screen updates and rendering them on its display. Since the traditional multihomed wireless thin client architecture does not provide this type of service provisioning, we propose the VAN architecture that facilitates seamless connectivity to the network for the multihomed wireless thin client device. The proposed architecture introduces a virtual access layer which ensures better access network selection by virtual access selection manager as well as ensures service mobility.

Acknowledgement

This research was supported by Future-based Technology Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0020725). The corresponding author is Eui-Nam Huh.

References

1. Kim, J., Lee, J.H., Lee, H., Kwon, T., Choi, Y.: A network selection scheme for multicast applications in wireless network convergence. Wirel. Commun. Mob. Comput (2010), doi:10.1002/wcm.1032
2. Richardson, T., Stafford-Fraser, Q., Wood, K.R., Hopper, A.: Virtual Network Computing. IEEE Internet Computing 2(1) (1998)
3. Nieh, J., Yang, S.J., Novik, N.: A Comparison of Thin-Client Computing Architectures. Technical Report CUCS-022-00
4. Tang, W., Lee, J.H., Song, B., Islam, M.M., Sangho, N., Huh, E.N.: Multi-Platform Mobile Thin Client Architecture in Cloud Environment. In: International Conference on Innovative Computing and Communication, Macao, March 5-6 (2011)
5. Citrix Systems, "Citrix Metaframe 1.8 Backgrounder", Citrix White Paper (June 1998)
6. Microsoft Corporation, Remote Desktop Protocol, `http://msdn.microsoft.com/en-us/library/aa383015.aspx`
7. Sun Microsystems, Sun Ray 1 Enterprise Appliance, `http://www.sun.com/products/sunray1`
8. Kafle, V.P., Kamioka, E., Yamada, S.: Maximizing User Satisfaction Based on Mobility in Heterogeneous Mobile Multimedia Communication Networks. IEICE Trans. Commun. E88-B(7) (July 2005)
9. Simones, P., Ali, F.A., Vankeirsbilck, B., Deboosere, L., Truck, F.D.: Cross-Layer Optimization of Radio Sleep Intervals to Increase Thin Client Energy Efficiency. IEEE Communications Letters 14(12) (December 2010)

New Optimization Techniques for Shear-Warp Volume Rendering

Ki-Young Choi, Sung-Up Jo, and Chang-Sung Jeong*

School of Electrical Engineering, Korea University, Seoul, Korea
{2xx195,easyup,csjeong}@korea.ac.kr

Abstract. A shear-warp algorithm with run-length-encoded volume, one of the fastest CPU-based speed acceleration techniques developed so far for direct volume rendering, has various defects such as increase in memory consumption as well as preprocessing time and deterioration in image quality. This paper provides two kinds of new optimization techniques which can solve such defects without degrading rendering speed. One optimization technique concentrates on enhancing image quality while decreasing memory consumption without reducing rendering speed by making direct access to the memory space where the initially loaded volume data is stored. The other concentrates on decreasing preprocessing time and hence accelerating rendering speed by creating only one run-length-encoded volume and combining non-photorealistic rendering techniques with shear-warp algorithm. We shall show a novel result that both optimization techniques efficiently decrease the memory consumption and preprocessing time while enhancing rendering speed and image quality simultaneously.

Keywords: volume rendering, CPU-based speed acceleration technique, shear-warp algorithm, non-photorealistic rendering.

1 Introduction

Volume rendering [1, 2] is a technique for visualizing volume data on the screen. One of the most widely used volume rendering techniques is direct volume rendering, which calculates the color of screen pixels by processing voxels directly. Such techniques as ray-casting [2] and splatting [3] are based on this method. Major research areas on direct volume rendering can be classified into two fields: (1) acceleration techniques for rendering speed, and (2) rendering techniques for high quality image. Direct volume rendering results in slow speed, as it requires traversing the entire volume data whenever the viewing direction changes. CPU-based speed acceleration techniques used to overcome such speed limitation create and use the spatial coherence information of volume data, in order to skip the transparent region at high-speed. The most representative examples based on this approach are distance transforms [4], shear-warp algorithm with run-length-encoded volume [5] and shell structures [6]. Recently, many GPU-based speed acceleration techniques [7, 8] have

* Corresponding author.

J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185, pp. 415–422, 2011.

been proposed to implement real-time volume rendering for large size volume data with the help of high-performance H/W equipment. GPU-based speed acceleration techniques show excellent performance in rendering speed as well as in image quality. However, CPU-based speed acceleration techniques are still useful in some circumstances, especially when the portability among heterogeneous systems are important, high-performance H/W equipment such as GPU are not prepared or a moderately sized volume data is rendered. The shear-warp algorithm with run-length-encoded volume has some defects, such as the increases in memory consumption and in preprocessing time as well as the deterioration in image quality, yet it is one of the fastest algorithms among CPU-based acceleration techniques developed so far. Therefore, if only such defects are solved, shear-warp algorithm can be used as the most efficient and useful CPU-based speed acceleration technique for direct volume rendering. This paper proposes two kinds of new optimization techniques which can solve such defects. One optimization technique concentrates on enhancing image quality and decreasing memory consumption without reducing rendering speed by making direct access to the memory space where initially loaded volume data is stored. The other concentrates on decreasing preprocessing time and hence accelerating rendering time by creating only one run-length-encoded volume and combining non-photorealistic rendering (NPR) techniques with shear-warp algorithm. We shall show that both optimization techniques efficiently decrease memory consumption and preprocessing time of shear-warp algorithm, and produce the excellent results in rendering speed and image quality.

The contents of our paper are as follows: In section 2, the shear-warp algorithm with run-length-encoded volume is summarized and examined. In sections 3 and 4, two optimization techniques for shear-warp algorithm are proposed respectively. In section 5, performance evaluations are presented. Finally, in section 6, the conclusion is given.

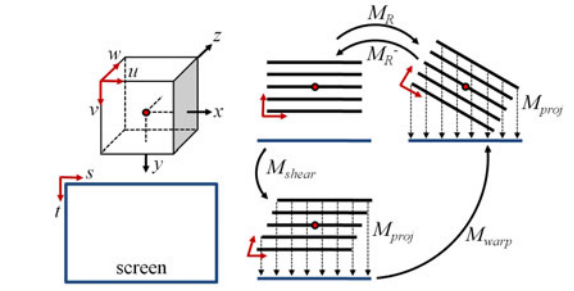
2 Shear-Warp Algorithm

2.1 Shear-Warp Algorithm

Shear-warp algorithm is a method applying shear-transform and warp-transform, instead of 3D rotation-matrix, to generate an image of rotated volume data.

$$\begin{aligned} M_{image} &= M_{proj} M_R(\theta_z) M_R(\theta_x) M_R(\theta_y) = M_R(\theta_z) M_{proj} M_R(\theta_x) M_R(\theta_y) \\ &= M_R(\theta_z) M_{skew} M_{scale} M_{proj} M_{shear} = M_{warp} M_{proj} M_{shear} \end{aligned} \quad (1)$$

Eq. 1 shows that the image obtained by applying the rotation-matrix M_R to the volume slices is the same as the image obtained by applying the shear-matrix M_{shear} and warp-matrix M_{warp} to the volume slices. When rendering a volume image by using shear-warp algorithm, the image rendered by applying M_{shear} and M_{proj} is called intermediate image, and the image created by applying M_{warp} to the intermediate image is called final image.



(a) three coordinate system (b) shear-warp algorithm

Fig. 1. Shear-warp algorithm

2.2 Fast Shear-Warp Algorithm

Shear-warp algorithm needs to create three different run-length-encoded volumes in the preprocessing stage. During the course of traversing the run-length-encoded volume to render an intermediate image, consecutive transparent voxel region can be skipped as fast as possible by using the run-length-encoding; and the compressed-opaque-voxel allows faster access to opaque voxels, as its allocated size is smaller than the size of the memory for entire volume data and all opaque voxels are arranged in the order they are traversed. However, they have the following defects.

Memory consumption & preprocessing time: Whenever the transfer-function which maps voxel intensity to opacity and RGB color is changed, shear-warp algorithm needs to create three different run-length-encoded volumes.

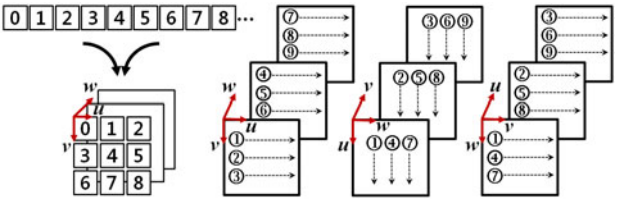
Resampling error: When compressed-opaque-voxel is decompressed to the four neighboring voxels to perform bilinear-interpolation, the values of the transparent voxels which are not included in the compressed-opaque-voxel are all set to 0. This results in resampling error, acting as a cause for the poor accuracy of the final image.

3 Shear-Warp Algorithm without Compressed-Opaque-Voxel

In order to resolve such problems, this section provides an optimized shear-warp algorithm that does not involve the creation and use of compressed-opaque-voxel. The optimized algorithm skips transparent voxel region by using optimized run-length-encoding, and directly accesses to the volume memory where the initially loaded volume data is stored to get voxels needed for bilinear-interpolation. As the optimized algorithm makes direct access to original volume data, neither is the decompression process needed nor do resampling errors occur, which reduces the memory consumption and preprocessing time. But, when the proportion taken up by opaque voxels in the entire volume data is very low, it may be more effective to use compressed-opaque-voxel for accelerating rendering speed.

3.1 Fast Direct Access to Volume Memory

In general, volume data is stored in 1-dimensional array of memory space. When voxels are stored in 1-dimensional array of memory space, the fastest way to traverse all voxels is to visit them in the order of their storage or of their reverse storage, since it does not cause back-and-forth movement between physically separated memory spaces, minimizing the traversing path. This can be easily shown and confirmed by simply traversing 1-dimensional array in various directions and in different order.



(a) voxel storage order (b) fast voxel traversal scheme for each principal viewing axis

Fig. 2. Fast voxel traversal scheme

3.2 Integrated Run-Length-Encoding

Instead of creating run-length-encoding by traversing voxels one by one, we created it by traversing four neighboring voxels simultaneously to optimize it for bilinear-interpolation; and the result was defined as integrated run-length-encoding. If the four neighboring voxels are all transparent, the current run state of the integrated run-length-encoding is defined as transparent; otherwise, it is defined as opaque. The use of integrated run-length-encoding has the advantage of eliminating the processes of traversing and comparing two neighboring run-length-encodings simultaneously in order to skip the transparent voxel region where all four neighboring voxels are transparent.

4 Shear-Warp Algorithm with One Run-Length Encoded Volume

This section provides an optimized shear-warp algorithm that uses mapping to the nearest pixel, without bilinear-interpolation in order to accelerate rendering speed. As this algorithm creates only one run-length-encoded volume by traversing volume data in the order of voxel storage, it efficiently reduces preprocessing time and memory consumption. Also, as it uses mapping to the nearest screen pixel by projecting one voxel (or one pixel) to its nearest corresponding pixel without bilinear-interpolation, there is no need to traverse and compare two run-length-encodings simultaneously and further speed acceleration is possible. But, as this algorithm combines automatic stippling technique [9] with warp-transform to hide and relieve the aliasing caused by mapping to the nearest pixel, it produces final image with NPR style which shows not the exact and detailed shape but rough shape of volume data. Therefore, this algorithm can be used efficiently for quick preview of volume data.

4.1 Creating Only One Run-Length-Encoded Volume

As described in Section 2, the fast shear-warp algorithm needs three run-length-encoded volumes. But, when rendering intermediate image by mapping each opaque voxel to its nearest pixel without bilinear-interpolation, shear-warp algorithm can be optimized so that it creates and uses only one run-length-encoded volume. When volume data is stored in 1-dimensional voxel array, the fastest way to traverse all voxels in order to create run-length-encoded volume is to visit voxels in the order of their storage. Therefore, in order to minimize preprocessing time, the run-length-encoded volume for the principal viewing axis $\pm w$ was selected and created. .

5 Experimental Results

In Table 1, the shear-warp 1 represents the fast shear-warp algorithm described in section 2; shear-warp 2, the shear-warp algorithm without compressed-opaque-voxel described in section 3; and shear-warp 3, the shear-warp algorithm with one run-length-encoded volume described in section 4. Table 1 shows the results of the three categories for four different volume data sets used in performance evaluation: memory consumption in Kbytes, preprocessing time in seconds and rendering speed in frame rate.

Table 1. Performance Evaluation

Data	Algorithm	Memory(Kbytes)	Preprocessing(sec)	Framerate
engine (256×256×110)	shear-warp 1	6424(100.00%)	0.984(100.00%)	32.55
	shear-warp 2	1309(20.38%)	0.469(47.66%)	32.55
	shear-warp 3	2588(40.29%)	0.125(12.70%)	66.66
carp (256×256×512)	shear-warp 1	65909(100.00%)	7.422(100.00%)	21.27
	shear-warp 2	1575(2.39%)	3.062(41.26%)	21.73
	shear-warp 3	23466(35.60%)	0.531(7.15%)	30.25
stag beetle (416×416×247)	shear-warp 1	18537(100.00%)	5.922(100.00%)	21.27
	shear-warp 2	2148(11.59%)	4.766(80.48%)	21.27
	shear-warp 3	7813(42.15%)	0.625(10.55%)	30.17
piggy bank (512×512×361)	shear-warp 1	84933(100.00%)	13.875(100.00%)	9.40
	shear-warp 2	5217(6.14%)	10.047(72.41%)	9.17
	shear-warp 3	32697(38.50%)	1.422(10.25%)	21.73

Memory consumption: As shown in Table 1, when the memory consumption of the shear-warp 1 is set to 100.0%, the ratio of the memory consumption of the shear-warp 2 decreased to the maximum of 2.39% and the minimum of 20.38%, while the shear-warp 3 decreased to the maximum of 35.60% and the minimum of 42.15%. The reason for the memory consumption of the shear-warp 3 being larger than that of the shear-warp 2 is that the shear-warp 3 includes compressed-opaque-voxel and the pointers to the internal position of the compressed-opaque-voxel that are not created by the shear-warp 2.

Preprocessing time: As shown in Table 1, when the preprocessing time of the shear-warp 1 is set to 100.0%, the ratio of the preprocessing time of the shear-warp 2

decreases to the maximum of 41.26% and the minimum of 80.48%. This is because the shear-warp 2 traverses the volume data three times in order to create three run-length-encodings, and does not create compressed-opaque-voxel. The ratio of the preprocessing time of the shear-warp 3 decreases to the maximum of 7.15% and the minimum of 12.70%, because the shear-warp 3 traverses the volume data only once in the order of voxel storage in order to create one run-length-encoded volume.

Rendering speed: As shown in Table 1, the shear-warp 2 does not create and use compressed-opaque-voxel for fast access to opaque voxels, the rendering speed of the shear-warp 2 is almost the same as that of the shear-warp 1. It also shows that the rendering speed of the shear-warp 3 is about two times as fast as that of the shear-warp 1 and shear-warp 2, enhancing the rendering speed in real time. The reason is that the shear-warp 3 makes use of mapping to the nearest pixel without bilinear-interpolation during rendering process.

Resulting Image: Fig. 3 shows the resulting images rendered by the fast shear-warp algorithm (top row) and the shear-warp algorithm without compressed-opaque-voxel (bottom row) described in section 2 and 3 respectively. Fig. 3 shows that the details of the original volume data, which are very clearly shown in the resulting images generated by the shear-warp algorithm without compressed-opaque-voxel, are not displayed well in the resulting images generated by the fast shear-warp algorithm. This is because the fast shear-warp algorithm sets the value of all the transparent voxels which are not included in the compressed-opaque-voxel to 0. Also more aliasing is seen in the resulting images generated by the fast shear-warp algorithm, due to the resampling error. Fig. 4 shows the resulting images of six volume data sets generated by the shear-warp algorithm without compressed-opaque-voxel; and Fig. 5 shows the resulting images generated by the shear-warp algorithm with one run-length-encoded volume. Fig. 4 shows the photorealistic rendering style generated by using bilinear-interpolation; and Fig. 5 shows the non-photorealistic rendering style (stipple images) generated by using the mapping to the nearest pixel. In Fig. 5 random values with a quasi-random distribution are used, based on a sobol sequence for implementing automatic stippling algorithm.

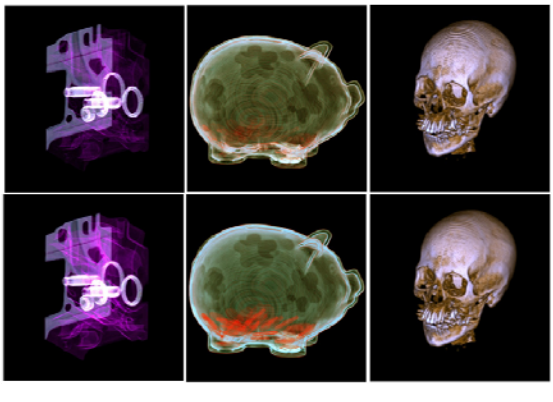


Fig. 3. Resulting images by fast shear-warp algorithm (top row) and shear-warp algorithm without compressed-opaque-voxel (bottom row). Left is engine, middle is piggy bank and right is CT head data.

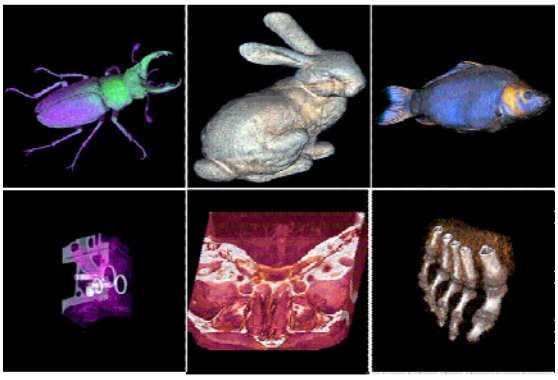


Fig. 4. Resulting images by shear-warp algorithm without compressed-opaque-voxel. Stag beetle, bunny, carp, foot, CTA brain and engine data from left top most to clock wise direction.

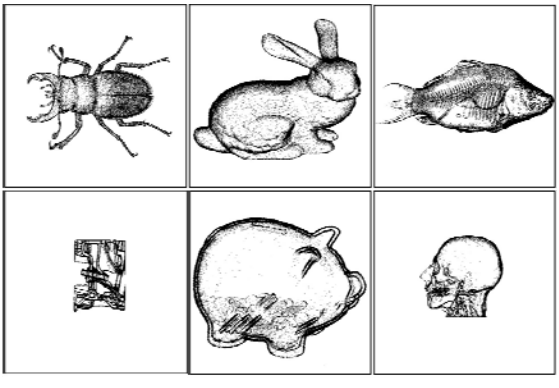


Fig. 5. Resulting images by shear-warp algorithm with one run-length-encoded volume. Stag beetle, bunny, carp, CT head, piggy bank and engine data from left top most to clock wise direction.

6 Conclusion

This paper has presented two kinds of novel optimization techniques for the shear-warp volume rendering algorithm on CPU: (1) the shear-warp algorithm without compressed-opaque-voxel, and (2) the shear-warp algorithm with one run-length-encoded volume. The former does not create compressed-opaque-voxel, making direct access to the memory space where initially loaded volume data are stored in order to enhance image quality by eliminating resampling error. The latter creates only one run-length-encoded volume by combining stippling techniques, which depicts image by a collection of points, with warp-transform in order to accelerate rendering speed without bilinear-interpolation and reduce aliasing in the final image. Also, we have shown that both optimization techniques efficiently decreased the memory consumption and preprocessing time for shear-warp algorithm, and produce the excellent results in rendering speed and image quality.

Acknowledgements. This research was supported by Samsung electronics and MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-C1090-1001-0008) and Future-based Technology Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0020-732).

References

1. Drebin, R., Carpenter, L., Hanrahan, P.: Volume Rendering. In: Computer Graphics 1988, vol. 22, pp. 65–74 (1981)
2. Levoy, M.: Display of surfaces from volume data. IEEE Computer Graphics & Applications 8(3), 29–37 (1988)
3. Lee, W.: Footprint evaluation for volume rendering. In: Proceedings of SIGGRAPH 1990 (1990); Computer Graphics 24(4), 367–376 (1990)
4. Zuiderveld, Karel, J., Koning, A.H.J., Viergever, M.A.: Acceleration of ray-casting using 3D distance Transforms. In: Proceedings of Visualization in Biomedical Computing 1992, Chapel Hill, North Carolina, pp. 324–335 (October 1992)
5. Lacroute, P., Levoy, M.: Fast volume rendering using a shear-warp factorization of the viewing transformation. In: Proceedings of SIGGRAPH 1994, pp. 451–458 (1994)
6. Udupa, J.K., Odhner, D.: Shell Rendering. IEEE Computer Graphics and Applications, 58–67 (1993)
7. Krueger, J., Westermann, R.: Acceleration techniques for GPU-based volume rendering. In: Proceedings IEEE Visualization 2003 (2003)
8. Hadwiger, M., Kniss, J.M., Rezk-salama, C., Weiskopf, D., Engel, K.: Real-time Volume Graphics. A. K. Peters, Ltd., Natick (2006)
9. Strothotte, T., Schlechtweg, S.: Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation. Morgan Kaufman, San Francisco (2002)

Author Index