Advances in Intelligent Systems and Computing 448

Shahram Latifi *Editor*

Information Technology: New Generations

13th International Conference on Information Technology

Springer

Advances in Intelligent Systems and Computing

Volume 448

Series editor

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

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

More information about this series at http://www.springer.com/series/11156

Shahram Latifi
Editor

Information Technology: New Generations

13th International Conference on Information Technology

Springer

Editor
Shahram Latifi
PHASE
Las Vegas, NV
USA

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-32466-1 ISBN 978-3-319-32467-8 (eBook)
DOI 10.1007/978-3-319-32467-8

Library of Congress Control Number: 2016935571

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Welcome to the 13th International Conference on Information Technology- New Generations- ITNG 2016. It is a pleasure to report that we have another successful year for the ITNG 2016. Gaining popularity and recognition in the IT community around the globe, the conference was able to attract many papers from authors worldwide. The papers were reviewed for their technical soundness, originality, clarity and relevance to the conference. The conference enjoyed expert opinion of 89 author and non-author scientists who participated in the review process. Each paper was reviewed by at least two independent reviewers. A total of 96 articles were accepted as regular papers and 14 were accepted as short papers (posters). The overall acceptance rate was 35 %.

The articles in this proceeding address the most recent advances in such areas as Wireless Communications and Networking, Software Engineering, Information Security, Data Mining, Informatics, Wavelets, High Performance Computing Architectures, Internet, and Computer Vision.

As customary, the conference features two keynote speakers on Monday and Tuesday. The presentations for Monday and Tuesday are organized in 2 meeting rooms simultaneously, covering a total of 16 technical sessions. Poster presentations are scheduled for the morning and afternoon of these days. The award ceremony, conference reception and dinner are scheduled for Tuesday evening. Two parallel sessions have been organized for Wednesday.

Many people contributed to the success of this year's conference by organizing symposia or technical tracks for the ITNG. Dr. Yenumula Reddy deserves much credit for spearheading the review process and running a symposium on Wireless Communications and Networking. My sincere thanks go to other symposium and major track organizers and associate editors namely- Drs. Kohei Arai, Alessio Bucaioni, Glauco Carneiro, Narayan Debnath, Luiz Alberto Vieira Dias, Ray Hashemi, Teruya Minamoto, Kashif Saleem, Basit Shahzad, Christoph Thuemmler, and Fangyan Shen.

Others who were responsible for solicitation, review and handling the papers submitted to their respective sessions include Drs Azita Bahrami, Doina Bein, Wolfgang Bein, Federico Ciccozzi, Saad Mubeen, Armin Schneider and Mei Yang.

The help and support of the Springer in preparing the ITNG proceedings is specially appreciated. Many thanks are due to Michael Luby, the Senior Editor and Nicole Lowary, the Assistant Editor of the Springer supervisor of Publications for the timely handling of our publication order. Finally, the great efforts of the conference secretary, Ms. Mary Roberts who dealt with the day to day conference affairs, including timely handling volumes of emails, are acknowledged.

The conference venue remains the same as last year- the Flamingo Hotel as the conference. The hotel, conveniently located in the heart of Las Vegas Strip, provides an easy access to other major resorts and recreational centers.

I hope and trust that you have an academically and socially fulfilling stay in Las Vegas.

April 2016

Shahram Latifi
The ITNG General Chair

ITNG 2016 Organization

Editor and General Chair

Shahram Latifi University of Nevada, Las Vegas, USA

Associate Editors

Kohei Arai	Saga University, Japan
Glauco Carneiro	University of Salvador, Brazil
Narayan Debnath	Winona State University, USA
Luiz Alberto Vieira Dias	Instituto Tecnológico de Aeronáutica, Brazil
Ray Hashemi	Armstrong State University, USA
Yenumula Reddy	Grambling State University, USA

Track Organizers

Kohei Arai	Saga University, Japan
Alessio Bucaioni	Mälardalen University, Sweden
Glauco Carneiro	University of Salvador, Brazil
Narayan Debnath	Winona State University, USA
Luiz Alberto Vieira Dias	ITA, Brazil
Ray Hashemi	Armstrong State University, USA
Teruya Minamoto	Saga University, Japan
Kashif Saleem	King Saud University, Saudi Arabia
Basit Shahzad	King Saud University, Saudi Arabia
Yenumula Reddy	Grambling State University, USA
Christoph Thuemmler	Edinburgh Napier University, UK
Fangyan Shen	New York City College of Technology, USA

Session Chairs

Azita Bahrami	IT Consultation, USA
Doina Bein	California State University, USA
Wolfgang Bein	University of Nevada, USA
Federico Ciccozzi	Mälardalen University, Sweden
Saad Mubeen	Mälardalen University, Saudi Arabia
Armin Schneider	Technical University Munich, Germany
Mei Yang	University of Nevada, Las Vegas, USA

Publicity Chair

Basit Shahzad King Saud University, Saudi Arabia

Conference Secretary

Mary Roberts PHASE Inc., USA

Reviewers

Abbas, Haider
Ahmed, Adel
Ahmed, Iqbal
Al-Muhtadi, Jalal
Arai, Kohei
Araujo, Marco Antonio
Ashino, Ryuichi
Autili, Marco
Awan, Fahim
Baguda, Yakubu
Barina, David
Behnam, Moris
Braga, Regina
Brito e Abreu, Fernando
Caetano, Paulo
Camargo, Valter
Campos, Fernanda
Carlson, Jan
Carneiro, Glauco
Carreiro, Bruno
Challah, Yacine
Chaudhry, Imran
Christos, Kalloniatis
Cicchetti, Antonio
Clincy, Victor
Colaco Jr., Methanias
Costa, Heitor
Crespo, Yania
Dai, Qing
David, Jose Maria
Delamaro, Marcio Eduardo
Derhab, Abdelouahid
Dias, Luiz
Eldefrawy, H.
Eramo, Romina
Fabbri, Sandra
Fernandes, Joao M.
Figueiredo, Eduardo
Frederic, Jean
Fujinoki, Kensuke
Garcia, Rogerio
Garcia-Valls, Marisol
Gawanmeh, Amjad
Hashemi, Ray
Huang, Yingsong
Ji, Yanqing
Kimura, Takuma
Kinoshita, Takehiko
Liebel, Grischa
Maia, Marcelo
Malavolta, Ivano
Marques, Johnny
Meghanathan, Natarajan
Mesit, Jaruwan
Mialaret, Lineu
Minamoto, Teruya
Mirza, Abdulrahman
Monteiro, Miguel
Montini, Denis
Morimoto, Akira
Munson, Ethan V
Novais, Renato

Oliveira, Toacy
Orgun, Mehmet
Pathan, Al-Sakib
Paulin, Alois
Pontin, Renata
Reddy, Yenumula
Rehmani, Mubashir
Rodrigues, Joel
Rossi, Gustavo
Saito, Norikazu
Saleem, Kashif
Saleem, Kashif
Sampaio, Paulo
Sarijari, Adib
Satyanarayana, Ashwin
Schneider, Armin
Shahzad, Basit
Shen, Fangyang
Soares, Michel
Solari, Martin
Sreekumari, Prasanthi
Suzana, Rita
Takano, Shigeru
Thuemmler, Christoph
Tolle, Herman
Tsuchiya, Takuya
Watanabe, Yoshitaka

Contents

Model-Driven Engineering for Cyber-Physical Systems

Data Mining

New Trends in Wavelet and Numerical Analysis

Computer Vision, HCI and Image Processing/Analysis

Part I
Networking and Wireless Communications

Understanding User's Acceptance of Personal Cloud Computing: Using the Technology Acceptance Model

Mohamed Eltayeb and Maurice Dawson

Abstract Personal Cloud Computing (PCC) is a rapidly growing technology, addressing the market demand of individual users for access to available and reliable resources. But like other new technologies, concerns and issues have surfaced with the adoption of PCC. Users deciding whether to adopt PCC may be concerned about the ease of use, usefulness, or security risks in the cloud. Negative attitudes toward using a technology have been found to negatively impact the success of that technology. The purpose of this study was to understand users' acceptance of PCC. The population sample consisted of individual users within the United States between 18 and 80 years of age. The theoretical framework utilized in this study was based on the technology acceptance model (TAM). A web survey was conducted to assess the measurement and understanding of patterns demonstrated by participants. Our results shows that in spite of the potential benefits of PCC, security and privacy risks are deterring many users from moving towards PCC.

Keywords Personal Cloud Computing · Cloud computing · Technology Acceptance Model

1 Introduction

Recently, cloud computing has been injected with new life from companies such as Google, Microsoft, IBM, Amazon, and others who now offer cloud computing

M. Eltayeb
Colorado Technical University, 4435 N. Chestnut St., Colorado Springs, CO 80907, USA
e-mail: Mohamed.Eltayeb@coloradotech.edu

M. Dawson(✉)
University of Missouri-St. Louis, 1 University Drive, St. Louis, MO 63121, USA
e-mail: Maurice.Dawson@umsl.edu
http://www.umsl.edu/

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_1

services for personal use. In this paper, we refer to PCC as the private and hybrid cloud computing for the individual user's usage. PCC has emerged as a new paradigm model for individual users. It has transformed the user's computer device from being device-centric to information-centric [1-7]. PCC enables individuals to share music, photos, videos, documents, and applications using any computer devices connected to the network. It provides portable access from anywhere in the world to information held centrally.

Though PCC demonstrated great potential in terms of scalability and agility, and is therefore enjoying great popularity and attention, storing data and applications in the cloud is becoming very risky [8]. Many cloud users have become concerned about security and protection of privacy [9], perhaps not surprisingly as the cloud functions to store and share private data [10]. The main challenge and disadvantage users see in the adoption of cloud computing is that they lack full control over the hardware [11]. There are, therefore, several issues related to privacy and security that must be addressed before PCC can be considered reliable.

Very often, data stored in the cloud is seen as valuable to those with malicious intent [12]. Storing data and applications in the cloud is becoming very risky. It is very important for users to take personal precautions in order to secure personal, sensitive information—information they would have stored in their local computer device (laptop, desktop, phone, tablet, etc.), and now store in the cloud. It is critical for users to understand the security measures that the cloud provider has in place [13].

Predicting technology usage and adoption has become a topic of mainstream study [14]. The rapidly growing importance of technologies has led researchers to study user technology acceptance intensively. Negative attitudes toward a technology, its ease of use, and its usefulness may negatively impact the success of that technology. Indeed, several IT scholars considered the user's acceptance of a technology to be the key success factor for the adoption of that technology [15].

The remaining of the paper is organized as follows: in the next section, we present the objectives of the study. Section 2 discusses and articulates our theoretical framework. Section 3 presents the research questions and hypothesis. The population sample information are discussed in section 4. A detailed discussion of the study's data collection and validation are provided in section 5. Section 6 discusses the data analysis of the study.

Section 5 presents methods used for data collections. Section 6 discusses the data analysis. Section 7 discusses our findings; and additional findings are discussed in section 8. Our recommendation and discussion are presented in section 9. Finally, section 10 concludes the paper.

2 Theoretical Framework

This research was deductive in nature. Therefore, a quantitative method was selected to understand user's acceptance of PCC. In particular, a non-experimental quantitative approach was used. The theoretical framework utilized in this study

was based on TAM—a widely recognized model in the field of IT, introduced by Davis in 1986 [16]. Davis [16, 17] recommended TAM as the best model for investigating end user attitudes towards and acceptance or rejection of a technology. According to Gao [18], "TAM can serve the purpose of predicting user acceptance of a technology before the users get heavily involved in the technology and thus is a cost-effective tool in screening potential candidate systems or programs" (p. 3). This theoretical framework was used in this study to examine factors influencing PCC adoption.

TAM proposes two independent variables—perceived ease of use, and perceived usefulness (PU) of a technology—to predict the user's attitude toward adopting a technology [18]. Davis [16, 17] defined PU as the degree to which a user believes that the adoption of a particular technology will improve the performance of his or her job. The perceived ease of use, on the other hand, is the degree to which a user believes that using a particular technology would be effortless [17].

Since the purpose of this study was to understand users' thoughts, experiences, expectations, and security concerns with respect to the adoption of PCC, TAM was found to be the most suitable theoretical approach due to its validity and reliability in exploring and explaining users' attitudes toward using technologies [16]. Cloud computing is a new approach to computing [19], and as new technologies are introduced and are adopted, concerns and challenges emerge [20]. Therefore, this study on PCC technology has a legitimate place in research on user technology acceptance.

This research study extended TAM to include attitude towards using (ATU) as an additional independent variable. This study adapted a survey instrument tested by Davis [16], and validated by Venkatesh [21]. Overall, this study included three independent variables and one dependent variable. The independent variables are perceived ease of use (PEOU), attitude toward using PCC (ATU) and perceived usefulness of PCC (PU). One dependent variable—Intention to Use PCC (ITC)—was used broadly in this study.

3 Research Questions and Hypothesis

3.1 Research Question 1: How Does Perceived Ease of Use Influence the Acceptance of PCC?

$H1_A$: Perceived ease of use positively influences attitude toward the acceptance of PCC.

$H1_0$: Perceived ease of use has no correlation to attitude toward the acceptance of PCC.

The independent variable PEOU determines the degree to which a user believes that using PCC would be effortless. PEOU is measured on a typical scale of seven ordered continuum of response categories: Strongly disagree, Disagree, Slightly disagree, Neutral, Slightly agree, Agree, Strongly agree.

3.2 Research Question 2: How Does Perceived Usefulness of PCC Influence the Acceptance of PCC?

H2A: Perceived usefulness positively influences attitude toward the acceptance of PCC.

H20: Perceived usefulness has no correlation to attitude toward the acceptance of PCC.

The independent variable PU determines the degree to which a user believes that the adoption of PCC would improve the performance of his or her job. PU is measured on a typical scale of seven ordered continuum of response categories: Strongly disagree, Disagree, Slightly disagree, Neutral, Slightly agree, Agree, Strongly agree.

Research Question 3: How does attitude toward using PCC influence acceptance of PCC?

$H3_A$: There will be a correlation between attitude toward using PCC and the user's acceptance of PCC.

$H3_0$: There will be no correlation between attitude toward using PCC and the user's acceptance of PCC.

The independent variable (ATU) determines the degree to which a user's acceptance of PCC relates to that user's attitude toward using PCC. ATU is measured across a typical scale of seven ordered continuum of response categories: Strongly disagree, Disagree, Slightly disagree, Neutral, Slightly agree, Agree, Strongly agree.

4 Population Sample

This study used a simple random sample from Survey Monkey engineering and IT professionals panels. The minimum representative sampling size (n) was estimated to be 384 participants for this study, based on a confidence interval of 5% and a confidence level of 95%. As many participants as possible were included to increase the accuracy of statistical sampling data.

A total of 464 participants were randomly chosen to participate in the study. However, only 399 of the chosen participants fully completed the survey; 20 participants opted out and 45 were disqualified either with unanswered questions or with providing partial data. Only fully completed responses were admitted. There was a 95.69% completion rate of surveys. Thus, the number of completed responses exceeded the estimated, representative minimal sample size of 384. Several items in the survey—gender, age range, education, state, and census region—were included to further describe demographic characteristics of the sample.

5 Data Collection and Validation

This study used a web survey to collect data from participants. The survey was hosted by SurveyMonkey.com, which is considered one of the leading providers of online survey solutions. The identity of participants was kept confidential and anonymous. Participant had the right to withdraw from participation at any time. Participants were required to complete a consent form as soon as they login to the SurveyMonkey.com. The consent form store collected information confidentially. Information collected by the survey included: gender, age range, state of residence, and level of education.

The survey questions in this study were designed to express a clearly negative or positive opinion and avoided neutral opinions to solicit more definitive responses. The researcher transcribed the data collected through Survey Monkey into digital documents. Data was downloaded by the researcher after the completion of data collection. Only fully completed surveys were employed in this study. Some participants were disqualified either with unanswered questions or with providing partial data. Incomplete responses and partial data were completely discarded. Additionally, Survey data was assessed for multicollinearity, homoscedasticity, linearity, and normality [22].

6 Data Analysis

The multiple regression test was performed as follows: (1) predictors were selected using their semi-partial correlation with the outcome; (2) predictors were entered into the model based on a mathematical method using Statistical Package for the Social Sciences (SPSS) Version 19.0; (3) the selection of predictors was performed in steps [23].

Descriptive statistics, ANOVA, Chi-square, and Multiple regressions were performed to explore the relationship between the continuous dependent variable (ITU) and independent variables (PEOU, PU, and ATU). Several descriptive statistics techniques were used to depict the sample population characteristics. ANOVA was used to analyze both dependent and independent variables. In addition, the multiple regression test was performed to assess the impact of each independent variable (PU, PEOU, and ATU) on the dependent variable (ITU).

7 Findings

TMultiple regressions were performed to explore the relationship between the continuous dependent variable (ITU) and independent variables (PEOU, PU, and ATU). This research study tested the relationship between perceived ease of use, usefulness, attitude toward using PCC, and the acceptance of PCC. The model's adequacy was tested in terms of normality, linearity, and collinearity [24].

7.1 First Regression Model

The first regression was performed to determine the relationship between the independent variable PEOU and the dependent variable ITU. Six predictors of the independent variable PEOU were tested to address the first research question. The relationship between the predictors and the outcome were measured using a stepwise method, so that the unique contribution of each predictor could be assessed to explain variance of the outcome [24]. The following research question and hypothesis were tested at this stage:

How does perceived ease of use influence the acceptance of PCC?

$H1_0$: Perceived ease of use has no correlation to the acceptance of PCC.

The results of the regression analysis demonstrated a positive correlation between PEOU and the acceptance of PCC. It was found that approximately 23.5% of the variance in the acceptance of PCC can be accounted for by PEOU. Therefore, the null hypothesis was rejected and the alternative was supported.

This result is consistent with other studies showing that there is a positive correlation between PEOU and the user's acceptance of a technology. Davis [16] argued that perceived ease of use positively influence attitude toward using a technology. Perceived ease of use plays an important role in determining and predicting the user's intention to use a technology [25]. Ramgovind [26] stated "The success of modern day technologies highly depends on its effectiveness of the world's norms, its ease of use by end users and most importantly its degree of information security and control" (p. 1).

7.2 Second Regression Model

The second regression was performed to determine the relationship between the independent variable PU and the dependent variable ITU. Five predictors of the independent variable PU were tested to address the first research question. A stepwise regression was performed to build the model. The unique contribution of each predictor was measured to explain the variance of the outcome [24]. The following research question and hypothesis were tested at this stage:

How does perceived usefulness of PCC influence the acceptance of PCC?

H20: Perceived usefulness has no correlation to attitude toward the acceptance of PCC.

The results of the regression analysis demonstrated a positive correlation between PU and the acceptance of PCC. It was found that approximately 78.2% of the variance in the acceptance of PCC can be accounted for by PU. Therefore, the null hypothesis was rejected and the alternative was supported. This result is consistent with other studies showing that there is a positive correlation between PU and the user's acceptance of a technology.

7.3 Third Regressions Model

The third regression was performed to determine the relationship between the independent variable ATU and the dependent variable ITU. Four predictors of the independent variable ATU were tested to address the first research question. The relationship between the predictors and the outcome were measured using stepwise regression. The unique contribution of each predictor was measured to explain the variance of the [24]. The unique contribution of each predictor is measured to explain the variance of the outcome. The following research question and hypothesis were tested at this stage:

How does attitude toward using PCC influence user's acceptance of PCC?

$H3_0$: There will be no correlation between attitude toward using PCC and user's acceptance of PCC.

The independent variable ATU was used for determining the degree in which there is a correlation between attitude toward using PCC and users' intention to use. The results of the regression analysis demonstrated a significant correlation between ATU and the acceptance of PCC. It was found that approximately 38.2% of the variance in the acceptance of PCC can be accounted for by ATU. Therefore, the null hypothesis was rejected and the alternative was supported.

8 Additional Findings

Statistically significant differences were found between age group means, as indicated by results of the one-way ANOVA (F (5, 2.801)=2.801, p=.017). The output result of the one-way ANOVA showed that the age group significantly contributed to the variation in the respondent's acceptance of the PCC technology. Because the output result of the one-way ANOVA test demonstrated significant differences between the age groups, a Post Hoc test was conducted to further investigate which means were significantly different from each other. The Post Hoc test result

indicated that the score of the age group 40-49 was much higher than the other age groups. Also, age groups 21-29 and 30-39 appeared to accept PCC more easily compared with age groups 18-20, 50-59, and 60-80.

A chi-square test of independence was performed to examine the relation between independent variables (PU, PEOU, and ATU) and the acceptance of PCC. A significant relationship was found between the independent variables and the frequency of acceptance of PCC at p-value <0.01, and degrees of freedom of 36. Additionally, a chi-square test of independence was performed to examine the relation between age group and acceptance of PCC. No significant relationship was found between age group and the frequency of acceptance of PCC, X2(10, N=399) =17.24, p-value=.069. The acceptance of PCC among the age group was as follows: approximately 66.6% of the age group "18-20;" 75% of the age group of "21-29;" 73% of the age group "30-39;" 80.5% of the age group "40-49;" 69.9% of the age group "50-59;" and 67.3% of the age group "60-80." Overwhelming majority of users that accepted PCC were in the "40-49" age group. On the other hand, the majority of users that did not have intention to use PCC were in the age group "60-80."

9 Discussion and Recommendation

Given the results of the data analysis in this study, it is arguable that PCC is here to stay. However, there are several issues related to security and privacy that must be tackled to increase user's acceptance of PCC. Hardly a day goes by without hearing news about privacy and security concerns of users. The findings of this study demonstrated that users are concerned about the way security and privacy is handled in the cloud. Protecting users' privacy in the cloud is big business for cloud providers. The success of a technology depends greatly on its ease of use but even more crucially its degree of security [26]. According to Koved [27], "When end-users' perceptions of risk are not aligned with those on which the system is based, there is a mismatch in perceived benefit, leading to poor user acceptance of the technology" (p. 1). Thus, users' understanding of privacy and security in the cloud rapidly becomes obsolete as the PCC technology progresses. Privacy and security must be evaluated continuously to achieve information privacy and security objectives and, thereby, user acceptance.

Prior research has shown that perceived privacy and security risks in the cloud negatively influence attitude toward using cloud computing. There are always security risks when moving towards cloud computing [28]. Though computer users would like to use PCC, they are concerned with utilizing a system they do not control. Data stored in the cloud is often seen as valuable to those with malicious intent [12]. It is very important for users to take precautions in order to secure personal, sensitive information—information they would have stored in their local computer device and now store in the cloud. It is critical for users to understand the security measures the cloud provider has in place before moving toward PCC [13].

10 Conclusion

The study's objective was to understand users' acceptance of PCC and a web survey was conducted to assess it. Several key findings emerged from this research study. The results of the data analysis showed that the majority of respondents had a positive view about PCC, but the participants seemed particularly concerned about security, privacy, and data theft in the cloud. User's attitude towards PCC appeared to be the most critical factor among the three suggested determinants of PCC acceptance in this study.

References

1. Ularu, E.G., Puican, F.C., Suciu, G., Vulpe, A., Todoran, G.: Mobile computing and cloud maturity Introducing machine learning for ERP configuration automation. Informatica Economica **17**(1), 40–52 (2013)
2. Onyegbula, F., Dawson, M., Stevens, J.: Understanding the need and importance of the cloud computing environment within the National Institute of Food and Agriculture, an agency of the United States Department of Agriculture. Journal of Information Systems Technology & Planning **4**(8), 17–42 (2011)
3. Dhar, S.: From outsourcing to cloud computing: Evolution of IT services. Management Research Review **35**(8), 664–675 (2012)
4. Okezie, C.C., Chidiebele, U.C., Kennedy, O.C.: Cloud computing: A cost effective approach to enterprise web application implementation (A case for cloud ERP web model). Academic Research International **3**(1), 432–443 (2012)
5. Reavis, D.: Information evaporation: The migration of information to cloud computing platforms. International Journal of Management & Information Systems (Online) **16**(4), 291 (2012)
6. Lin, A., Chen, N.C.: Cloud computing as an innovation: Perception, attitude, and adoption. International Journal of Information Management **32**(6), 533–540 (2012)
7. Tian, Y., Song, B., Huh, E.N.: Towards the development of PCC for mobile thin-clients. In: 2011 International Conference on Information Science and Applications (ICISA). IEEE (2011). doi:10.1109/ICISA.2011.5772368
8. Dutta, A., Peng, G.C.A., Choudhary, A.: Risks in enterprise cloud computing: The perspective of its experts. The Journal of Computer Information Systems **53**(4), 39–48 (2013)
9. Han, Y., Sun, J., Wang, G., Li, H.: A cloud-based BPM architecture with user-end distribution of non-compute-intensive activities and sensitive data. Journal of Computer Science and Technology **25**(6), 1157–1167 (2010)
10. Wang, H.: Privacy-preserving data sharing in cloud computing. Journal of Computer Science and Technology **25**(3), 401–414 (2010)
11. Katzan Jr., H.: On the privacy of cloud computing. International Journal of Management and Information Systems **14**(2), 1–12 (2010)
12. Chakraborty, R., Ramireddy, S., Raghu, T.S., Rao, H.R.: The information assurance practices of cloud computing vendors. IT Professional Magazine **12**(4), 29–37 (2010)
13. Alzain, M.A., Soh, B., Pardede, E.: A new model to ensure security in cloud computing services. Journal of Service Science Research **4**(1), 49–70 (2012)

14. Selamat, Z., Jaffar, N.: IT acceptance: From perspective of Malaysian bankers. International Journal of Business and Management **6**(1), 207–217 (2011)
15. Cocosila, M.: Role of user a priori attitude in the acceptance of mobile health: An empirical investigation. Electronic Markets **23**(1), 15–27 (2013)
16. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of IT. MIS Quarterly **13**(3), 318–342 (1989)
17. Davis, F.D.: User acceptance of IT: System characteristics, user perceptions and behavioral impacts. International Journal of Man-Machine Studies **38**(3), 475–487 (1993)
18. Gao, Y.: Applying the technology acceptance model (TAM) to educational hypermedia: A field study. Journal of Educational Multimedia and Hypermedia **14**(3), 237–247 (2005)
19. Khan, S., Khan, S., Galibeen, S.: Cloud computing an emerging technology: Changing ways of libraries collaboration. International Research: Journal of Library and Information Science **1**(2), 151–159 (2011)
20. Paquette, S., Jaeger, P., Wilson, S.: Identifying the security risks associated with governmental use of cloud computing. Government Information Quarterly **27**, 245–253 (2010)
21. Venkatesh, V.P.: An assessment of security vulnerabilities comprehension of cloud computing environments: a quantitative study using the unified theory of acceptance and use. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 3564324, Capella University) (2013)
22. Vogt, P.: Quantitative research methods for professionals. Pearson Education, Boston (2007)
23. Field, A.: Discovering statistics using SPSS. Sage, Thousand Oaks (2009)
24. Fabozzi, F.J., Focardi, S.M., Rachev, S.T., Arshanapalli, B.G.: Building and testing a multiple linear regression model. In: The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications, pp. 81–102. John Wiley & Sons, Hoboken (2014)
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 & management **40**(3), 221–232 (2003)
26. Ramgovind, S., Eloff, M. M., Smith, E.: The management of security in cloud computing. In: Information Security for South Africa (ISSA), pp. 1–7. IEEE (2010)
27. Koved, L., Trewin, S., Swart, C., Singh, K., Cheng, P.C., Chari, S.: Perceived security risks in mobile interaction. In: Symposium on Usable Privacy and Security (SOUPS), July 2013. http://cups.cs.cmu.edu/soups/2013/risk/Koved-RP-IT-2013.pdf
28. Aleem, A., Christopher, R.S.: Let me in the cloud: Analysis of the benefit and risk assessment of cloud platform. Journal of Financial Crime **20**(1), 6–24 (2013)

Cognitive Spectrum Decision via Machine Learning in CRN

A.M. Koushik, Fei Hu, Ji Qi and Sunil Kumar

Abstract In this research, we propose cognitive spectrum decision model comprised of spectrum adaptation (via Raptor codes) and spectrum handoff (via transfer learning) in Cognitive Radio Networks(CRN), in order to enhance the spectrum efficiency in multimedia communications. Raptor code enables the Secondary User (SU) to adapt to the dynamic channel conditions and maintain the Quality of Service (QoS) by prioritizing the data packets and learning the distribution of symbols transmission strategy called *decoding-CDF* through the history of symbol transmissions. Our scheme optimizes the acknowledgement (ACK) reception strategy in multimedia communications, and eventually increases the spectrum decision accuracy and allows the SUs to adapt to the channel variations. Moreover, to enhance spectrum decision in a long term process, we use Transfer Actor Critic Learning (TACT) model to allow the newly joined SU in a network to learn the spectrum decision strategies from historical spectrum decisions of the existing 'expert' SUs. Experimental results show that our proposed model works better than the myopic spectrum decision which chooses the spectrum decision actions based on just short-term maximum immediate reward.

Keywords Cognitive Radio Network (CRN) · Cognitive Spectrum Decision (CSD) · Raptor codes · Machine learning · Transfer Actor-Critic Learning (TACT)

1 Introduction

As per the Federal Communication Commission (FCC) the current frequency spectrum is under-utilized [1]. To utilize the vacant spectrum called as *Back-hauls*, the

A.M. Koushik(✉) · F. Hu · J. Qi
Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA
e-mail: koushik@crimson.ua.edu

S. Kumar
Electrical and Computer Engineering, SDSU, San Diego, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_2

concept of Cognitive Radio Networks (CRNs) has been brought up in which the unlicensed Secondary Users (SUs) use the idle spectrum when the licensed Primary Users (PUs) are not using it. But due to the random arrival of PUs, the SUs cannot utilize the idle spectrum efficiently.

In this research we target spectrum decision (SD) problem, which is to manage the SU transmission behaviors in order to adopt to the dynamic CRN spectrum conditions. Especially we focus on two sub-issues in SD (see Fig. 1): The first one is spectrum adaptation, that is, how a SU adjusts its transmission behaviors (such as packet sending pace) under highly dynamic channel conditions. The second one is spectrum handoff, which requires a SU switches to a new channel if the PU comes back.

On the issue of spectrum adaptation, it is difficult to adjust the sending rates based on the immediate observation of the channel conditions that could change so quickly. To overcome such an issue, we will use raptor codes [2] to control the sending pace. The raptor code is a rateless and regret free, i.e. the lost packet can be recovered using the earlier symbols received, and it makes use of all the symbols received for packet recovery (i.e., no symbols are wasted). To further improve spectrum efficiency, we employ *decoding-CDF* called as ratemore protocol [3] to transmit the sufficient amount of symbols in a dynamic and time varying channel. Without decoding CDF, the sender may send many redundant symbols when the channel condition is good which consumes more bandwidth unnecessarily or the symbols sent may not be enough to decode the packet successfully in a poor channel condition.

For spectrum handoff, we propose to use machine learning algorithms to manage the channel switching under dynamic channel conditions. Especially we adopt node-to-node transfer learning for spectrum handoff control, (i.e., a node learns from another node on how to switch to a new channel), instead of using self-learning, (i.e., a node learns the channel switching control by itself). A new joined SU can learn from a neighboring SU (called "expert") that has similar radio conditions as itself. A benefit of using node-to-node knowledge transfer is that we can reduce the time taken to converge to optimal value in a global optimization algorithm. Particularly, we employ Transfer Actor Critic Learning (TACT) [4] to allow a "student" SU to make correct spectrum decision by learning from an "expert" SU, which shares the optimal policy acquired from its historical spectrum decision with the learning SU. One more advantage of TACT is that, as the learning progresses, the student SU can start adapting to the CRN environment by its own. This is beneficial since each SU may experience different, dynamic channel conditions due to the independent and random mobility of each SU. Figure 1 shows our cognitive spectrum decision model.

The rest of this paper is organized as follows: Section 2 briefly summarizes the related work. Section 3 then details our CDF and TACT based spectrum decision model. The simulation results are given in Section 4, and Section 5 concludes this paper.

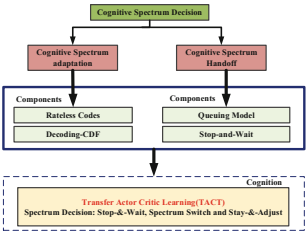


Fig. 1 The concept of Cognitive Spectrum Decision

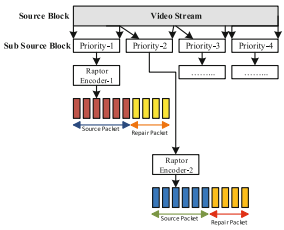


Fig. 2 Prioritized Raptor codes.

2 Related Work

Teaching Based Spectrum Decision Enhancement: Teaching based machine learning algorithms for wireless protocol design have been proposed in Apprenticeship learning (AL) [5] and Docitive learning models [6]. Docitive learning [6] have been used for interference management in femtocells. But those models cannot provide the explicit channel selection parameters. They also do not describe how to search an expert node. Besides, in AL model [5], a student node learns from the expert node all the time, and the student node should encounter exactly the same channel conditions as the expert node does during the entire learning process. In our TACT-based approach, a SU learns from expert at the initial stage of transmission and later learns on its own.

Rateless Codes with Decoding CDF for Enhanced Wireless Transmission: Rateless codes enable to decode the packets with low Packet error rate in wireless transmissions. Rateless codes in CRN have been proposed in [7] in which SUs act as relay nodes to forward PUs data packets. Whereas in our approach we employ rateless codes with decoding CDF to optimize symbol transmission strategy adjust to the time varying channel condition adaptively.

3 Cognitive Spectrum Decision

3.1 Spectrum Adaptation Based on CDF-Enhanced Raptor Codes

Each SU selects a channel for communication by considering various parameters such as channel holding time, channel sensing accuracy, Packet dropping rate, etc. The reason of considering all those parameters is due to the dynamic nature of the CRN channels that exhibit time-varying link quality. Poor link quality can induce high packet loss rate. To reduce the packet loss rate the sender needs to adjust the packet

transmission rate based on channel conditions. As an example, in 802.11 links, the sender determines the sending rate by measuring link signal-to-noise (SNR) ratio and uses it to estimate the constellation points associated with the respective modulation modes. It is hard to achieve smooth rate adaption since we typically have only a few pre-determined rates available in the system. Therefore, it is challenging to design a dynamic spectrum adaptation scheme in CRNs since channel conditions vary so quickly, even within the very short individual packet transmission time.

Rateless codes are also called as *regret-free* codes, which treat all transmitted symbols equally and no symbol is discarded since any symbol can be used to decode the later received symbols. Rateless codes have shown promising performance in multimedia over CRNs. In addition, in the sender side, each packet is disintegrated into symbols with added small redundancy. This enables the receiver to decode the symbols successfully as long as it has enough symbols to decode. Moreover, the sender does not need to make changes in its modulation and encoding schemes. In other words, it is "rateless" since the sending rates do not need to be micro-adjusted based on the channel conditions. The sender simply keeps sending symbols until the receiver is able to decode all the packets, and then the sender sends the next window of packets. For a well-designed rateless codes, the number of symbols transmitted closely tracks the variation in channel conditions.

Raptor codes assign different priorities to different packets, and the packets with higher priority are given more redundancy of symbols. From Fig. 2 we can see that, initially the packet is decomposed into pieces (group of symbols). Those pieces first pass through outer encoder called as LDPC codes, and they then pass through inner encoder called as LT codes. Hence, encoding can be characterized by $(K, C, \theta(x))$, here K means the number of message blocks (pieces), C is the outer code result (with block size L). Thus we have L intermediate symbols after passing outer encoder. The last $L - K$ symbols are redundant symbols. $\theta(x)$ is the degree distribution of LT codes. The L intermediate symbols are encoded with LT code to generate N encoded symbols. In total N symbols are transmitted over the lossy wireless channel. Even some packets are received with errors, it can be successfully decoded by using the previously received symbols. Let N_r be the number of received encoded symbols. The decoding failure probability, $P_e(\xi_r)$, is very low. Here $\xi_r = Nr - K$ is the encoding overhead (i.e., redundancy level) of Raptor codes. Then we have:

$$Pe(\xi_r) = 0.85 \times 0.567^{\xi_r} \tag{1}$$

The average communication overhead, induced due to the extra added symbols among the source symbols is:

$$\rho = \frac{1}{K} \sum_{i=0}^{\infty} (i.(P_e(i-1) - P_e(i))) \approx \frac{2}{K} \tag{2}$$

According to (2), we can see that we just need to transmit approximately 2 extra symbols to decode the transmitted packet successfully because extra added symbols (in average) should be: $K \times \rho = 2$. In addition, we can generate more symbols for high priority packets so that the receiver can have ample amount of symbols to decode the high priority packets successfully. Let L_1 denote the highest priority, L_2 the second highest, and so on. And we have K_i source symbols with priority L_i . Also $\xi_r(Ki)$ denotes the number of extra symbols for the priority L_i . Then the minimum coding overhead induced is $\rho(K_i)$, and the percentage of additional symbols among the total source symbols for priority L_i, should be:

$$\rho(K_i) = \frac{K_i \times PER + \xi_r(K_i)}{(1 - PER) \times K_i} \tag{3}$$

Here PER is the packer error rate. Conventional raptor codes treat all packets equally. Hence, our Unequal Error Protection (UEP) based Raptor codes can adaptively adjust the overhead of raptor codes based on PER for data with different priority levels.

Decoding CDF for Enhanced Raptor Codes: Achieving higher throughput in wireless networks is an important goal. In rateless codes, without careful scheduling and estimation the sender will send the data, meaning that in ratless codes the amount of symbols sent may be too redundant when the channel quality is good or the sent symbols may not be sufficient to decode the packet successfully in a lossy channel. Sending too many redundant symbols consumes bandwidth unnecessarily and the bandwidth cannot be used efficiently. On the otherhand, when packets are not being decoded successfully the QoE detoriates. Hence, both will not help to attain the required QoS and QoE. Hence a link layer protocol called *Ratemore*, is used to obtain the cumulative distribution for the probability of successful decoding of a sent packet. Using such a distribution (called as decoding CDF) the sender can estimate the number of symbols to be in the present channel condition. Using Ratemore protocol [3] we can design a proper spectrum adaptation strategy to estimate how many symbols can be transmitted before pausing for the feedback. In addition, the proper pausing intervals can also be determined to maintains the good QoS without introducing much communication overhead.

$$F(x) = \int_0^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{4}$$

Where, x, μ and σ are the Number of samples(NS) , mean and variance estimated out of the symbols transmission history respectively.

The decoding CDF [3] can be learnt online using the history of transmitted symbols. The Gaussian approximation model can be used for CDF learning. We just need to estimate the parameters including Gaussian mean, μ, and variance, σ^2. Both can be estimated in exponentially weighted accumulator fashion. Equation (4) shows CDF learning principle by using Gaussian approximation with learning rate $\alpha \in [0, 1]$.

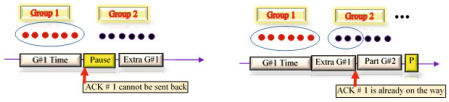


Fig. 3 (left) No CDF ; (right) With CDF.

Figure 3 shows the benefit of using CDF. The receiver cannot send the ACK unless there are ample amount of symbols to decode the G#1 packets successfully (see left figure); but when it has sufficient symbols to decode the packet (see right figure), the receiver sends the feedback. By using those received symbols, the receiver can draw a distribution of packet decoding.

3.2 Spectrum Handoff Control Based on TACT Model

The spectrum decision can be modeled as a Markov Decision process (MDP). It can be represented by a tuple as (S, A, T, R), where S depicts the set of system states, A is the set of system actions at each state. T represents the transition probability, where $T = P(s, a, s')$ is the probability of transition from state s *to* $s\prime$ when the action a is taken, and $R : S \times A \mapsto R$ is the reward or cost function, which depicts the reward for taking an action $a \in A$ in state $s \in S$. In MDP we intend to find the optimal policy $\pi^*(s) \in A$, i.e. a series of actions $\{a_1, a_2, a_3, ...\}$ for state s, in order to maximize the total discount reward function.

States, S : For SU-i, the network state at $(j+1)th$ channel assignment stage is $s_{ij} = \{\chi_{ij}^{(k)}, \xi_{ij}^{(k)}, \rho_{ij}^{(k)}, \phi_{ij}^{(k)}\}$. Where k is the channel being used. $\chi_{ij}^{(k)}$ is the idle or busy status of the channel. $\xi_{ij}^{(k)}$ is the channel quality determined using Packet Error Rate(PER). $\rho_{ij}^{(k)}$ is the channel traffic load condition determined using non-preemptive M/G/1 queueing model[5][4] and $\phi_{ij}^{(k)}$ denotes priority.

Actions, A : Three actions are considered for iSM scheme: 1. stay-and-await:stay in the same channel and hold the traffic until the channel condition is above a pre-set threshold; 2. stay-and-adjust (transmit more or less symbols) until the stable reward,Mean Opinion Score(MOS) value is met using decoding-CDF , and 3. spectrum handoff: switch to a new channel; We denote $a_{ij} = \{\beta_{ij}^{(k)}\} \in A$ as the candidate of actions for SU_i on state s_{ij} after the assignment of $(j+1)^{th}$ channel. $\beta_{ij}^{(k)}$ represents the probability of choosing action a_{ij}. A particular action is selected based on softmax policy as below,

$$\pi^{(k)}(s^{(k)}, a) = \frac{exp(\frac{Q(s,a)}{\tau})}{\sum_{a\prime \in A} exp(\frac{Q(s,a')}{\tau})} \tag{5}$$

Where τ is called temperature. The high temperature indicates the exploration of the unknown state-action pairs.

Reward : We adopt Mean opinion score (MOS) as the reward function using the following equation:

$$MOS = \frac{a_1 + a_2 FR + a_3 ln(SBR)}{1 + a_4 PER + a5(PER)^2} \tag{6}$$

Where FR is frame rate, SBR is Sender Bit Rate, and TPER is Total Packet Error Rate. Here $a_1, a_2, a_3, a_4, and\ a_5$ are coefficients estimated using regression analysis.

Self-Learning via Q-Learning: When the new SU (denoted as SU-i) joins the network and if there is no expert SU that has similar QoS requirement to itself, then it can learn spectrum decision actions by itself via Q-learning algorithm, as shown in Fig. 4. Q-learning aims to find the optimal action to maximize the MOS at the current policy $\pi^*(s_{i,j}, a_{i,j})$ in the process of $(j+1)th$ channel assignment to SU-i. The fairness of the action, a_i, taken in state s_i given policy π, can be found by action-value function, $Q^{\pi}(s, a)$, which is given by

$$Q^*(s,a) = E(R_{i,j+1}) + \gamma \sum_{s\prime} P_{s,s\prime}(a) \max_{a\prime \in A} Q^*(s,a), \quad \gamma \in (0,1) \tag{7}$$

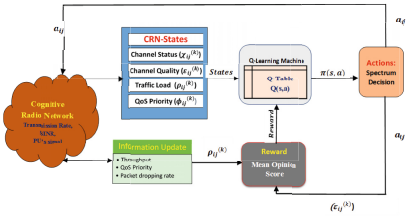


Fig. 4 Q-learning based CSH

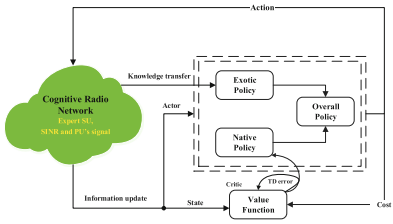


Fig. 5 TACT based CSH.

TACT-Based Spectrum Decision: Q-learning could take a long time to converge to an optimal solution due to the difficulty of selecting suitable state and initial parameters in the Markov model. To increase the efficiency of spectrum decision, the learning process has to be made fast. We thus propose to use Transfer Actor Critic Learning (TACT) to enhance the learning process. In TACT, the newly joined SU finds an expert SU which has similar QoS requirement as itself. For example, a SU with video transmission could find an expert SU with video applications instead of a SU with data transmission tasks. The expert SU exchanges the optimal policy learnt during its previous spectrum decisions with the student SU. Moreover, if the SU cannot learn from the expert anymore due to radio condition mismatch, the SU can learn from its own historical spectrum decision records.

TACT comprises of three components: *actor, critic, and environment*. For a given state, the actor selects and executes an action using softmax policy. This causes a

transition from one state to another with certain reward that is fed back to actor. Then the critic calculates the $time\ difference$ (TD error) to evaluate the action taken and updates the value function. After receiving the feedback from the critic, the actor updates the policy. Figure 5 depicts the TACT principle.

1. Transferring the Knowledge: Both the policy and value function are updated separately. This makes the transfer of policy knowledge easier than Q-learning model.
i. Action Selection: Initially at state s_{ij} the node is using channel k, and $SU - i$ chooses an action using eqn(6) to find an optimal policy in a current channel.
ii. TD error calculation and State-value function update: TD time difference can be calculated as,

$$\delta(s,a) = R_{s,a} + \gamma \sum_{s' \in S} P(s'|s,a)V(s') - V(s) = R_{s,a} + \gamma V(s') - V(s)$$

Where, R(.) is the reward, and V(.) is the state value function. Subsequently, the state-value function can be updated as

$$V(s') = V(s) + \alpha(\nu_1(s,m))\delta(s,a) \tag{8}$$

Where $\nu_1(s,m)$ indicates the occurrence time of the state s in these m stages. $\alpha(.)$ is a positive step-size parameter that affects the convergence rate.
iii. Policy Update: The policy is updated using the feedback from the critic as follows,

$$p(s,a) = p(s,a) - \beta(\nu_2(s,a,m))\delta(s,a) \tag{9}$$

Where $\nu_2(s,a,m)$ denotes the occurrence time of action, a at state s in these m stages. $\beta(.)$ denotes the positive step size parameter.

2. Overall Policy Update: The overall spectrum decision policy compares the policy of the expert with that of the student SU, as shown in (10). The overall policy is determined by the combination of the native policy p_n and an exotic policy (expert policy) p_e . Assume at channel k, the state is $s^{(k)}$ and the chosen action is $a^{(k)}$. Accordingly, the overall policy can be updated as [4]

$$p_o^{(k+1)}(s,a) = [(1 - \omega(\nu_2(s,a)) p_n^{(k+1)}(s,a) + \omega(\nu_2(s,a)) p_e^{(k+1)}(s,a)]_{-p_t}^{p_t} \tag{10}$$

Where $[x]_m^n$ with $n > m$, indicates the Euclidean distance of interval $[m,n]$, i.e. $[x]_m^n = m \text{if } x < m$; $[x]_m^n = n$ if $x < n$; and $[x]_m^n = x$ if $m \leq x \leq n$. In this scenario, $m = -p_t$ and $n = p_t$. $p_0^{(m+1)}(s,a) = p_0^{(m)}(s,a)$, $\forall a \in A$ but $a \neq a_{ij}$. Apart from that, $p_n(s,a)$ also updates itself according to (10). Interestingly, at the beginning of the process, the exotic policy $p_e(s,a)$ dominates the overall learning policy, i.e., the learning SU-i takes the action which is optimal for expert SU; and as the learning progresses, the effect of exotic policy starts decreasing with parameter $\omega \in (0,1)$. The transfer rate follows $\omega \mapsto 0$ as the number of iterations goes to ∞. This makes

student SU-i get rid of negative guidelines from expert SU and can learn by itself according to new channel conditions.

4 Simulation Results

In our simulation, we assume that the selected channel has high idle duration, channel condition is generally good (i.e., low packet error rate),and the network congestion(Packet dropping rate) is not serious. In addition, we assume the video transmission has highest priority and its waiting time in queue is very low.

4.1 Spectrum Adaptation Based on Decoding CDF

In this section, we evaluate the decoding CDF distribution and learning performance after utilizing raptor codes. Figure 6 shows the decoding CDF learning result using Algorithm 1 with learning rate specified in Table 1 for different SRN values. i.e., from -5 dB to 25 dB. The graph clearly depicts that, with high SNR (say 25 dB), less symbols (100 from Fig. 6) have to be transmitted in order to decode the packet successfully; whereas at low SNR (-5 dB) large number of symbols (>2300) decode the transmitted packet successfully.

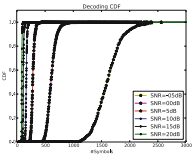


Fig. 6 decoding CDF

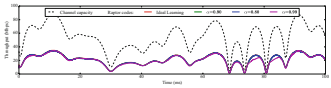


Fig. 7 Channel throughput.

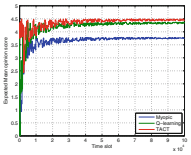


Fig. 8 Learning Performance.

In Fig. 7, we estimate the throughput that can be achieved using the decoding CDF strategy and pausing intervals in raptor codes over a Rayleigh fading channel in the duration of 100ms. In addition, it is assumed that the SU is moving at a speed of 10m/s with average link SNR=15dB. The raptor code has the throughput almost half the Shannon capacity due to the fact that it needs to transmit some *empty* packets (no payload) in order to meet the latency requirements.

4.2 TACT Enhanced Spectrum Decision

In this section we examine the performance of our TACT based spectrum decision with the Q-learning and Myopic spectrum decisions. We assume that the peak sending

rate of each SU is 3Mbps. Assume that all SUs adopt rateless codes, and the expert SU teaches a learning SU about its spectrum decision policy as determined in section III-B. We examine 3 cases. 1). Fast moving SU, 2). Learning without decoding CDF, and 3) learning with decoding CDF. Myopic spectrum decision considers only the immediate maximum reward without considering the long-term effect. Whereas Q-learning scheme is a self-learning scheme which learns about the spectrum decision from the scratch. It eventually takes longer time to converge than TACT-based scheme. TACT learns from the expert SU to reduce the learning time and to get adapted to the dynamic channel conditions quickly in order to enhance the spectrum efficiency. We consider MOS as the reward since at the receiver it is hard to measure the video PSNR (peak SNR). Our scheme has been verified in Fig. 8, in which the reward, MOS (from (8)) achieved in TACT is high compared to Q-learning and myopic. TACT learns from the expert SUs policy. Thus it takes less time to achieve optimality. Q-learning takes the action with the best performance in the future but takes initial time to converge, whereas myopic scheme takes the action only considering the immediate reward.

Figure 9 depicts the learning performance when the SU is moving fast and does not use Ratemore protocol. The SU experiences variation in channel conditions frequently due to channel fading, signal attenuation and small coherence time, etc. Those factors eventually affect the reward value as shown in Fig. 9. If the Ratemore protocol (decoding CDF) is not employed, the SU cannot recover from the low reward and it continues with the same lower value. If the decoding CDF is employed, the SU can pause properly for the feedback from the receiver, and it eventually learns and executes the strategy, i.e., transmitting enough symbols for the receiver to decode the packet successfully. This conclusion is shown in the Fig. 10 in which the SU is able to recover its optimal value with the aid of the strategy learnt from decoding CDF.

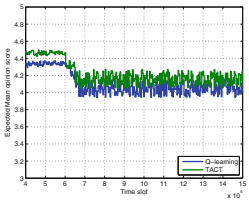


Fig. 9 SU Without decoding CDF

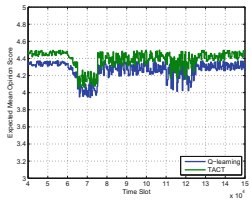


Fig. 10 SU with decoding CDF.

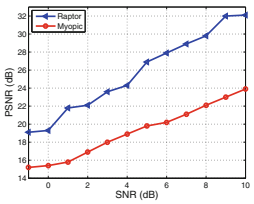


Fig. 11 PSNR v/s SNR.

Figure 11 shows the variation of PSNR with respect to different SNRs for the video frames shown in Fig. 12. Compared to general myopic scheme (without using learning model), our method improves the PSNR all the times. Figure 12 shows the video resolution comparisons between our scheme and myopic spectrum decision. We can see our scheme outperforms the myopic one.

Fig. 12 Video effect comparisons of learning-based and myopic spectrum decision schemes.

5 Conclusions

In this paper we have demonstrated our cognitive spectrum decision scheme using rateless codes, decoding CDF, and machine learning algorithm called TACT. It is observed that the rateless codes along with decoding CDF can maintain the QoS over dynamic channel conditions and optimize the feedback and symbol transmission strategies. TACT-based learning algorithm enhances the process of adaptation to the channel conditions. This increases the throughput and spectrum utilization. Our cognitive spectrum decision can be applied in multimedia communications over CRNs.

References

1. Commission, F.C., et al.: Spectrum Policy Task Force Report, FCC 02-155 (2002)
2. Shokrollahi, A.: Raptor codes. IEEE Transactions on Information Theory **52**(6), 2551–2567 (2006)
3. Iannucci, P.A., Perry, J., Balakrishnan, H., Shah, D.: No symbol left behind: a link-layer protocol for rateless codes. In: Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, pp. 17–28. ACM (2012)
4. Hu, F., Koushik, A.M.: Intelligent Spectrum Mobility based on Transfer Actor-Critic Learning for Rateless Transmissions in Cognitive Radio Networks. IEEE Transactions on Mobile Computing, Submitted (2015)
5. Wu, Y., Hu, F., Kumar, S., Sun, Q.: Apprenticeship Learning based Spectrum Decision in Multi-Channel Wireless Mesh Networks with Multi-Beam Antennas. IEEE Transactions on Mobile Computing, Submitted (2015)
6. Giupponi, L., Galindo-Serrano, A., Blasco, P., Dohler, M.: Docitive networks: an emerging paradigm for dynamic spectrum management [dynamic spectrum management]. IEEE on Wireless Communications **17**(4), 47–54 (2010)
7. Zhi, X.-Y., He, Z.-Q., Wu, W.-L.: A Novel Cooperation Strategy Based on Rateless Coding in Cognitive Radio Network. IJACT: International Journal of Advancements in Computing Technology **4**(8), 333–347 (2012)

Elastic Edge-Overlay Methods Using OpenFlow for Cloud Networks

Amer Aljaedi, C. Edward Chow and Jia Rao

Abstract The virtualization of cloud network requires flexible and effective techniques to accommodate the rapid changes in the network configurations and updates. OpenFlow protocol has attracted attentions for cloud networks since it facilitates managing and sharing the network resources, and it can be utilized to create an overlay abstraction on top of the network infrastructure for the flow setup in the cloud. However, the traditional reactive flow setup of OpenFlow introduces higher flow latency and overhead on the network controller. This paper discusses the issues of the reactive flow setup and presents two optimized overlay network virtualization methods that leverage OpenFlow to control and forward the tenants' traffic. The proposed methods enable tenants to use their own MAC/IP addresses in the cloud. We have implemented and evaluated the proposed overlay methods, and the experimental results show that our methods have less flow latency than the traditional reactive approach, and higher performance than the popular overlay tunneling protocols such as VXLAN, STT, and NVGRE.

Keywords OpenFlow · Overlay · Cloud network · Flow rules · Reactive rules · Proactive rules · Flow setup · Virtual network

1 Introduction

Network virtualization is one of the key components for the multi-tenancy services in the cloud that enables the cloud provider to satisfy the customer requirements. As thousands of tenants subscribe to the cloud services on a daily basis, the cloud has to handle a massive number of network configurations and updates. Therefore, the network virtualization in the cloud requires effective techniques for sharing the network infrastructure and maintaining an efficient

A. Aljaedi(✉) · C.E. Chow · J. Rao
Department of Computer Science, University of Colorado, Colorado Springs, USA
e-mail: {aaljaedi,cchow,jrao}@uccs.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_3

isolation of the tenants' traffic. It should support VMs migration to arbitrary locations in the cloud while it allows the cloud tenants to use their own addressing scheme and configure their virtual subnets.

The traditional network virtualization techniques such as VLAN cannot accommodate large multi-tenant datacenter due to its VLAN ID, 12 bits, which supports only 4096 virtual networks. Therefore, the cloud industry has adopted tunneling (L2-in-L3) overlay protocols such as VXLAN, NVGRE, and STT to address the VLAN limitations. These tunneling protocols encapsulate the whole Ethernet frame of VM in an IP packet in order to transmit the VM frame to its destination through a tunnel in the cloud physical network. This encapsulation technique hides the MAC/IP addresses of the tenant virtual network (TVN) from the cloud physical network, and it maps the tenants' virtual addressing schemes to the cloud physical topology. These tunneling protocols can accommodate millions of tenants in the cloud since the virtual network ID (VNID in the outer headers) is 24 bits in VXLAN and NVGRE, and 64 bits in STT (Context ID field). On the other hand, these tunneling protocols have introduced manageability and compatibility issues in the traditional networks.

Usually, the VM fragments the packetinto standard MTU-size without considering the additional tunneling headers since the tunneling process is transparent to VMs. Consequently, the frame is fragmented again after the tunneling encapsulation [1], which affects the network performance [2]. Furthermore, VXLAN and NVGRE depend on a multicast-enabled network for forwarding the tenants' traffic, which adds more complexity on troubleshooting the network problems, besides the underlying network has to handle a large number of multicast trees [3]. The STT protocol was designed to utilize the standard offloading capabilities in the network interface cards (NIC) to improve performance. However, since it uses a TCP-like header in L4 of the outer headers (i.e., it does not engage in the usual TCP 3-way handshake), it is treated as an invalid packet by the traditional network security appliances. The NVGRE cannot utilize ECMP-based load balancing since it uses GRE protocol for encapsulation, which does not have a standard transport layer (TCP/UDP) header.

This paper presents two flexible edge-overlay methods for network virtualization in the cloud datacenter. Both methods apply the same principles and leverage OpenFlow to rewrite the addresses of the VM frame before transmitting the frame through the cloud physical network. Our methods allow the tenants to use their own MAC/IP addresses and forward the tenants' traffic without relying on multicasting or IP encapsulation.Consequently, they eliminate the limitations of the tunneling protocols. The first edge-overlay method can be utilized by the cloud provider that forward traffic based on layer two (i.e., L2 datacenter network fabric), while the second method is edge-overlay for cloud datacenters that rely on layer three network infrastructure. Also,performance evaluation of the proposed edge-overlay methods, compared to VXLAN, STT, and GRE, is included in this paper.

The rest of this paper is organized as follows. Section 2 provides an overview of OpenFlow, and it discusses the flow rules for traffic forwarding in the cloud. Section 3 presents the edge-overlay methods for L2 and L3 fabrics of the datacenter networks. Section 4 elaborates on the implementation of the proposed methods. Section 5 shows the experimental results of the proposed OpenFlow-based overlay with detailed analysis, and Section 6 surveys the related work. Finally, Section 7 concludes this research and highlights the future work.

2 OpenFlow

Under the current SDN paradigm, OpenFlow protocol has gained broad support throughout the networking industry as you can easily find OpenFlow-enabled commodity switches with reasonable prices in the networking market. OpenFlow allows more flexibility in managing, programming, and dynamically controlling the whole network devices by a logically centralized controller. The network operators/applications can specify a global network policy via the controller, and then the controller translates this high-level policy into low-level instructions that add/delete/modify flow entries in the flow tables of the related switches in the network. It has been utilized in the large datacenters such as Google B4 [4] and VMWare/Nicira NVP [5] to control and forward the network traffic efficiently.

The OpenFlow controller can install the forwarding flow rules in switches using either reactive or proactive flow setup. In the reactive approach, when the ingress switch receives a packet, it performs lookup for a match in its flow table (i.e., based on headers of the received packet) to forward that packet to one or

```
in_port=1, ip, nw_dst=0.0.0.2
actions=mod_dl_dst:68:54:A1:05:53:48,
mod_nw_dst:172.16.0.1, output:2
in_port=1, ip, nw_dst=0.0.0.3
actions=mod_dl_dst:52:54:00:af:87:2a,
mod_nw_dst:10.12.11.22, output:3
in_port=1, ip, nw_dst=0.0.0.4
actions=mod_dl_dst:42:34:01:ab:87:a1,
mod_nw_dst:192.168.0.1, output:4
```

Fig. 2 Flow rules for incoming packets to OVS in host 1(L2network fabric).

```
in_port=1, ip, dl_dst=00:00:00:00:00:02
actions=mod_dl_dst:68:54:A1:05:53:48,
mod_nw_dst:172.16.0.1, output:2
in_port=1, ip, dl_dst=00:00:00:00:00:03
actions=mod_dl_dst:52:54:00:af:87:2a,
mod_nw_dst:10.12.11.22, output:3
in_port=1, ip, dl_dst=00:00:00:00:00:04
actions=mod_dl_dst:42:34:01:ab:87:a1,
mod_nw_dst:192.168.0.1, output:4
```

Fig. 3 Flow rules for incoming packets to OVS in host 1 (L3 network fabric).

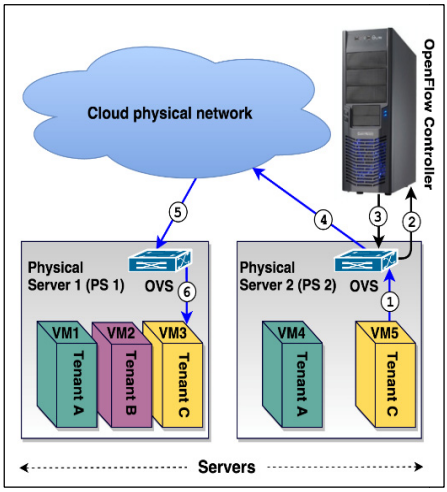


Fig. 1 Overview of the proposed OpenFlow edge-overlay.

more egress port(s). If it does not find the matching flow rule, it will drop the packet or send it to the controller via OFPT_PACKET_IN message for forwarding decision. Typically, there is a flow entry in flow table called table *miss*, which specifies how to process unmatched packets. After receiving PACKET_IN message, the controller installs the flow rules on the related switches along the flow path by sending OFPT_FLOW_MOD message to these switches.

This reactive flow setup provides a fine-grained flow visibility for the network controller, and it saves the switch memory since the reactive flow rules have an idle and hard timeout for their expiration. However, this traditional OpenFlow reactive approach increases the workload on the controller and the flow latency in the network [6,7]. For example, for each bi-directional flow setup, there will be: $2N_{FLOW_MOD} + 2_{PACKET_IN} + 2_{PACKET_OUT}$transmitted control messages between the controller and N number of switches along the flow path. Consequently, the flow latency is increased along with the increase of network diameter since each switch in the flow path has to process OFPT_FLOW_MOD message and install the flow rules.

Our proposed OpenFlow edge-overlay methods emphasize limiting the reactive control messages by using hybrid flow setup. Here, whenever VM is migrated to another location in the cloud and connected to a virtual switch, which is running on the physical host server, the controller proactively installs flow rules for the incoming packets to that VM in the virtual switch. Thus, the reactive flow setup is used only for the outgoing traffic in order to instruct the virtual switch to rewrite the headers of the outgoing packet before transmitting that packet to its destination through the cloud physical network (see Section 3).This hybrid flow setup reduces the control messages to $2_{PACKET\_IN} + 2_{FLOW\_MOD+PACKET\_OUT}$ for each bi-directional flow. Also, the flow latency is reduced in this approach since only the ingress virtual switch processes the FLOW_MOD message for each new flow. Note, the OpenFlow controller knows the location of every VM in the cloud (i.e., virtual switch ID, and the port number where the VM is connected to) as follows:

- When the cloud controller (e.g., OpenStack) configures and adds a new VM to the virtual switch, the switch sends an OFPT_PORT_STATUS message to notify the OpenFlow controller of the change. Also, when VM is migrated to another host, it sends gratuitous ARP, which is intercepted and sent by the virtual switch to the OpenFlow controller.Hence, the controller can obtain MAC/IP and location of themigrated VM(i.e., Sender Protocol Address (SPA) field in ARP has the IP address of the ARP sender).
- In addition, the OpenFlow controller can utilize the Neutron plug-in of OpenStack to obtain information about the VMs and their locations directly.

3 Edge-Overlay

This section presents the proposed edge-overlay for cloud infrastructure that uses L2 fabric (see Section 3.1) and elaborates on the rewriting techniques of the packets' headers using OpenFlow. It also shows how the same overlay principles

can be applied to the cloud that relies on L3 fabric (see Section 3.2). Our proposed overlay methods only require that the cloud physical switches should be OpenFlow-enabled, which is easily attainable nowadays as OpenFlow has been widely supported in the networking industry. This requirement is necessary to control the MAC-learning among the intermediate L2 switches and prevent the MAC table explosion problem as OpenFlow can help to reduce the forwarding tables by using hierarchical addressing scheme [8].

3.1 Overlay for L2 Fabric

In virtualized datacenter, the first switch that receives the VM packet is the virtual switch (e.g., OVS [9]) in the physical host. There we can rewrite the MAC/IP addresses of the packet in order to send it to its destination through the cloud physical network as shown in Figure 1. Here is the workflow:

1. VM5 in physical server 2 sends a packet to VM3 in physical server 1 (both VMs belong to Tenant C).
2. The virtual switch receives the outgoing packet to VM3. As the virtual switch does not know the destination, it sends `OFPT_PACKET_IN` to the controller.
3. As the controller keeps tracks of the location, MAC/IP addresses, and VNID of each VM in the cloud, it replies to the virtual switch and installs the reactive flow rule for forwarding the packet. The flow rule instructs the virtual switch to replace the destination MAC address of the packet with MAC address of the physical server 1 that hosts VM3 and replace the destination IP address with Virtual IP (VIP) address. This VIP address is used to tell the virtual switch in the destination physical server 1 how to forward the received packet (see Section 3.3).
4. After rewriting the destination MAC and IP addresses, the virtual switch transmits the packet to the physical network through the trunk port.
5. When the virtual switch in the physical server 1 receives the incoming packet, it knows that the packet should be forwarded to VM3 based on the VIP in the destination IP field.
6. The virtual switch in physical server 1 replaces the destination MAC and IP addresses in the received packet with the original MAC and IP of the VM3, and then it forwards the packet to the destination VM3. Figure 4 shows the addresses of the transmitted packet during VM-to-VM communication as described in the six steps above.

3.2 Overlay for L3 Fabric

As some cloud datacenters forward traffic based on L3, we also provide another edge-overlay for L3 datacenter network fabric. This overlay method applies the same 1, 2, 4, and 6 steps above-mentioned when forwarding the tenants' traffic. The difference here is in the steps 3 and 5. In step 3, the controller replies to the

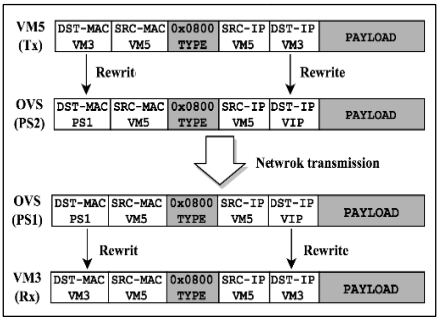


Fig. 4 Rewriting the packet headers during transmission (overlay for L2 fabric).

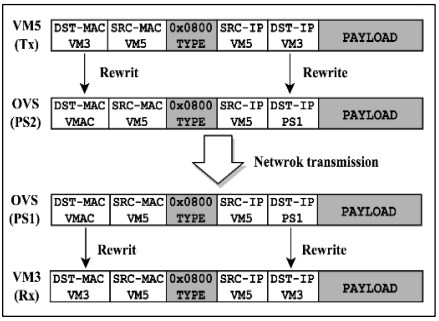


Fig. 5 Rewriting the packet headers during transmission (overlay L3 fabric).

virtual switch and installs flow rule to replace the destination IP address of the packet with IP address of the physical server 1 that hosts VM3 and replace the destination MAC address with Virtual MAC (VMAC) address. The VMAC is similar to VIP. It is just used to inform the virtual switch in the destination physical server 1 how to forward the received packet (see Section 3.3). In step 5, the virtual switch in the destination, where the target VM3 is connected, knows that the received packet should be forwarded to VM3 based on the VMAC in the destination MAC field. Figure 5 shows the addresses of the transmitted packet during VM-to-VM communication with edge-overlay for L3 datacenter network fabric.

3.3 Virtual Addresses

As it is highlighted earlier, the OpenFlow switch is relatively simple because the forwarding decision is defined by a controller, rather than by switch firmware. Whenever it receives a new packet, it checks the flow table(s) to find the corresponding action of the matching flow entry. The action field in the flow entry specifies how to forward the received packet. If it does not find a matching entry in its table(s), it will forward the packet to the controller. Our proposed methods were designed to reduce the number of the exchanged control messages between the virtual edge switches and the controller, which consequently reduces the workload on the controller for forwarding the tenants' traffic. We achieved this by using hybrid flow setup. In the hybrid approach, the controller proactively installs flow rules for the incoming traffic to the VMs, which are connected to the virtual switch. Thus, the virtual switch only sends the reactive `OFPT_PACKET_IN` message to the controller for the outgoing traffic. Here, when the virtual switch receives an incoming packet via its trunk port, it can forward the received packet to destination VM based on VIP or VMAC as described below:

Virtual IP (VIP). VIP is a simplified methodto instruct the virtual switch in the destination on how to forward the received packet from the trunk port. For example, in step 3 of Figure 1, the controller knows that the destination VM3 is connected to port 4 of OVS in physical server 1. Consequently, it installs flow rule in the OVS of physical server 2 that rewrites the original destination IP address of the outgoing packet with VIP = "0.0.0.4". We used only the first byte of the VIPto encode the switch port number where the destination VM is connected to. Now, when the OVS in physical server 1 receives that packet, it replaces the destination MAC and IP fields in the incoming packet with the original MAC and IP of VM3, and then it forwards the packet through port 4. Thus, the header rewriting is transparent to VMs in the cloud. Figure 2 shows an example of flow rules for incoming packets into OVS of the physical server 1, which are installed proactively by the controller when VMs were connected to the virtual switch. In this example, the port number one of the OVS is a trunk port while VM1, VM2, and VM3 are connected to ports two, three, and four respectively. When the OVS receives a packetwith VIP = "0.0.0.4" in the destination IP field, the third flow rule in Figure 2 instructs the OVS to change the destination MAC address to "42:34:01:ab:87:a1", IP address to "192.168.0.1", and forward the packet through port 4 where the destination VM3 is connected. Note, in the hybrid flow setup, the controller deletes the proactive flow rules for the disconnected VMs to save the switch memory and keep the flow tables updated.

Virtual MAC Address (VMAC). VMAC is used for the second proposed edge-overlay in L3 fabric instead of VIP. Since the destination IP is used for forwarding purposes in the cloud physical network, which relies on layer three network infrastructure, we used the destination MAC field to tell the virtual switch in the destination how to forward the received packet from the trunk port. In our implementation, only the first byte of the VMAC is used as the port number where the destination VM is connectedto the virtual switch. Figure 3 shows an example of the flow rules for incoming packets into OVS of physical server 1, which can be installed proactively with the deployment of edge-overlay for L3 fabric. For example, when the OVS receives a packet with VMAC = "00:00:00:00:00:02" in the destination MAC field, the first flow rule for incoming traffic in Figure 3 instructs the OVS to change the destination MAC address to "68:54:A1:05:53:48", IP address to "172.16.0.1", and forward the packet through port 2 where the destination VM1 is connected to.

3.4 ARP Processing

When the OVS receives an ARP request from VM, it sends `OFPT_PACKET_IN` message to the controller. As mentioned earlier, the controller knows the MAC/IP addresses of all VMs in the cloud, so it replies with the flow rule that instructs the OVS to create an ARP reply (i.e., the ARP reply is created by using OVS-specific actions which are `move`, `load`, and`mod_dl_src` [10]) and send it to the same port where the ARP request came from.

4 Implementation

The proposed OpenFlow edge-overlay methods are implemented as Floodlight applications [11]. Both methods utilize `StaticFlowPusher` and `Forwarding` modules in the Floodlight controller to implement the aforementioned hybrid flow setup. Also, The `IOFMessageListener` module of the Floodlight is used to create the VM profile, which maps each VM to its location in the cloud (i.e., switch ID and port number where the VM is connected to) and the MAC/IP addresses of the physicalhost server. Our methods use the `StaticFlowPusher`module to proactively install and update the flow rules for incoming traffic into the virtual edge-switches, which is OVS in our experiments.These proactive flow rules are installed based on the information that is collected from `TopologyManager` and `DeviceManager` modules in Floodlight.

In our implementation, we modified the `Forwarding` module of the Floodlight to reactively install flow rules for the outgoing traffic in the virtual ingress switch only, which sends the `OFPT_PACKET_IN`message, and the ARP request flood is prevented as ARP is handled directly by the controller. Also, the `LinkDiscoveryManager` module in the Floodlight controller was utilized to notify our network applications of the changes in the network topology (i.e., added/removed switches) and the status of the network links.

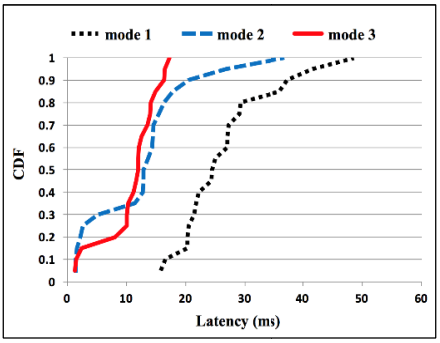


Fig. 6 Latency variation in a tree topology.

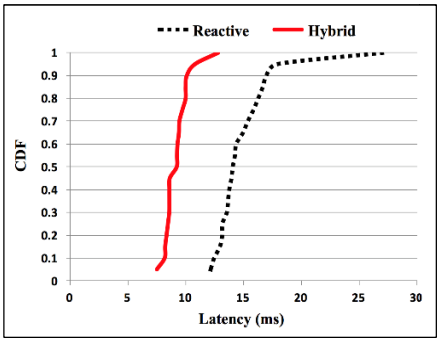


Fig. 7 Latency variation in a fatigued path.

5 Evaluation

This section presents and discusses the performance evaluation results of the proposed edge-overlay methods. We tested our overlay methods through a series of experiments. The first set of the experiments shows how the proposed hybrid flow setup helps to reduce the flow latency compared with the traditional reactive flow setup. The second set of experiments compares the network throughput of our edge-overlay to the tunneling protocols VXLAN, GRE, and STT.

5.1 Flow Latency

This set of experiments was conducted to investigate the impact of the traditional reactive flow setup on the flow latency and validate the enhancement using the hybrid flow setup. All the flow latency experiments were conducted in the same experimental environment. We used the Mininet emulator to configure and run different types of SDN topologies with using software OpenFlow switches. Mininet integrates Open vSwitch in its framework, which is a popular OpenFlow-compliant switch. Note, Mininet can create many OVS instances and configure them to run in kernel space. In our experiments, the network emulator was running in a server that is equipped with four CPUs, each Intel(R) Xeon(R) E5530 @ 2.40GHz, and 16 GB of RAM. Another server with the same above-mentioned resources was used to run the Floodlight controller.

In the first experiment, 243 hosts and 121 switches were emulated by Mininet to create a tree network topology with fanout 3 and depth 5. The emulated hosts generate traffic randomly using *Iperf* and create synthetic workload on the network. We chose two hosts in the network that are 9 hops apart and used ping responses to measure the latency between them in three modes. The first mode is the reactive flow setup for all flows in the network, including the flows from the two selected hosts. The second mode is same as the first mode except that the flow rules for the path between the selected two hosts are installed proactively in the intermediate switches. In the third mode, we deployed the hybrid approach and used our implemented applications on the controller. Here, the flow tables of the intermediate switches were populated proactively for traffic forwarding between the edge switches, and the flow rules for incoming traffic are installed proactively in the edge switches as described in Section 3.3. Every run was repeated 20 times, and the results are plotted in Figure 6. The ping response latency between the selected hosts is significantly high in mode one, compared to mode two and three. Also, the standard deviation of the ping responses' latencies in mode one was 8.4, while it was 4.6 in mode three. The maximum latency in mode one was 48.6 *ms,* whereas it was approximately 64% less in mode three. In this experiment, we did not configure any transmission or propagation delays between the network nodes in the emulator in order to focus only on the flow setup latencies in the reactive and hybrid approaches. As shown in Figure 6, the reactive flow setup introduces higher latency because all switches in the flow path have to process `OFPT_FLOW_MOD` message before forwarding the traffic.

The second experiment was aiming to investigate the impact of processing additional flow rules while there is traffic in transit, and testing the flow latency for this scenario in the reactive and hybrid approaches. Therefore, a linear topology with two edge switches and 12 intermediate switches was created in Mininet. We configured the controller to push randomly and continuously additional flow rules in the intermediate switches, and the ping responses were used to measure the flow latency between hosts, which are connected to edge switches. Each run was repeated 20 times. As you can see in Figure 7, generally the flow latency is lower in the hybrid approach, and the standard deviation of the responses' latencies in the hybrid setup is approximately 62% less than the reactive approach.

5.2 Network Throughput

The second set of experiments was designed to assess the performance of the proposed edge-overlay methods in an emulated cloud environment. Therefore, we used three physical servers and Kernel-based Virtual Machine(s) (KVM). Two servers were used for virtualization, server one and server two were hosting VM1 and VM2 respectively, and the third server was used as a controller. Figure 8 shows the machine specifications of the virtualized environment. OVS, version 2.4, was used as the virtual switch in server one and server two, and it was configured with OVS Linux-kernel modules, which are included in the OVS distribution. All the three servers were connected via Gigabit Ethernet switch. In this experiment, *Iperf* version 2.0.8 was used to send/receive UDP and TCP packets for 25 seconds in each run. The *Iperf* client was running in VM1 while the *Iperf* server in VM2. All the conducted experiments share the same environment.

VM 1 (Sender)	
OS	Fedora 22 (4.0.4)
CPU	1 core
Memory	2 GB
vNIC	virtio-net
Offloading	TSO, GSO, GRO

Physical Server 1	
OS	CentOS 7 (3.10.0)
CPU	8 × i7-2600 CPUs (3.40GHz)
Memory	8 GB
VMM	KVM
vSwitch	Open vSwitch 2.4
NIC	1000BASE-T
Offloading	TSO, GSO, GRO

VM 2 (Receiver)	
OS	Fedora 22 (4.0.4)
CPU	1 core
Memory	2 GB
vNIC	virtio-net
Offloading	TSO, GSO,GRO

Physical Server 2	
OS	CentOS 7 (3.10.0)
CPU	8 × i7-2600 CPUs (3.40GHz)
Memory	8 GB
VMM	KVM
vSwitch	Open vSwitch 2.4
NIC	1000BASE-T
Offloading	TSO, GSO, GRO

Fig. 8 Machine specifications.

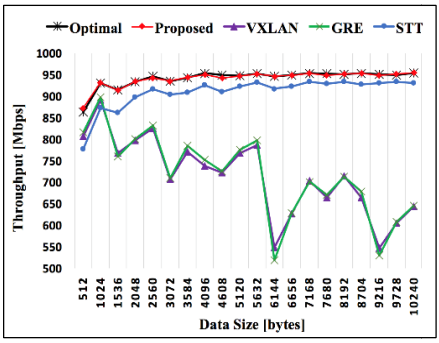


Fig. 9 UDP throughput.

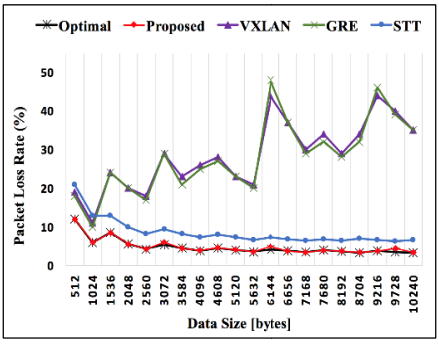


Fig. 10 UDP loss rate.

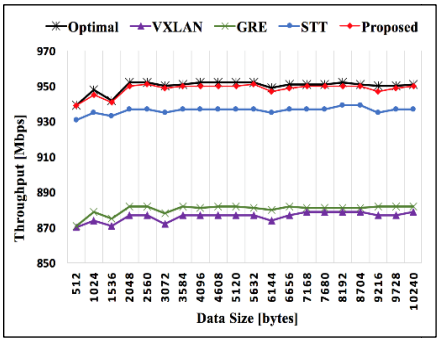


Fig. 11 TCP throughput.

The network throughput of the proposed edge-overlay methods are similar. Therefore, we present the performance results of the edge-overlay for L3 fabric and omit the results for the other overlay method due to the space limit.

The results of UDP throughput in the proposed edge-overlay are plotted in Figure 9 with the UDP throughput in VXLAN, STT, and GRE for comparison. Besides, the optimal throughput for VM to VM communication is included as the upper performance bound in the experimental environment. Here, the VMs were bridged to the physical network (i.e., no headers rewriting or tunneling encapsulation). The x-axis is the size of the transmitted bytes in UDP, excluding headers, by *Iperf* client. The y-axis is the throughput measured by the *Iperf* server. As you see in Figure 9, the STT performance was close to the optimal, while the proposed method almost matched the optimal throughput. The overhead of the outer headers in STT does not affect its performance as it utilizes the offloading capabilities of NIC. In our method, there are no outer headers, so there are no any additional fragmentations, and with the support of OVS-kernel module, the performance was high. Contrarily, the throughput of VXLAN and GRE was far below the optimal, especially for large data chunks in UDP. The VXLAN specification recommends setting MTU size to a value that can accommodate the outer headers and avoid fragmentation. In our experiment, we kept the default MTU (i.e., 1500 bytes) to test all protocols under the same conditions and obtain a fair comparison. Comparatively, Figure 10 shows the packet loss rate, which is high in VXLAN and GRE. Whereas TCP retransmits the missingfragments, UDP drops the whole packet when it misses fragments, especially when processes send packets rapidly as it has a limited frame buffer. In our experiments, the *Iperf* client in the sender side was configured to consume all the available bandwidth. Thus, Figure 10 shows the packet loss rate increases gradually with VXLAN, notably when data chunk is multiple of 1500 bytes. The TCP in our overlay method can use the offloading capabilities of NIC same as STT. Figure 11 shows the throughput of TCP in the proposed method, which is even higher than STT and very close to the optimal throughput.

6 Related Work

NetLord [12] architecture encapsulates the VM frame with additional MAC/IP headers. The outer MAC headers have the addresses of the edge switches where the sender and receiver VMs are connected, and the outer destination IP has the tenant ID. Our solution does not require any additional headers as it rewrites the addresses in the original headers and restores them in the destination. VL2 [13] uses L3 fabric for forwarding traffic in Clos topology, and it uses IP-in-IP encapsulation for network virtualization. It relies on IP multicasting to handle the virtual network broadcasts. Kawashima et al. [2] proposed non-tunneling edge-overlay model for the cloud network. His model utilizes OpenFlow to rewrite the MAC addresses (source and destination) of the VM frame and replace them with the MAC addresses of the physical servers in the cloud. It uses VLAN tag as VM

identifier in the host. However, his model does not hide the addresses of the physical servers in the cloud from the tenants, which exposes the cloud network infrastructure to threats. Chen et al. [14] proposed a scalable L2 architecture that spans multiple datacenters across diverse geographical locations. Their architecture only serves the cloud provider that relies on L2 fabric datacenter networks. Guenender et al. [19] published non-encapsulation overlay technique for network virtualization, which replaces the addresses in VM packet with the addresses of the edge-switches, and they used the source TCP port number to identify the VM in the destination host. Their research focused only in TCP, while our overlay methods consider both TCP and UDP protocols.

7 Conclusion

In this paper, we have presented two OpenFlow-based overlay methods for the cloud networks. The first method is designed to serve L2 fabric datacenter network, while the second method can be used for L3 cloud networkinfrastructure. Our methods utilize OpenFlow for rewriting VM packet headers in order to forward the tenants' traffic through the cloud physical network with maintaining sufficient isolation. Both methods do not use any additional encapsulation headers, and they can handle broadcast traffic such as ARP with less overhead. We are planning to extend our design to forward traffic between multiple datacenters in different locations as future work.

References

1. Kawashima, R., Matsuo, H.: Performance evaluation of non-tunneling edge-overlay model on 40GbE environment. In: IEEE 3rd Symposium on Network Cloud Computing and Applications, pp. 68–74. IEEE Xplore (2014)
2. Kawashima, R., Matsuo, H.: Non-tunneling edge-overlay model using openflow for cloud datacenter networks. In: IEEE 5th International Conference on Cloud Computing Technology and Science, pp. 176–181. IEEE Xplore (2013)
3. Nakagawa, Y., Hyoudou, K., Shimizu, T.: A management method of IP multicast in overlay networks using OpenFlow. In: First Workshop on Hot Topics in Software Defined Networks, pp. 91–96. ACM, New York (2012)
4. Jain, S., Kumar, A., Mandal, S., Ong, J., Poutievski, L., Singh, A., Venkata, S., Wanderer, J., Zhou, J., Zhu, M., Zolla, J., Hölzle, U., Stuart, S., Vahdat, A.: B4: experience with a globally-deployed software defined wan. In: ACM SIGCOMM 2013, pp. 3–14. ACM, New York (2013)
5. Koponen, T., Amidon, K., Balland, P., Casado, M., Chanda, A., Fulton, B., Ganichev, I., Gross, J., Gude, N., Ingram, P., Jackson, E., Lambeth, A., Lenglet, R., Li, S., Padmanabhan, A., Pettit, J., Pfaff, B. Ramanathan, R., Shenker, S., Shieh, A., Stribling, J., Thakkar, P., Wendlandt, D., Yip, A., Zhang, R.: Network virtualization in multi-tenant datacenters. In: 11th USENIX NSDI, pp. 203–216. ACM, Berkeley (2014)

6. Curtis, A., Mogul, J., Tourrilhes, J., Yalagandula, P., Sharma, P., Banerjee, S.: DevoFlow: scaling flow management for high-performance networks. In: ACM SIGCOMM 2011, pp. 254–265. ACM, New York (2011)
7. Bu, K.: Gotta tell you switches only once: toward bandwidth-efficient flow setup for SDN. In: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 492–497. IEEE Xplore (2015)
8. Chen, C., Liu, C., Liu, P., Loo, B., Ding, L.: A scalable multi-datacenter layer-2 network architecture. In: 1st ACM SIGCOMM Symposium on Software Defined Networking Research, Article No. 8. ACM, New York (2015)
9. Pfaff, B., Pettit, J., Koponen, T., Jackson, E., Zhou, A., Rajahalme, J., Gross, J., Wang, A., Stringer, J., Shelar, P., Amidon, K., Casado, M.: The design and implementation of Open vSwitch. In: 12th USENIX NSDI, pp. 117–130. ACM, Berkeley (2015)
10. OVS ARP Responder – Theory and Practice. http://assafmuller.com/2014/05/21/ovs-arp-responder-theory-and-practice/
11. Floodlight Project. https://floodlight.atlassian.net/wiki/display/floodlightcontroller/For+Developers
12. Mudigonda, J., Yalagandula, P., Mogul, J., Stiekes, B., Pouffary, Y.: Netlord: a scalable multi-tenant network architecture for virtualized datacenters. In: ACM SIGCOMM 2011, pp. 62–73. ACM, New York (2011)
13. Greenberg, A., Hamilton, J., Jain, N., Kandula, S., Kim, C., Lahiri, P., Maltz, D., Patel, P., Sengupta, S.: VL2: a scalable and flexible data center network. In: ACM SIGCOMM 2009, pp. 51–62. ACM, New York (2009)
14. Guenender, S., Barabash, K., Ben-Itzhak, Y., Levin, A., Raichstein, E., Schour, L.: NoEncap: overlay network virtualization with no encapsulation overheads. In: 1st ACM SIGCOMM Symposium on Software Defined Networking Research, Article No. 9. ACM, New York (2015)

An Approach to Generate Automatic Variable Key to Assure Perfect Security in Cryptosystem

Subhasish Banerjee, Manash P. Dutta and C.T. Bhunia

Abstract Due to advancement and worldwide rapid deployment of computer networks, the security becomes a major issue. Many mechanisms have been devised and more new ideas are about to propose in coming era. But, all the existing techniques have only one important criterion that key must be sustained and protected at any circumstances. In this regard, the AVK is one of the novel and unbreakable approaches to fulfill such criteria as being experimented so far. In this paper, we have proposed a new technique to generate AVKs which can provide higher security by enhancing the randomness among the generated successive keys. To adhere to our claimed, the comparative studies with existing techniques have also been cited at the end.

Keywords Common divisor · Variable key · Randomness · Security

1 Introduction

As computer networks becomes an essential and important part of our daily lives, protecting information and information system from unauthorized access, disruption, modification, use, disclosure or destruction becomes a complicated and challenging task among the researchers. From last few decades, after putting rigorous effort, many new schemes and ideas have been proposed to fulfill such requirements. But, due to the advancement and growth of computer technology day by day, attaining the security becomes an abundant issue with a single key. In this regard, Shannon [1, 2] devised that if key changes from session to session or time to time or make the size of key as same as that of plain text then to break the code with cipher text only attack becomes infeasible. Herein, lies the necessity of Automatic Variable Key (AVK) [3–5]. AVK makes the key dynamic in nature rather than static one throughout the transfer of data communication between sender and receiver. AVK mechanism

S. Banerjee(✉) · P. Dutta · C.T. Bhunia
Department of CSE, National Institute of Technology, Yupia 791112, Arunachal Pradesh, India
e-mail: subhasishism@gmail.com, manashpdutta@gmail.com, ctbhunia@vsnl.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_4

generates the keys in such a way that key changes every time whenever a new block of data is exchanged. The main key features of AVK is that due to random variation among the successive keys the probability of the cipher text attacks such as frequency attack, differential frequency attack, brute force attack etc. has been reduced tremendously according to theoretical survey by the researchers. Thereafter, many researchers like Chakraborty et al. [6], Goswami et al. [7, 8], Banerjee et al. [9–11] and Dutta et al. [12–14] had proposed their ideas and gave their contributions towards the security over insecure communication channel. In this paper we have proposed a new technique for key generation which can upgrade the level of security by increasing the randomness among the successive keys. The rest of the paper has been organized as: proposed scheme with key generation examples have been included in section 2 and 3 respectively. In section 4, we have summarized our experimental results. To prove its excellency we have compared our scheme with the existing schemes in section 5. Finally, we have concluded in section 6.

2 Proposed Scheme

In this section, we have described our proposed technique namely Automatic Variable Key based on Common Divisor (AVKCD) to generate the automatic variable keys. In this technique, the keys used to change based on common divisor of greatest and the smallest (Other than One) between the previous key and previous block of data. The explanation of our proposed algorithm is mentioned in the following way:

$Key\ Generation(Initial\ Key, Data\ set)$
{
$K_1 = Initial\ Key$
$i \leftarrow 1$
$repeat$
{
$R \leftarrow GCD(K_{i-1}, D_{i-1})$
$S \leftarrow SCD(K_{i-1}, D_{i-1})$
$if(S == 0)then$
$S \leftarrow 1$
$K_i = CRS(K_{i-1}, R) \oplus CLS(D_{i-1}, S)$
$i++$
}$until(Data\ set \neq \emptyset)$
}

Where:
$CLS(x, y)$ = Circular Left Shift of x by y bit positions.
$CRS(x, y)$ = Circular Right Shift of x by y bit positions.
$GCD(x, y)$ = Greatest Common Divisor of x and y
$SCD(x, y)$ = Smallest Common Divisor of x and y other than one

3 Thoretical Example of Key Generation

In this section, we have summarized our proposed technique by defining some key generation examples. In this scheme, the key generator starts the operation to produce the successive keys after the establishment of initial key between the sender and receiver. For simplification, we have considered the block size of plain text and initial key are 8 bits only. Here, we have assumed the initial key K_1 as 10001010 and initial data D_1 as 10101010. The necessary steps for generating the successive keys are given as below:-

Step 1 The GCD and SCD of decimal equivalent of $K_1(= 138)$ and $D_1(= 170)$ are 2 and 2 respectively. Therefore, 2^{nd} key, K_2 can be calculated by $CRS(10001010, 2) \oplus CLS(101010, 2) = 10100010 \oplus 1010100 = 00001000$.

Step 2 To compute 3^{rd} key K_3, assumed 2^{nd} data block D_2 was 11011000, hence GCD and SCD of decimal equivalent of $K_2(= 8)$ and $D_2(= 216)$ are 8 and 2. Henceforth, K_3 is $CRS(00001000, 8) \oplus CLS(11011000, 2) = 00001000 \oplus 01100011 = 01101011$.

Step 3 To yield the next key K_4, 3^{rd} data block D_3 was assumed as 11100100, therefore GCD and SCD of decimal equivalent of $K_3(= 107)$ and $D_3(= 228)$ are found as 1 and 0 respectively. But, as per the algorithm, if SCD is 0 then it is assigned as 1. Hence, the computed value of K_4 is $CRS(01101011, 1) \oplus CLS(11100100, 1) = 10110101 \oplus 11001001 = 01111100$.

In this way, successive keys can be generated based on the availability of data set.

4 Experimental Results

In this section, we have demonstrated various experiments to show the efficiency of our proposed scheme. In all the following experiments, we have assumed that the length of all the individual datasets is 8 bits long. For simplification, we have taken 8 bit initial key $K_1 = 10001010$. While plotting the graph, we have considered first 100 keys only.

Experiment 1:-The dataset for this experiment is taken as *A brute-force attack involves trying every possible key until an intelligible translation of the cipher text into plaintext is obtained.* The generated successive keys shows the randomness which are depicted in Fig. 1.

Experiment 2:-Here, we have taken *Frequency analysis is based on the fact that, in any given stretch of written language, certain letters and combinations of letters occur with varying frequencies.* as data set to get the randomness and the graph is given in Fig. 2.

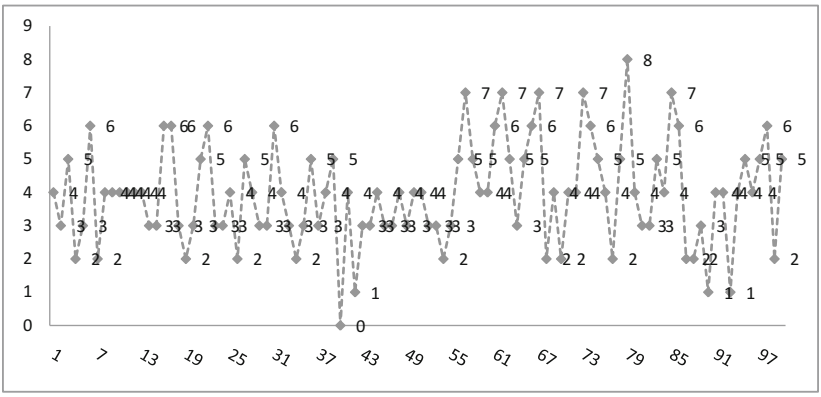


Fig. 1 Randomness of the successive keys

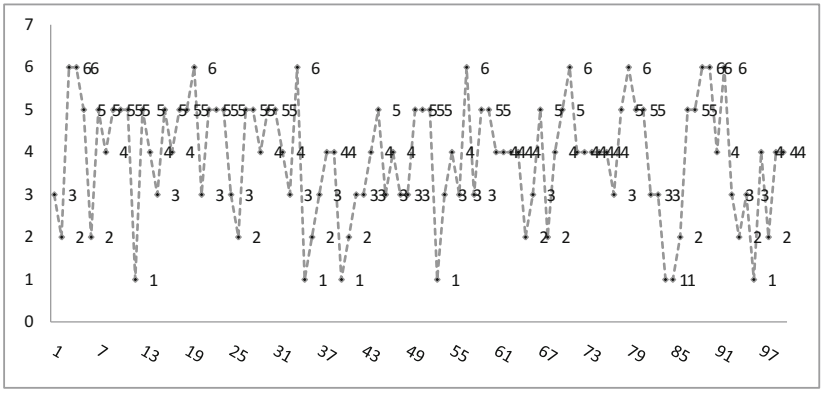


Fig. 2 Randomness of the successive keys

Experiment 3:-In the last experiment, the data set are taken as *An encryption scheme is said to be computationally secure if either of the foregoing two criteria are met.* to calculate the randomness of the auto generated successive keys and the generated graph is included in Fig. 3.

5 Performance Comparison

In this section, we have compared the efficiency of our proposed scheme by comparing the average randomness with the other related schemes. The average randomness has been calculated as the same fashion which has been described in the

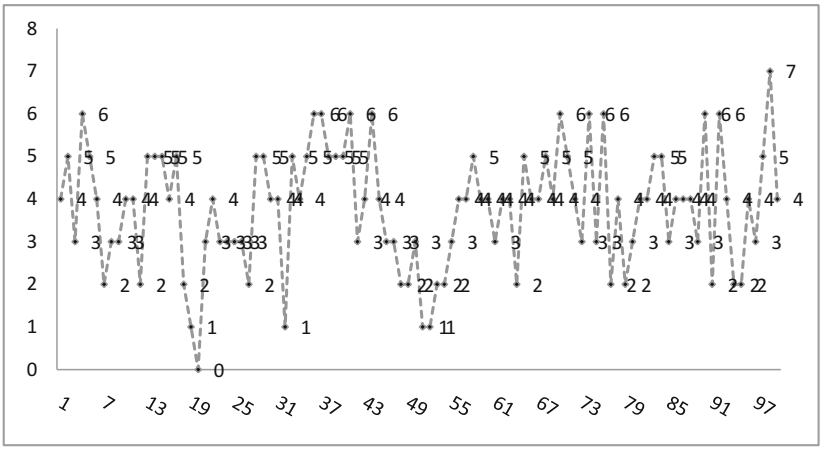


Fig. 3 Randomness of the successive keys

compared schemes. As because, the randomness among the keys is highly dependent on the initial key and data set pairs, we have taken all the above experiments into our account. The average randomness comparison graphs for the experiments 1 to 3 have been plotted in Fig. 4.

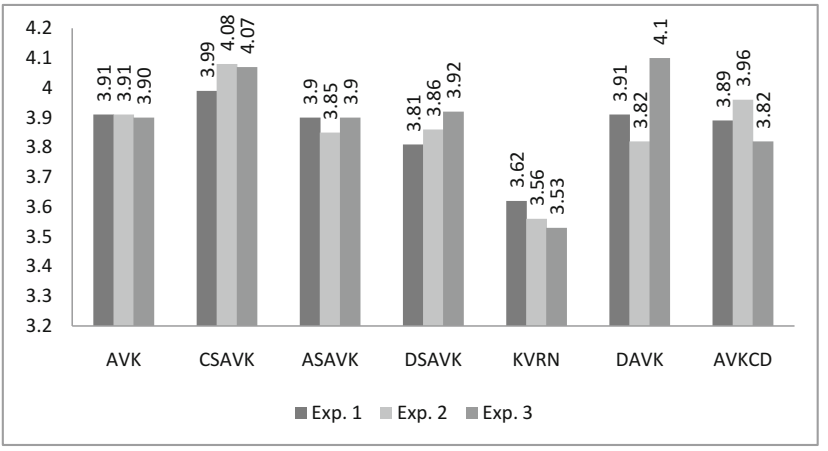


Fig. 4 Average Randomness Comparison with the Various Existing Techniques

To define the actual degree of heterogeneity, we have also compared our scheme using experiments 1 to 3 based on standard deviation as well. The corresponding graphs have been plotted from Figure 5 to 7.

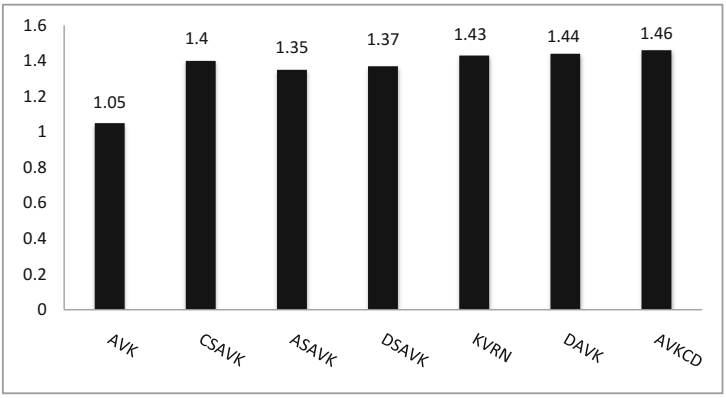


Fig. 5 Standard deviations comparisons for the experiment 1

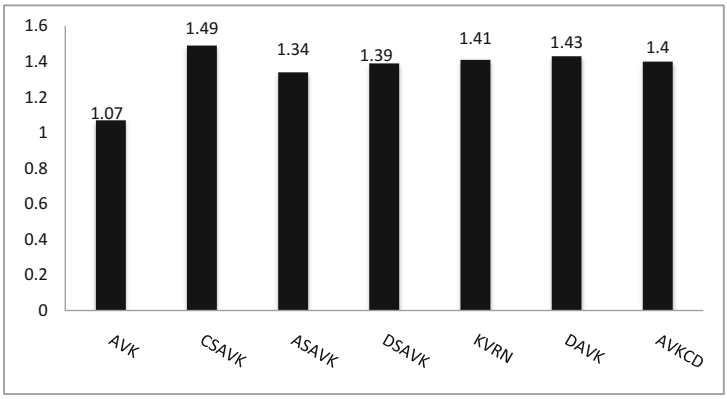


Fig. 6 Standard deviations comparisons for the experiment 2

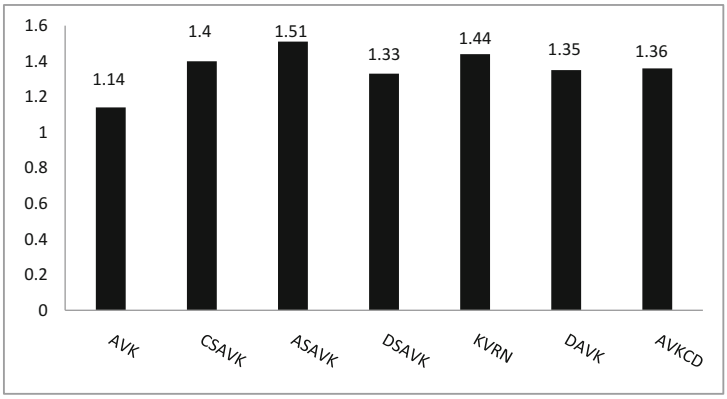


Fig. 7 Standard deviations comparisons for the experiment 3

6 Conclusion

In this paper, we have proposed a new mechanism to generate the AVKs and named AVKCD. Based on the comparisons performed with various related experiments given under performance analysis section, it can be stated that our technique is superior to other techniques in terms of standard deviation as it is the best way to define the actual degree of heterogeneity than average randomness. Therefore, it can be concluded that our proposed scheme can enhance the security by changing the keys from session to session during data communication.

References

1. Shannon, C.E.: Mathematical theory of communication. Bell System Technical Journal **27**(379–423), 623–656 (1984)
2. Shannon, C.E.: Communication theory of secrecy system. Bell System Technical Journal **28**, 656–715 (1949)
3. Bhunia, T.C.: New approaches for selective AES towards tackling error propagation effect of AES. Asian Journal of Information Technology, 1017–1022(2006)
4. Bhunia, T.C.: Information technologynetwork and internet, vol. 1. New Age Publication (2005)
5. Bhunia, T.C., Mondal, G., Samaddar, S.: Theories and application of time variant key in RSA and that with selective encryption in AES. Proc. EAIT, 219–221 (2006)
6. Chakraborti, P., Bhuyan, B., Chowdhuri, A., Bhunia, C.T.: A novel approach towards realizing optimum data transfer and automatic variable key (AVK). International Journal of Computer Science Network Security **8**(5), 241–250 (2008)
7. Goswami, R.S., Chakraborty, S.K., Bhunia, A., Bhunia, C.T.: Generation of Automatic Variable Key under various Approaches in Cryptography System. Journal of Institute Engineering India Series B **94**(4), 215–220 (2014)
8. Goswami, R.S., Chakraborty, S.K., Bhunia, A., Bhunia, C.T.: Approach towards Optimum Data Transfer with various automatic variable key techniques to achieve perfect security with analysis and comparison. International Journal of Computer Applications **82**(1), 28–32 (2013)
9. Banerjee, S., Dutta, M.P., Bhunia, C.T.: A Novel Approach to Achieve the perfect security through AVK over Insecure Communication Channel. Journal of Institute Engineering India Series B (Communicated)
10. Banerjee, S., Dutta, M.P., Bhunia, C.T.: A New Three Dimensional Based Key Generation Technique in AVK. Journal of Institute Engineering India Series B (Communicated)
11. Singh, B.K., Banerjee, S., Dutta, M.P., Bhunia, C.T.: Generation of automatic variable key to make secure communication. In: International Conference on Recent Cognizance in Wireless Communication and Image Processing-ICRCWIP-2014 (2015)
12. Dutta, M.P., Banerjee, S., Bhunia, C.T.: An Approach to Generate 2-Dimensional AVK to Enhance Security of shared Information. International Journal of Security and Its Applications **9**(10), 147–154 (2015)
13. Dutta, M.P., Banerjee, S., Bhunia, C.T.: Two new schemes to generate automatic variable key (avk) to achieve the perfect security in insecure communication channel. In: Proceedings of the International Conference on Advanced Research in Computer Science Engineering and Technology (ICARCSET, Eluru, India) (2015)
14. Dutta, M.P., Banerjee, S., Bhunia, C.T.: Generation of variable session keys based on piggy-backing strategy. In: Proceedings of 3rd International Conference on Advances in Computing, Electronics and Communication (Zurich, Switzerland) (2015)

Airborne Networks with Multi-beam Smart Antennas: Towards a QOS-Supported, Mobility-Predictive MAC

Xin Li, Fei Hu, Lei Hu and Sunil Kumar

Abstract Airborne networks require throughput-efficient MAC for mission-oriented communications. The use of multi-beam smart antennas (MBSAs) could provide the network with better throughput performance since all beams can send out data concurrently. In this paper, we propose a two-layer MAC design for MBSA-based airborne network. In the upper layer, we use TDMA-like schedule control to separate the packet collision domains in multi-beam data transmissions. In the lower layer, we use the CSMA/CA based scheme that is compatible with conventional 802.11 protocols. Such a two-layer scheme significantly reduces the packet contentions and thus improves the throughput. In addition, in order to support mission priorities in the network, we introduce the QoS-oriented MAC control in both upper and lower layers. Furthermore, a mesh network time synchronization method is proposed to guarantee the beam synchronization. A Hierarchical Dirichlet Process (HDP) enhanced Hidden Markov Model (HMM) is used for mobility prediction in each beam (direction) of a node. These approaches could better exploit the benefits of MBSAs. The simulation results show that our proposed MAC protocol outperforms the standard 802.11 DCF and general MBSA-based MAC designs. The validation of the QoS control, synchronization and prediction schemes are also evaluated. It turns out that these schemes could greatly improve the overall performance of the airborne networks.

Keywords Airborne networks · Multi-beam smart antennas (MBSAs) · MAC · QoS · Synchronization · HDP-HMM

X. Li(✉) · F. Hu · L. Hu
Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA
e-mail: xli120@crimson.ua.edu

S. Kumar
Electrical and Computer Engineering, SDSU, San Diego, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_5

1 Introduction

1.1 Multi-beam Antenna Airborne Networks

Airborne network performance is important in modern battlefields. Generally, an airborne network consists of wireless nodes (e.g., aircrafts, unmanned aerial vehicles (UAVs), etc). Figure 1 shows a typical scenario. Such a network has a wireless mesh network (WMN) topology. The large aircrafts serve as mesh routers (MRs). The nodes can change mission areas. The data to be transmitted in the WMN could be in multimedia format, such as video or images of different terrains. Those MRs form a wireless backbone with high data rate links among them. Other planes, usually UAVs, are regarded as mesh clients (MCs). MCs are connected to the closest MR directly or through other MCs to a MR.

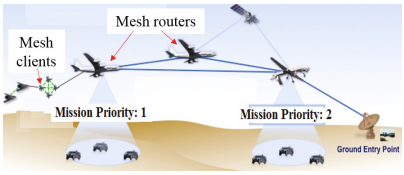


Fig. 1 An illustration of a mission-oriented airborne mesh network.

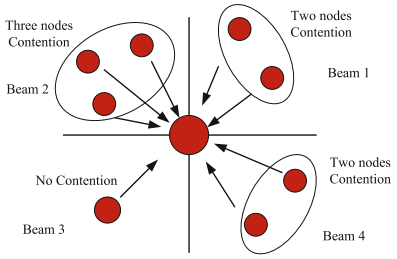


Fig. 2 Channel contention within beams

The Ku-band channel is used in today's airborne links. Ku-band[1], across 10GHz to 15GHz, has conventionally been reserved for satellite communication. Now it has been released for general long-range military or civilian applications. Its high frequency makes the data rate faster than Wi-Fi links. It also provides better anti-interference capability, which is due to its good directionality, especially after using directional antennas. But the main drawback is that it is difficult for such a short-wavelength signal to propagate through some objects such as trees, buildings, vehicles. This means that the signal can suffer from high path loss when no clear line-of-sight (LOS) is present.

Nowadays, the antenna technologies have made great progress. The size of the multi-beam antennas could be very small. It is feasible to equip the aircrafts in today's airborne networks with multi-beam smart antennas (MBSA). It enables cost-effective, concurrent multi-beam transmissions. Thus it significantly improves the WMN throughput.

1.2 Mission-Oriented Mobile Networks

The mission-oriented network has two important features: (1) Priority-aware communications: If any node captures important RoI data (such as an intrusion event), its traffic is marked with a higher priority than general scene data; (2) Predictable mobility trajectory: any node can establish a state transition model based on its history communication patterns. Unfortunately, none of the existing MBSA MAC protocols supports the above mission-oriented communications. Even though in higher layers such as routing layer and transmission layer, the protocols could support QoS, the performance of the network is still poor. There is a demand of the MBSA-based MAC layer which is able to handle both the QoS and node mobility.

In this paper, we propose an efficient MAC design for MBSA-based airborne networks. We design a hierarchical MAC for multi-beam antenna communications. In the lower layer, we propose the enhanced distributed/point coordination function (DCF/PCF) scheme. In the upper layer, we use time-slot-based overlay MAC control scheme. We then propose the time synchronization scheme to ensure the beam and node synchronization in multi-beam concurrent transmissions. Finally, beam communication pattern prediction model is used to predict the node mobility and plan the beam queue sizes.

The rest of the paper is organized as follows: In Section 2, we review some relevant work. The hierarchical MAC layer design is presented in Section 3. Section 4 provides our time synchronization scheme. Section 5 discusses the mission QoS control. The HDP-HMM based node mobility prediction is explained in Section 6. The simulation results are shown in Section 7. Section 8 has the concluding marks.

2 Literature Review

There has been much research work on directional antenna MAC designs. However, most of them simply assume that the antenna could communicate in only one direction each time [2]. With the development of antenna technology, researchers start to consider multi-beam antennas for high-frequency (such as millimeter-wave or Ku-band) communications. Among those works, [3] proposed a hybrid MAC protocol which assumed only 802.11 distributed coordination function (DCF) was used. This is not suitable to some contemporary MAC implementations (such as 802.11e) that emphasize the use of point coordination function (PCF).

In [4] a study on QoS-based MAC is conducted in Wi-Fi. It designed a set of polling-based MAC control protocols. Note that only PCF mode is improved in [4] compared to standard 802.11 protocols. The DCF mode has not been explored in multi-beam antennas. In [5] a distributed, receiver-oriented MAC with multibeam antennas, is designed. However, many practical MAC issues are not considered in [4] and [5]. The QoS support is not discussed there either.

3 Two-Layer Hierarchical MAC Design

3.1 Higher Level MAC

In the higher layer, we use a coarse time slot management to divide the time into different intervals. Each interval is called a superframe. It only allows one node to be the receiver or the sender. The reason of separating the collision domains (by only allowing one node to receive data from all beams), is that existing CSMA-based MAC tries to give each node equal chance to access the medium, which has been shown to significantly under-utilize the nominal network throughput [6], especially when there are very different sending rates among the neighbors. Another shortcoming of existing 802.11 protocols is that they do not give the nodes that need to help to relay other nodes′ traffic more opportunities to access the resource, and hence generate a suboptimal resource allocation. The TDMA-like MAC can also help to handle prioritized flows by controlling slot allocations among nodes.

We use a time slot T to divide the time into different superframes. Note that T is a very coarse time duration and is much larger than the clock synchronization error. Usually during time T, the node could send/receive several packets. If one node takes one role (e.g. being a receiver) in a slot, this node may not be the same role in the next slot, since the role selection is based on probability calculations. This approach thus guarantees fairness between different nodes.

As mentioned before, in each slot, there is only one node allowed to be the receiver and all the remaining nodes are transmitters. The receiver node is called as the 'star' node in this paper. The 'star' node could also be a transmitter. Obviously, being the 'star' node means that it could achieve the maximum throughput within the specific time slot T. Thus one main issue in the higher layer is how to choose the 'star' node. When we try to address this issue, we need to make a balance between the total throughput of the network, the data priority and the fairness between the nodes.

We assume that each node in the network has its own ID. For simplicity, we assign an unique integer number to each node ID. The IDs in a neighborhood are known to each node by using neighbor discovery protocols. Each node in a MAC neighborhood can calculate a pseudo-random hash function for each node ID nearby:

$$Position(i) = Hash(ID(i), timestamp), i = 1, 2, ..., N \tag{1}$$

The value of the Hash function is normalized between $(0, 1]$. The variable $timestamp$ is the current time. It serves as the random function seed. A node with $ID = J$ wins the time slot only if the following equation holds true. In other words, the node with the maximum hash result wins the slot occupancy.

$$arg \max_{1 \le x \le b} Hash(ID(i), timestamp) = J \tag{2}$$

In this way, each node knows who is the current 'star' node and who will be the next. The Hash function is carefully chosen so that nodes have an equal chance to be the 'star' node. When we need to consider the QoS for different kinds of data, we should make some nodes have higher chances to be the 'star' node than others.

3.2 Lower Level MAC

The latest 802.11 standards (such as 802.11e) recommend the use of PCF for schedule control of each neighbor′s transmissions. Such a centralized control can poll each node to ask for desired data rates for each neighbor. In our study, we also keep PCF phase in the superframe for a few reasons: first, the existence of PCF is compatible with the latest 802.11 MAC standards; Second, in the above-mentioned upper MAC layer, we have selected a node as the star node, which is the only receiver or sender in that time slot. Naturally the star node can serve as point coordinator (PC) in the PCF mode.

From the above PCF mode in the upper layer, it seems that we could establish an effective MAC layer scheme simply based on the hash function. However, there are still some challenging issues. Figure 2 shows one possible problems. From the figure we see that in beams 1, 2 and 4, there are more than one node that want to send packets to the star node. According to the basic principles of MBSA communication [3], each beam can only communicate with one node at one time. So the problem is how to solve the contentions within the beam. The lower MAC layer needs to be compatible with the traditional CSMA/CA backoff scheme. Meanwhile, the channel contention within the beam could be addressed by the CSMA backoff algorithm.

In this paper, we propose an enhanced PCF / DCF scheme for the lower layer, which is shown in Figure 3. In the enhanced PCF phase, the 'star' node first sends out QoS query messages to the other nodes within the beam area to ask for the QoS information. The transmitting nodes reply with the QoS response messages. If there are more than one sender in one beam, the contention is solved in the collision resolution phase. Then, the star node knows the priority level of the data from each beam.

In the enhanced DCF phase, the first enhancement we made (compared to conventional DCF) is that we control the backoff timer for concurrent multi-beam transmissions. Unlike the conventional 802.11 DCF operations that use random backoff after each DIFS, here we use multi-beam antennas that require all nodes to be ready for receiving data simultaneously. To achieve such a requirement, first, all neighbors which are supposed to receive the data from the star node should synchronize their clocks (later on we will discuss our clock synchronization scheme); second, we need to make the entire node (including all of its beams) perform CW-based backoff together, instead of performing backoff in each beam individually. This is because that we cannot guarantee that the entire node is under accurate timing control if each beam waits for different contention windows (CWs). Particularly, we need

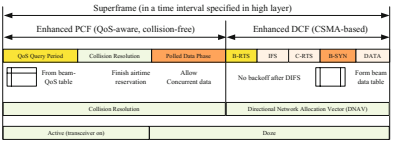


Fig. 3 Enhanced PCF and DCF in each superframe

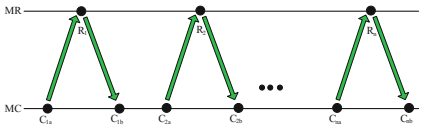


Fig. 4 Message exchange for synchronization

to remove CW-based backoff after DIFS for beam-synchronized communications. However, we still need to keep backoff timer for 802.11 compatibility.

The CW waiting time is expressed as:

$$W\_time = Random(seed) \times Delta\_Delay \tag{3}$$

4 Network Synchronization

As shown in figure 4, when a MR replies to a message from a MC, it checks the timestamp from the packet that contains the time at which this message is sent out. Meanwhile, MR knows its local time. So it is easy for the MR to figure out the time from the MC to MR. In the response message, MR tells MC the round trip time (RTT). The RTT is expressed as:

$$\Theta_i = C_{ib} - C_{ia} \tag{4}$$

The clock offset of the ith round trip is estimated as the mean of the time difference between MR and MC, which is:

$$\Delta_i = \frac{(R_i - C_{ia}) + (R_i - C_{ib})}{2} \tag{5}$$

The time in MC can be estimated as:

$$T_{MC} = T_{MR} - \Delta + \Theta \tag{6}$$

where $\Delta = \frac{1}{N}\sum_{i=1}^{N} \Delta_i$ and $\Theta = \frac{1}{N}\sum_{i=1}^{N} \Theta_i$.

5 Mission QoS Control

In airborne networks, the applications are mission-oriented, which means that different applications have different QoS requirements. So the MAC layer should support QoS control. Since our proposed MAC has two layers, we design QoS control for the two layers respectively.

5.1 Upper-Layer QoS Control

In mission-oriented networks, we assume that the applications have different priorities. In the upper layer, we assign nodes with different priorities with certain weights. According to [5], one of the effective approaches is that the priority is represented by the winning probability of the node:

$$P_i = H(ID(i), t)^{1/\omega_t} \tag{7}$$

$H(ID, t)$ is the function used in equation (1), while ω_t is the weight we assigned to the node.

5.2 Lower-Layer QoS Control

In the lower layer, our goal is to exploit the benefit of MBSA. For VBR, the desired airtime a node could have is [4]:

$$\Omega = \left(\mu + \sigma\sqrt{\frac{1-\xi_{VBR}}{\xi_{VBR}}}\right) \times \frac{SF}{R_{ch}} \tag{8}$$

Here ξ_{VBR} is the individual nodes tolerable degree ($0 \leq \xi_{VBR} < 1$) for insufficient airtime. SF is the length of SuperFrame, while R_{ch} is channel capacity.

Likewise, for CBR, the desired time will be:

$$\Omega = \mu(1-\xi_{CBR}) \times \frac{SF}{R_{ch}} \tag{9}$$

To sum up, based on equation (8) and (9), we could adjust the airtime of each node to meet QoS requirements.

6 Intelligent Communication via HDP-HMM

Airborne networks often transmit surveillance video streams among them. In the upper MAC layer, the entire interval should be dedicated to useful data transmissions. It will seriously decrease the communication efficiency if we ask the upper layer to exchange neighboring node arrival information frequently. Therefore, each node should use other ways to prepare the beam transmissions. The node state estimation and mobility prediction will be a good approach to low-overhead communications. A good news is, for mission-oriented communications the system typically has predictable mobility [7]. In the following descriptions, we will discuss our proposed HDP-HMM based prediction scheme.

HMM [8] can be used to deduce the internal state transition even if the observed values have noise or even missing data. In finite HMM, the system mode can only switch between pre-set finite number of states. Finite HMM can be simply described as a generative process as follows:

$$Z_t|Z_{t-1} \sim \pi_{Z_{t-1}} \tag{10}$$

$$y_t|Z_t \sim F(\theta_t) \tag{11}$$

We need to extend conventional finite-state HMM to an infinite-state scenario, where the number of states could change from time to time. Here we build *infinite* HMM via Dirichlet Process (DP):

$$G_0 = \sum_{i=1}^{\infty} \beta_i \delta_{\theta_k}, \quad \theta_k \sim H \tag{12}$$

The above DP-HMM adapts to *variable-state* Markov chain. However, it has a serious drawback: It assumes that the transition spectrum has non-overlapping support. This can cause large Markov state subspace and intolerable HMM convergence time. To reduce the prediction time and remove redundant states, we define a hierarchical specification $G_j \sim DP(\alpha, G_0)$, and G_0 itself is a draw from a $DP(\gamma, H)$. Therefore, we can extend (12) to a HMM with a prior distribution of hierarchical DP (HDP) as follows:

$$\begin{cases} G_0 = \sum_{l=1}^{\infty} \beta_i \delta_{\theta_k}, \ \beta|\gamma \sim DP(\gamma, H) \\ G_j = \sum_{l=1}^{\infty} \pi_{ji} \delta_{\theta_k}, \ \beta|\gamma \sim DP(\gamma, H) \end{cases} \quad and \ \theta_k \sim H \tag{13}$$

7 Performance Evaluation

In this section, we show the simulation results of our proposed MAC protocol. To evaluate the network performance, the two-layer MAC design in Section 3 is implemented. We also estimate the performance of the time synchronization and prediction schemes.

7.1 Simulation Setup

In our simulation model, we assume that there are 15 nodes in the network. Each node is equipped with a MBSA that contains 4 beams. All beams in one node could send out data simultaneously. The link capacity is set to be $2Mbps$ in each beam. So, ideally,

the maximum data rate that one node would achieve is $2Mbps \times 4 = 8Mbps$. The size of each packet is 1500Bytes. The average slot length in the upper layer is 100ms. In the following results, we mainly use network throughput and average packet delay to be our performance indicators.

7.2 Simulation Results

Figure 5 shows the throughput of the network with the average packet arrival rate. In Fig.5 we compare our proposed MAC protocol with another two schemes: traditional 802.11 DCF and multi-beam DCF (MB-DCF).

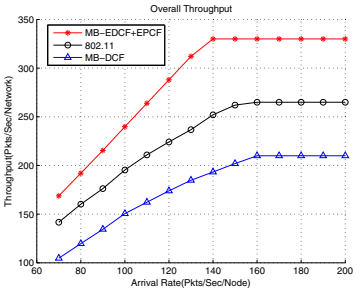


Fig. 5 Network Throughputs Comparison

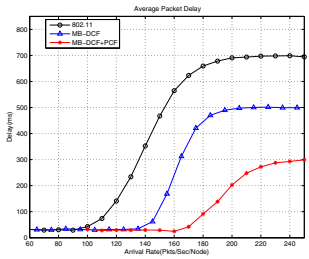


Fig. 6 Network Delay Comparison

We can easily see that our proposed scheme has the highest throughput compared to the remaining two ones. The conventional 802.11 DCF has the lowest throughput.

Figure 6 shows the average delay performance. It is also obvious that the proposed MAC has the shortest average time delay, while 802.11 DCF has the longest one.

In Fig.7, we show the benefit of the QoS control. In our model, all the nodes could send out data with 3 different priorities. The blue curves represent the throughput of data with different priorities when we do not have any QoS control in MAC layer. In that case the performances of those flows are similar to each other. However, with QoS control, the performance is greatly different with respect to priorities, which are shown in red curves. Our scheme could guarantee the QoS of the applications with higher priorities.

The synchronization scheme is also evaluated in this paper. Figure 8 displays the effect of using our proposed time synchronization algorithm. The throughput increases dramatically while the delay drops greatly. It is close to the performance under accurate clock synchronization.

Figure 9 shows the effect of using node state prediction schemes, including ARMA (Auto-Regressive Moving Average) and HMM based prediction schemes. As we can see, both HMM and ARMA achieve a higher throughput than no-prediction case. This is because each node can prepare the data queue better after predicting the

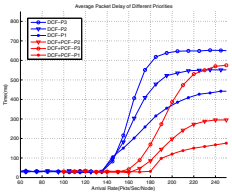


Fig. 7 Network Delay for Different Priorities

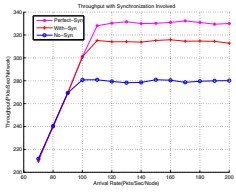


Fig. 8 Throughput with Synchronization

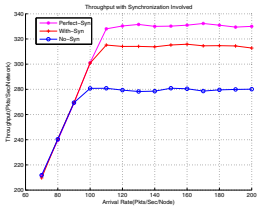


Fig. 9 State prediction Schemes

traffic profile and node mobility behaviors. Figure 9 also shows that both HMM and ARMA have the performance that is closer to the ideal prediction case. But HMM outperforms ARMA in most cases since it can overcome the impacts of noisy, missing or even erroneous measurements. HMM can deduce the internal state from the inaccurate measurements. While ARMA simply assumes that the measurement represents the internal true state.

8 Conclusions

In this paper, we have proposed an two-layer hierarchical MBSA-based MAC design. Out MAC layer protocols can be used in the MBSA-based airborne network. We presented the 2-layer MAC architecture. Then the QoS support and time synchronization approach were proposed, which could greatly improve the performance of the network. Finally we used HDP-HMM to predict the node mobility. This helps the star node to better prepare the queue content and schedule order. Our results have validated the proposed MAC designs.

References

1. Nagase, F., Mitsugi, J., Nakayama, M., Ueba, M.: Ku band mobile multimedia satellite communications system for trains. In: AIAA-paper-2003-2205, ICSSC (2003)
2. Vilzmann, R., Bettstetter, C.: A survey on MAC protocols for ad hoc networks with directional antennas. In: EUNICE 2005: Networks and Applications Towards a Ubiquitously Connected World, pp. 187–200. Springer (2006)
3. Jain, V., Gupta, A., Agrawal, D.P.: On-demand Medium Access in Multihop Wireless Networks with Multiple Beam Smart Antennas. IEEE Transactions on Parallel and Distributed Systems **19**(4), 489–502 (2008)
4. Chou, Z.-T., Huang, C.-Q., Chang, J.: QoS Provisioning for Wireless LANs With Multi-beam Access Point. IEEE Transactions on Mobile Computing **13**(9), 2113–2127 (2014)
5. Bao, L., Garcia-Luna-Aceves, J.: Transmission scheduling in ad hoc networks with directional antennas. In: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking, pp. 48–58. ACM (2002)

6. Rao, A., Stoica, I.: An overlay MAC layer for 802.11 networks. In: Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services, pp. 135–148. ACM (2005)
7. Xie, J., Wan, Y., Kim, J.H., Fu, S., Namuduri, K.: A survey and analysis of mobility models for airborne networks. IEEE on Communications Surveys & Tutorials **16**(3), 1221–1238 (2014)
8. Rabiner, L.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE **77**(2), 257–286 (1989)

CR Based Video Communication Testbed with Robust Spectrum Sensing / Handoff

Ji Qi, Fei Hu, Xin Li, A.M. Koushik, Lei Hu and Sunil Kumar

Abstract As radio spectrum is becoming congested, wireless communications require more efficient spectrum usage. Recently, the cognitive radio (CR) techniques have become attractive as they can utilize any unused spectrum. In this paper, we build a CR video communication test-bed, which implements spectrum sensing and spectrum handoff functionalities in USRP boards. We have implemented compressive spectrum sensing, intelligent spectrum handoff, multi-video-flow transmission, under TDMA scheduling and Raptor codes for reliable video transmissions. By using compressive sensing in spectrum sensing method, the spectrum detection accuracy is improved without much algorithm complexity, and it is also robust to the noise uncertainty due to the use of cyclostationary features. To realize intelligent spectrum handoff, we have designed a real-time jamming detection scheme, as well as synchronized spectrum switching method. We have also implemented a multi-point TDMA-based communication system, which enables any node to send out multiple video flows to different neighbors with pipelined and scheduled data transmissions. We also proposed a special rateless codes called prioritized Raptor codes for more reliable video transmission and implemented in the GNU Radio applications. The proposed CR video transmission testbed can be used for new protocol testing purpose.

Keywords Cognitive radios · USRP · Compressive spectrum sensing · Intelligent spectrum handoff · TDMA · Raptor code

1 Introduction

Due to the fixed frequency allocation policy in today's wireless networks, the spectrum has become a scarce and precious resource in wireless communications.

J. Qi(✉) · F. Hu · X. Li · A.M. Koushik · L. Hu
Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA
e-mail: jqi5@crimson.ua.edu

S. Kumar
Electrical and Computer Engineering, SDSU, San Diego, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_6

However, a large portion of the assigned spectrum is not efficiently used. The utilization of assigned spectrum varies sporadically and geographically, and ranges from 15% to 85% in time [1]. Many Novel techniques have been investigated to solve this problem. Recently, the cognitive radio networks (CRN) [2] have become attractive as they can better utilize the existing wireless spectrum and provide high bandwidth through dynamic spectrum access (DSA).

According to the current fixed frequency allocation policy, the whole spectrum is divided into small bands with different ranges. Each band is exclusively used by a specific wireless system. In CRNs, there are two co-existing systems in the same frequency range, which are called primary system and secondary system [3]. Primary system is the licensed system with legacy spectrum. This system has the exclusive privilege to access the assigned spectrum. Secondary system represents the unlicensed cognitive system and can only opportunistically access the spectrum holes which are not used by the primary system. We call the user in the primary system as Primary User (PU) and the user in the secondary system as Secondary User (SU). By allowing the SUs to temporarily access the PU's under-utilized licensed spectrum, spectrum can be utilized effectively.

CRNs provide the capability of using or sharing the spectrum in an opportunistic manner. With dynamic spectrum access, a CRN can operate in the best available channel. More specifically, the CRN enables the users to (1) determine which portions of the spectrum are available and detect the presence of licensed users(PU) in licensed bands (spectrum sensing), (2) select the best available channel (spectrum management), (3) coordinate access to this channel with other SUs (spectrum sharing), and (4) vacate the channel when a PU is detected (spectrum mobility).

Different hardware test platforms have been built to study CRNs. The Universal Software Radio Peripheral (USRP) [4] is one of the most popular software defined radio (SDR) platforms: it implements front-end functionality and A/D and D/A conversion. It assumes that physical layer processing is done on the PC that hosts the device. The USRP connects to the PC through a high-speed interface, and the host-based software is used to control the USRP hardware to transmit/receive data. The USRP device has a motherboard that has the following subsystems: clock generation and synchronization, FPGA, ADCs, DACs, host processor interface, and power regulation. These are the basic components required for baseband processing of signals. A modular front-end, called a daughterboard, is used for analog operations such as up/down-conversion, filtering, and other signal conditioning. In stock configuration, the FPGA performs several DSP operations, which ultimately provide translation from real signals in the analog domain to lower-rate, complex, baseband signals in the digital domain. In most cases, these complex samples are transferred to/from applications running on a host processor that performs DSP operations. The code for FPGA is open-source and can be modified to allow high-speed, low-latency operations.

Recent years much research on CRN and SDR have been done with USRP devices [5–7]. However, very few of them have built a comprehensive video communication network to evaluate the performance of spectrum sensing and spectrum handoff strategies. Our test-bed serves as a real-world platform to validate experimental

ideas and verify protocols / algorithms in CRNs and real-time video transmission over wireless network. In our testbed, each laptop with its controlled USRP board, is called an USRP node .The USRP board is supported by an open-source software code repository, GNU Radio.

2 Compressive Spectrum Sensing

Compressive Sensing (CS) is developed in signal processing community, and is more efficient for the sparse signal sampling in the sensing step than the traditional Shannon-Nyquist sampling theorem. Other than using the Nyquist sampling which needs the sampling rate of more than double the highest frequency component in the signal, the CS uses much lower sampling rate by randomly collecting the samples from the entire sparse signal domain. Then the optimal method is used to iteratively reconstruct the original signal with little or no data loss. In our method, we directly sample the information from the cyclic domain without time-consuming signal reconstruction while reserving the important signature features for different types of modulated signals.

2.1 Cyclostationary Feature

Modulated signals are in general coupled with sine wave carriers, pulse trains, hoping sequences, or cyclic prefixes, and thus have built-in periodicity. Even though the data is a wide-sense stationary random process, these modulated signals are characterized as cyclostationary [8]. Therefore, cyclostationary domain can be used to analyze the feature of the signals which are not stationary but with periodical appearance in specific frequencies. Cyclostationary feature detectors have been introduced as a complex 2-D signal processing technique for recognition of modulated signals in the presence of noise and interference [9].

Cyclostationary signal $x(t)$ has the property as

$$m_x(t) = m_x(t + kT) = E[x(t)] \quad (k = 1, 2, ..., N) \tag{1}$$

where E is the expectation and estimation of the signal mean, T is the cycle period. Thus the signal autocorrelation is

$$R_x(t, \tau) = R_x(t + kT, \tau) \quad (k = 1, 2, ..., N) \tag{2}$$

Taking FT w.r.t τ, we get spectral correlation function (SCF) as

$$S_x^{\alpha}(f) = \lim_{T\to\infty} \lim_{\Delta t\to\infty} \int_{-\Delta t/2}^{\Delta t/2} \frac{1}{T} X_T\left(t, f + \frac{\alpha}{2}\right) X_T^*\left(t, f - \frac{\alpha}{2}\right) dt \tag{3}$$

The sufficient statistics used for the detection are obtained through non-linear squaring operation. Therefore, FFT-based methods are used in digital implementation of the cyclostationary detectors. Given N samples divided in blocks of T_{FFT} samples, a simplified SCF is estimated as

$$\widetilde{S}_x^{\alpha}(f) = \frac{1}{NT} \sum_{n=0}^{N} X_{T_{FFT}}\left(n, f + \frac{\alpha}{2}\right) X^*_{T_{FFT}}\left(n, f - \frac{\alpha}{2}\right) \tag{4}$$

where $X_{T_{FFT}}$ is the N-point FFT around sample n.

2.2 Compressive Measurement and Compressed Signal Processing

In time-random, the received signal is usually sensed by CS to get a low dimensional vector Y using a sensing matrix Φ as follows

$$Y = \Phi X \tag{5}$$

where X is an $M \times 1$ vector representing the Nyquist samples of $x(t)$, Y is the $N \times 1$ compressed measurement vector, and Φ is an $N \times M$ measurement matrix.

Therefore, a critical task in compressive sensing is to design the measurement matrix, which collects compressed signal measurements and fulfills the robust detection. Based on the matrix transformation theory we design the second-order measurement matrix. We not only detect the signals robustly under the colored noisy environment but also classify the different types of modulation signals in a wide spectrum.

For cyclostationary signal detection and classification, we use compressed signal processing (CSP) to reserve the signal geometry structure in the compressive domain. The classification we use here is defined as the CSP signal detector based on the hypothesis test, in order to distinguish different modulation signals between $\Phi(H_0)$ and $\Phi(H_1)$:

$$t_i = \min \| P_{\Phi}(t) X - P_{\Phi}(t) S_i \| \tag{6}$$

where the $P_{\Phi}(t)$ is the CSP detector and s_i is the signal of the ith user.

As shown in Fig. 1, our CSP detector is divided into two main parts. First, we add the cyclic feature into the CS random measurement and build a sampling matrix as described above. The sensing matrix is implemented via the low-rate sampling on the cyclic features, which are calculated via the filter banks.

Thus, with USRP hardware, we can incorporate autocorrelation function to visualize the cyclic features of various signals in real-time. By sending the RF signals from the USRP receiver to Matlab and then calculate the Cycle Domain Profile (CDP) via an efficient method to recognize the different modulated signals.

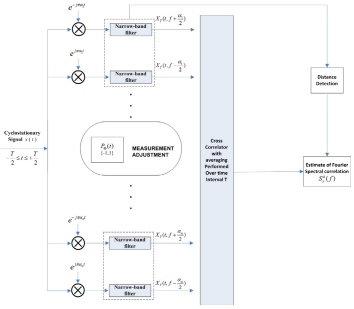


Fig. 1 Cyclostationary Compressed Detector.

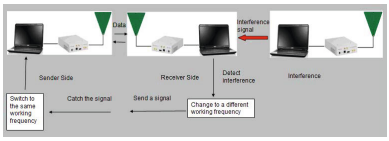


Fig. 2 Spectrum Handoff Framework.

3 Intelligent Spectrum Handoff

With the development of Cognitive radio techniques, spectrum handoff has become one of the hot topics recently. In CRN, when a primary user (PU) reappears on a channel, existing secondary users (SUs) must return the channel control to the PU and find other unused spectrum band(s) to switch. Spectrum handoff techniques can help the interrupted SU to vacate the occupied licensed channel and find a suitable target channel to resume its unfinished data transmission.

3.1 Framework Design

To realize spectrum handoff, we change the way that the conventional wireless system used to work. The third USRP node is used as an jamming node that will send jamming signals during the transmission process. We build a jamming event warning mechanism and make it possible for both sides to switch to the opposite mode (more details will be discussed in section B). When the receiver detects the jamming signal, it will send a warning signal to the sender side and changes to a different frequency. Then the sender will catch this signal and also switches to the same frequency. Fig. 2 shows our spectrum handoff framework.

During wireless video transmission, the quality of video can decrease sharply when the jamming signal comes. The reason is that the jamming signal will be in collision with the original video signal, which causes serious packet loss. We use the packet error rate (PER) to detect the invasion of the jamming signal.

In USRP implementation, we put a checksum into the head of each video packet. Thus the receiver can detect whether the current packet is correctly received or not. With this information, we can obtain the real-time PER of our system. However, the wireless link has fast fading. Even there is no jamming signal, the packets can still be discarded during the transmission. If every occurrence of a non-zero PER is judged as an invasion of the jamming signal, we will get many false alarms. Based on the above analysis, we can determine the invasion of the jamming signals by checking:

(1) PER is higher than a threshold δ, and (2) PER remains high for a sufficient time duration T. The whole system operates as follows:

- Both sender and receiver are set to a specific spectrum frequency f_1 and it starts video transmission.
- During the transmission, after every time interval ΔT, the sender side will switch to receiving mode with a different frequency f_2 for a short time Δt, and then change to sending mode again.
- The jamming node, starts to send jamming signals at frequency f_2.
- The receiver detects the jamming signal and switches to sending mode with the working frequency f_2.
- Once the sender side catches the signal, it will change to another working frequency f_3, and the receiver side will also switch to the same frequency f_3.

3.2 Multi-thread Scheduling

In GNU Radio programming, each action of the USRP hardware will be controlled by one thread. In the thread, users can build a *radio chain* by defining a flowgraph using the *connect* function. The connect function specifies how the output stream of a processing block connects to the input stream of downstream blocks. The flowgraph mechanism then automatically builds the flowgraph. However, as all of the inputs and outputs are predefined before running, it is impossible for the USRP hardware to perform two different actions (i.e., sending and receiving) in one thread. Therefore, to realize the mode switch in our spectrum handoff, we need to introduce multi-thread scheduling.

Both sender and receiver need to switch between sending mode and receiving mode. In software implementation, we adopt two-thread scheduling. Each thread runs independently according to the preset timetable. If the system notices that there is a jamming signal, one thread will send a warning message to the other one to change its RF frequency. The thread schedule on the sender and receiver is shown respectively in Figs. 3 and 4.

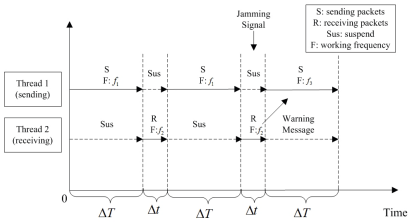


Fig. 3 Multi-thread Scheduling on Sender Side.

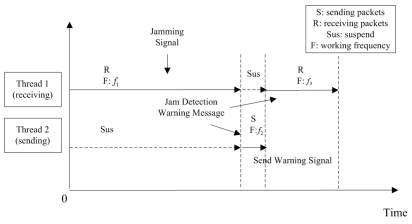


Fig. 4 Multi-thread Scheduling on receiver Side.

4 Network Protocols and Applications

4.1 Multi-video Transmission via TDMA Framework

Time division multiple access (TDMA) is a channel access method for a shared wireless medium. It is typically used in MAC layer. But scheduled communication can occur in any layer. It allows several users to share the same frequency channel by dividing the communication duration into different time slots. The users transmit in rapid succession, one after the other, each using its own time slot. This allows multiple stations to share the same transmission medium (e.g. radio frequency channel) while using only part of its channel capacity.

In our test-bed, we first use two USRP nodes to build a basic TDMA transmission network. We divide the whole transmission process into time slots and then send out two video streams alternately. According to the predefined time slot for each stream, the receiver can correctly decode and display two real-time video streams (we use ffmpeg, x264 and ffplay to encode, decode and display the video). The whole system is shown in Fig. 5. The whole system will work as the following steps:

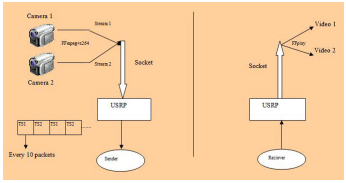


Fig. 5 TDMA Dataflow and Framework

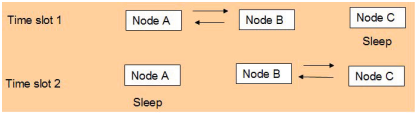


Fig. 6 Scheduled Data Transmission Framework.

– Set sender and receiver to a specific working frequency f and start transmission.
– The sender switches between two video streams every 10 packets.
– The receiver sorts two data-streams based on the predefined time slots and displays two real-time video streams.

Then we increase the number of USRP nodes to demonstrate how our test-bed performs the scheduled data transmission in a network. Each node can switch between sending, receiving and sleeping modes quickly and randomly. Taking 3-nodes case as an example (see Fig. 6), here we define the time slot schedule such that only two nodes can talk with each other at the same time while the third one goes to sleep.The whole system works as follows:

– All three nodes are set to a specific frequency f and start video transmission.
– Node A first begins to send packets and wakes up node B. Then node B starts to receive these packets, while node C keeps sleeping.
– After a time interval of Δt_1, node A changes to sleeping mode and node B starts to send packets and wakes up node C, at the same time node C will begin to receive.

– Again after a time interval Δt_2, node B will switch to sleeping mode and node C begins to send packets and wakes up node A, at the same time node A will change to receiving mode.
– After a time interval Δt_3, the system goes back to the second step.

4.2 Raptor Codes for Reliable Video Transmission

Fading and shadowing in wireless channels cause packet loss and deterioration of video quality. When packet loss occurs, the feedback from the receiver can be used for the request of the retransmission of the loss packets. This retransmission mechanism is bandwidth-costly. Raptor codes, which are a class of powerful rateless codes, can completely recover the source data with little overhead and linear encoding/decoding time.

The Raptor codes consist of a precode (usually a LDPC code) as the outer code and a weakened LT code as the inner code. They can be parameterized by $(K, C, \Omega(x))$, where K is the number of source symbols, C is a pre-code with block-length L and dimension K, and $\Omega(x)$ is a degree distribution of LT codes. Each encoded symbol is associated with an ID (ESI). The pre-code and weakened LT code can ensure a high decoding probability with a small coding overhead.

In our test-bed, we use the systematic Raptor codes [10]. If there are K source symbols $S[i]$ in one block, $i = 0, ..., K - 1$, the first K encoded symbols are constructed such that $E[0] = S[0]$, $E[1] = S[1]$, ..., $E[K - 1] = S[K - 1]$. The systematic Raptor codes can therefore correctly decode some source symbols even if the number of received encoded symbols N_r is less than the number of source symbols K.

As far as we know, no rateless codes (such as Raptor codes, LT codes) are implemented in GNU Radio and USRP platform. Only very simple FEC scheme, such as RCPC codes, is implemented in GNU Radio. Therefore, we built all Raptor codes programs from the scratch. In GNU Radio, applications are primarily written in Python programming language, while some performance-critical components are implemented in C++. Raptor codes, which include Gaussian Elimination and Belief Propagation, could consume many computing resources. Therefore we implemented the Raptor codes in C++. And then we used the SWIG software to transfer some Raptor codes C++ API into Python API to be called by GNU Radio applications. The implementation of Raptor codes for video transmission over USRP is shown in Fig. 7.

5 Experimental Results

This section presents the numerical and simulation results that verify the effectiveness of our proposed video CRN test-bed.

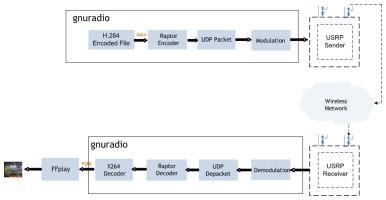


Fig. 7 Video Transmission Network with Raptor codes

5.1 *Spectrum Sensing*

Our proposed CSP detector is tested under different compressed rates and SNR levels, with the purpose of analyzing (1) the effects introduced by the sensing rate, (2) the detection accuracy, and (3) its robustness to the noise. As shown in Fig. 8, the classification errors quickly approach to zero under high SNR levels. The compression ratio is 50% for the worst case with a low error rate. It also means that half of the sensing energy can be saved.

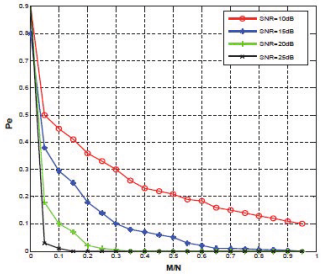


Fig. 8 CSP under different SNR levels

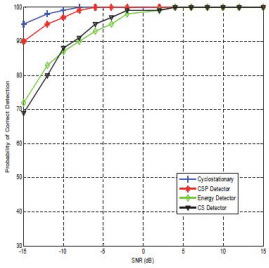


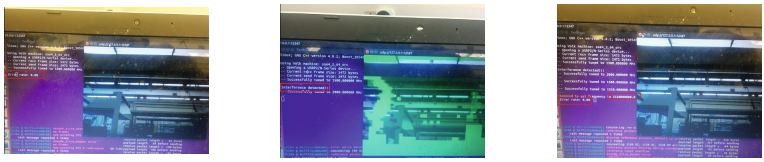
Fig. 9 Detection Probability for different detectors.

We also compare the receiver operating characteristic (ROC) of our proposed compressive cyclic feature detector with the ROC of the energy detector, traditional CS and cyclostationary detectors. ROC is typically used to evaluate the detectors sensitivity and accuracy, and a larger area under the curve means a better performance. As shown in Fig. 9, under a SNR of 0 dB and with a 50% compression ratio (M/N) for the CSP measurements, the performance of our system is much better than the energy detection and CS detection methods.

5.2 *Spectrum Handoff*

In this test, we use a real-time video transmission framework to test the channel switching accuracy and the response time of our intelligent spectrum handoff scheme.

In physical test the USRP sender node captures the real-time video with a built-in camera in the host computer, and then sends the video data out. And the jamming node sends out jamming signals. The performance of our spectrum handoff scheme in the receiver side is shown in Fig. 10.



(a) Before jamming (b) After jamming (c) Spectrum handoff

Fig. 10 Spectrum Handoff Hardware Test (Receiver Side)

Through our tests, we find out that the the jamming signal detection can be done in 2 seconds but the whole frequency switching process can last for 15 to 20 seconds. There are two reasons for this time consumption. Firstly, it takes a while for USRP hardware to suspend current working thread and change to another thread. Secondly, the sender side needs to switch to receiving mode for a short time Δt after every time interval ΔT. Thus if the sender side is working on sending mode, even the receiver side tries to tell the sender side that the jamming signal is detected, the sender side may not be able to hear this warning message within a short time.

5.3 TDMA Framework

To test and verify the performance of our TDMA scheme, we check both the two-node case and the three-node case. For multi-video transmission between two USRP nodes, we use a built-in camera in the host computer as well as a USB camera to capture two video streams simultaneously. The performance of our multi-video transmission scheme in receiver is shown in Fig. 11.

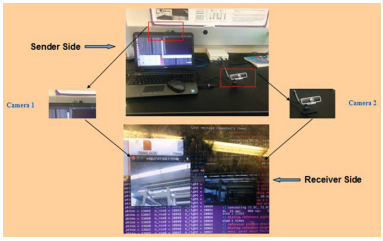


Fig. 11 Multi-video Transmission Hardware Test (Receiver Side)

Fig. 12 10-th frame of the received video. Left: video transmission without Raptor codes; Right: video transmission with Raptor codes

5.4 Raptor Codes

In our experiment, we evaluate the performance of video transmission over USRP with and without Raptor codes scheme. The raw video source is first encoded by H.264/AVC encoder and then encoded by Raptor codes at the Application layer. After that the encoded packets are encapsulated into UDP packets, which are modulated and transmitted over wireless networks using USRP. For comparison purpose, the raw video source is only encoded by H.264/AVC (without Raptor codes) and then gets transmitted by UDP protocol in USRP hardware.

From Fig. 12, we can see that the video quality with Raptor codes protection is much better than the one without Raptor codes protection.

6 Conclusions

In this work, we have built a CRN video communication test-bed based on USRP hardware. In our test-bed, we have incorporated some CRN techniques including spectrum sensing and spectrum handoff. To realize multi-video transmissions with high video quality, we have also adopted TDMA framework and Raptor codes. Experimental results showed that our test-bed provides a comprehensive CRN platform for video transmissions. In the future, we will continue to improve our test-bed framework, especially in CRN MAC layer and routing layer design.

References

1. FCC, E.: Docket no 03-222 notice of proposed rule making and order (2003)
2. Thomas, R.W., Friend, D.H., DaSilva, L.A., MacKenzie, A.B.: Cognitive networks. Springer (2007)
3. Akyildiz, I.F., Lee, W.-Y., Vuran, M.C., Mohanty, S.: A survey on spectrum management in cognitive radio networks. IEEE on Communications Magazine **46**(4), 40–48 (2008)

4. Ettus, M.: Usrp users and developers guide. Ettus Research LLC (2005)
5. Dhar, R., George, G., Malani, A., Steenkiste, P.: Supporting integrated MAC and PHY software development for the USRP SDR. In: 1st IEEE Workshop on Networking Technologies for Software Defined Radio Networks, SDR 2006, pp. 68–77. IEEE (2006)
6. Sarijari, M.A., Marwanto, A., Fisal, N., Yusof, S.K.S., Rashid, R., Satria, M.H., et al.: Energy detection sensing based on GNU radio and USRP: an analysis study. In: 2009 IEEE 9th Malaysia International Conference on Communications (MICC), pp. 338–342. IEEE (2009)
7. Marwanto, A., Sarijari, M.A., Fisal, N., Yusof, S.K.S., Rashid, R., et al.: Experimental study of OFDM implementation utilizing GNU radio and USRP-SDR. In: 2009 IEEE 9th Malaysia International Conference on Communications (MICC), pp. 132–135. IEEE (2009)
8. Gardner, W., et al.: Exploitation of spectral redundancy in cyclostationary signals. IEEE on Signal Processing Magazine **8**(2), 14–36 (1991)
9. Gardner, W.A., Napolitano, A., Paura, L.: Cyclostationarity: Half a century of research. Signal Processing **86**(4), 639–697 (2006)
10. Shokrollahi, A., Luby, M.: Raptor codes. Now Publishers Inc. (2011)

A Control-Message Quenching Algorithm in Openflow-Based Wireless Mesh Networks with Dynamic Spectrum Access

Xin Li, Xiaoyan Hong, Yu Lu, Fei Hu, Ke Bao and Sunil Kumar

1 Introduction

In recent years, Open Flow [1] is becoming a popular network architecture. As more and more users starts to join the conventional Internet, the drawbacks of the conventional networks are now gradually appearing. OpenFlow network provides us a brand new sight to the development of networks. This architecture separates the control plane and data plane from the hardware level(physically). OpenFlow has plenty of benefit compare to other network structures. Firstly, the control plane and data plane are decoupled, this means more flexible networks can be figure out with customized rules in the network. Secondly, OpenFlow provides us a new platform to design and test network protocols. Researchers could test new protocols in a real network environment. Thirdly, in OpenFlow architecture networks we could monitor flow traffic statistics. This fine-grained monitoring of flows enables us to better understanding the network protocols and scheme we applied.

Typically, the control plane is made up by the networking device named controller, which is the key part in OpenFlow architecture. The controller knows the overall topology of the network it manages. In addition, it figures out the path needed by routers and switches. When data plane devices have problems with the packets dealing. They will ask the controller for help. The major work for the data plane devices is relatively simple compared to that of controller. Routers and switches are responsible for executing the rule made by controllers and forwarding datagram in the network.

X. Li(✉) · Y. Lu · F. Hu · K. Bao
Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA
e-mail: xli120@crimson.ua.edu

X. Hong
Computer Science, University of Alabama, Tuscaloosa, AL, USA

S. Kumar
Electrical and Computer Engineering, SDSU, San Diego, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_7

However, there are still some challenges for OpenFlow architecture both in control plane and data plane. In data plane, one of the most challenging issue is the memory capacity requirements. As the network is growing lager, the data plane devices have to store more rules to handle the packets in the complicated network. The rules and other dynamic flow tables occupy TCAMs(Ternary Content Addressable Memory) in the devices. The TCAMs are precious and expensive. Therefore minimizing the usage of the storage is necessary for the data plane.

In this paper, we first modeled a complete wireless network architecture based on OpenFlow and Cognitive Radio Networks(CRN). The CRN is designed by FCC to improve the frequency utility efficiency by occupying the channel which the licensed user released. We employ the wildcard rule [2] to reduce the TCAM usage of the data plane devices. In CRN, a HDP model is applied to sense and classify the channel into different groups. Then we proposed an algorithm to reduce the number of request to the controller. In the proposed algorithm, we not only record the source-destination pairs, but also collect information from neighbors in order to reduce the request times. The simulation results show that our proposed algorithm improves the throughput and reduces the average packet delay of the network.

The rest of the paper is organizing as follows: In Section 2, we go over some of the related researches. The system design is proposed in section 3. We first introduce the network architecture and then explain the wireless wildcard rule. In section 4, we propose and analyze the controller bottleneck problem within the system. The Control Message Quenching(CMQ) algorithm with information from neighbors and its enhances version are provided in section 5. Section 6 shows the simulation results of channel selection and CMQ algorithms. Section 7 is the conclusion.

2 Related Work

There are plenty of research papers related to OpenFlow networks. Tie Luo in [3] proposed the CMQ algorithm. The algorithm could not only be applied to wired network, but also in sensor openflow network [7]. Min Lan proposed the DIFNE [9] system to solve the issues related to the rule of data forwarding in the data plane. The author in [5] proposed DeveFlow to devolve the control function back to switches. Other researches including enhance the OpenFlow scalability includes [11] and [12].

3 System Design

3.1 Wireless Mesh Networks(WMN)

In this paper, we mainly consider the Wireless Mesh Networks(WMN) architecture. WMN is one kind of network architecture with a wireless backbone network and plenty of wireless devices such as mesh routers(MR) and mesh clients(MC). Each MR is the ‘header’ of a certain region. Each MR is connected to one MR. When a MC

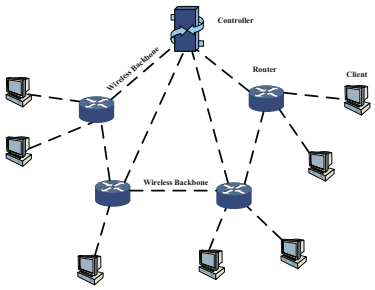


Fig. 1 Small Scale WMN

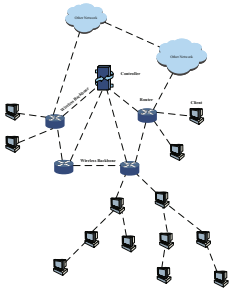


Fig. 2 Large Scale WMN

has data to send out to the network. Basically, WMN can be classified into two types according to its network size and scalability, namely small scale WMN and large scale WMN. For small scale WMN shown in Fig. 1, MCs are only one-hop from the MR in the region. The MR can contact all the MCs in its region within one-hop time. For large scale WMN shown in Fig. 2, some MCs are not connected to MR directly. There exits MC to MC connections. In this paper, we assume that the MCs form a 'tree-topology' within a mesh region since it is easy to implement synchronization method and routing algorithm. But the system complexity and performance will be compromised compare to small scale WMN.

3.2 Wireless Wildcard Action Flows

In [2], the author proposed a considerably useful approach called Wildcard identical action flows to simplify the data forwarding process. The fundamental idea is to implement a specified action on every flow. By adding the RouteHeader to each data packet, the router will find it much easier to forwarding the data to its destination. At the same time, the rule assigned to every router is almost the same command which is simple. The RouteHeader indicates the outgoing interface of the router to the next router in its path. This approach is designed for the wired networks. For Node i, the rule stored is:

$$Outgoing\_Interface = Number\_FirstHeader$$

We can apply it to wireless networks too. We use the Wireless Wildcard Action Flows, instead of marking the interface number as the RouteHeader. We employ the channel frequency and the ID number(e.g.MAC address) as the RouteHeader. The rules are changed to:

$$Outgoing\_Node = ID\_FirstHeader$$

$$Channel\_Frequency = Frequency\_FirstHeader$$

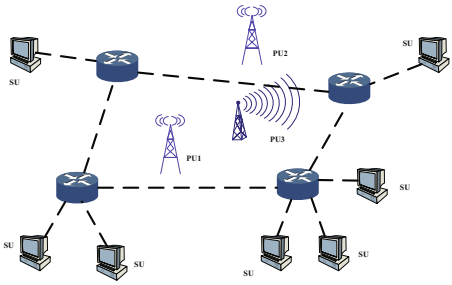


Fig. 3 Cognitive Radio based WMN

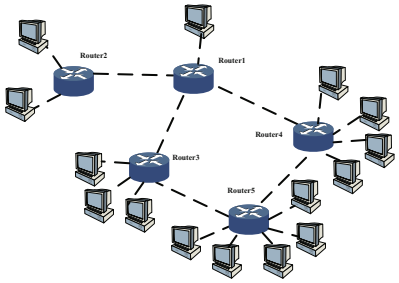


Fig. 4 Simulation Topology

Where $Frequency\_FirstHeader$ represents the frequency value picked out from the Frequency Flow Table. Here we simply consider the channel is ready for us. In the next section, we will explain in detail how we could classify the channel by the HDP model.

4 Problem Statement

The Control-Message Quenching scheme was first proposed in [3]. The goal of the scheme is to reduce the visiting times of the controller so that the controller is not overwhelmed. Now we consider an OpenFlow based wireless mesh network with one controller and N mesh routers. There are m_i mesh clients connected to the ith mesh router, where $1 \leq i \leq N$. The data traffic between the routers is expressed as:

$$\Lambda = \begin{bmatrix} \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,N} \\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,1} & \lambda_{N,2} & \cdots & \lambda_{N,N} \end{bmatrix} \tag{1}$$

Where $\lambda_{i,j}$ is the data arrive rate from $Router_i$ to $Router_j$ and $\lambda_{i,i} = 0$ since the data capacity within the same LAN can be large.

Assume that T_{c_i} is the controller processing time for a request coming from $Router_i$. Typically we have

$$T_{c\_i} >> \frac{1}{\lambda_{i,j}} \tag{2}$$

We know that the clients behind each router can generate data packets that are sent to different routers with probabilities. In order to simplify the analysis, we only consider the static probability. The probability from $Router_i$ to $Router_j$ is represented as

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,N} \end{bmatrix} \tag{3}$$

and with the constraints

$$\sum_{j=1}^{N} p(i, j) = 1 \tag{4}$$

For $Router_i$, during time T_{c_i}, the clients behind the router would still generate data and submit it to the router. The arrive rate, according to (1) and (3), will be

$$r_i = \sum_{j=1}^{N} \lambda_{i,j} p(i, j) \tag{5}$$

If these data packets have to be processed by the controller in order to get the routing path, during time T_{c_i}, the number of requests the controller will receive is $\frac{T_{c\_i}}{r_i}$. Substitute (5) into the expression, we get

$$\frac{T_{c\_i}}{r_i} = \frac{T_{c\_i}}{\sum_{j=1}^{N} \lambda_{i,j} p(i, j)} \tag{6}$$

The total number of requests received by the controller is

$$Total_{req} = \sum_{i=1}^{N} \frac{T_{c\_i}}{\sum_{j=1}^{N} \lambda_{i,j} p(i, j)} \tag{7}$$

This is a huge number. According to [5], in a network with average arrive $\lambda_{i,j} = 85.3k$ packet/sec, the messages received by the controller is about 2.7 Giga/second, which is far beyond the controller's capability to handle the messages.

The author in [2] proposed an algorithm to quench the control messages. The algorithm applies the recording and waiting scheme to the data plane devices. By establishing a table that contains the source and destination pair. The following messages that hit the record will not send the request again to the controller. The algorithm can reduce all the repeated requests thus reducing the burden of the controller significantly.

However, in cognitive radio based wireless mesh networks, the frequency and bandwidth are precious resources. We want to further quench the control messages to give the controller more opportunities to do something valuable.

5 Channel Selection and Quenching Algorithm

5.1 HDP Based Spectrum Access

In the previous sections we have known that Cognitive Radio(CR) technology could improve the spectrum utilization significantly. Recently, many researchers have proposed different kinds of models for spectrum sensing and accessing. Among those works, Xin-lin Huang in [4] proposed a Hierarchical Dirichlet Process(HDP) based spectrum access scheme. According to [4], the HDP is naturally fit for distributed spectrum sensing.

The 2nd-level DP is

$$G_0 = \sum_{k=1}^{+\infty} \beta_k \delta_{\widetilde{\lambda}^k}, \beta|\gamma \sim GEM(\gamma), \widetilde{\lambda}^k \sim H \tag{8}$$

The 1st-level DP:

$$G_j = \sum_{t=1}^{+\infty} \widetilde{\pi}_{jt} \delta_{\widetilde{\lambda}^{jt}}, \widetilde{\pi}_j|\alpha \sim GEM(\alpha), \widetilde{\lambda}^{jt} \sim G_0 \tag{9}$$

By applying the HDP model, the scheme could automatically sense the channel and classify the channel into a certain group. One of the major benefit of HDP model is that we don't have to indicate how many groups there. The model will update the group number as the probability indicated. The hidden parameter in this system is

$$\lambda^{ji}|G_j, X^{ji}|\lambda^{ji} \sim \prod_{k=W_H T/\pi}^{k=W_H T/\pi} Exponential(\lambda_k^{ji}) \tag{10}$$

The CR channel is recognized as a Rayleigh channel. After applying HDP model, we can find out the nodes which share the similar channel environment(i.e. channel frequency). One explanation of the HDP model is named Chinese Restaurant Franchise(CRF).

$$p(\lambda_i|\lambda_{-i}, \gamma, H) = \frac{\gamma}{\gamma + N - 1} H + \frac{1}{\gamma + N - 1} \sum_{k=1, k\neq i}^{N} \delta_{\lambda^k} \tag{11}$$

It shows that SUs can either be classified into one existed groups with probability $\frac{\gamma}{\gamma+N-1}$ or into a brand new group with probability $\frac{1}{\gamma+N-1}$. In the simulation section, we will show the results of the channel classification.

5.2 Quenching Algorithm

In the previous algorithm, only the repeated requests are quenched. We only record the source and destination pair in the local router for searching. Here we further explore the neighborhoods' resources in their flow tables. The basic idea of our proposed algorithm is that we can not only record the historical pair. When the new packets failed to hit the record. The router could automatically ask its neighbors for help. That is, one router could share its own flow table with its neighbors. Here are the rules involved in the idea.

Every router know its neighbors' ID and location.
The controller inform the router about the topology when there are changes in the network.
Each router can only ask for routing information from the nodes within one hop.
The items in the flow table in each router only exists for a certain period of time and will be killed by the router when the time is over.

Assume that each router maintains and updates a table list L. We denote a path from source(s) to destination(d) as $< s, d >$. Refer to Algorithm 1.

When we choose Algorithm 1, we could quench most of the redundant requests messages to controller. Thus saving plenty of controller's computational resources. As we mentioned in the previous sections, each router has the probability to visit

Algorithm 1 Control-Message Quenching(CMQ) with Information From Neighbors

```
L: = empty set
for each incoming packet do
    Check the destination node and mark it as d_com;
    Look up the flow table to find out whether there is a record matching
  the destination node and path < s,d >=< s,d_com >
    if Matched record found then
        if The path is ready then
            Handle the packet as the rule indicated
        else
            Wait for the path
        end if
    else
        Ask neighbors to find out the same path
        if Matched record found then
            Extract the rule
            Copy the rule and paste to local flow table
            Handle the packet as the rule indicated
        else
            Consult the controller for path configuration
            Record the destination node and update list L
        end if
    end if
end for
if Receive an answer from the controller then
    Record the path
    Record the node ID along the path
    Update list L
end if
if Item time expires then
    Kill item
end if
```

Fig. 5

Algorithm 2 A Function added to Controller

```
if Request from i to j arrived then
    Ask routers for path
    if Path found from k to j then
        ComputePath(i,k)
    else
        ComputePath(i,j)
    end if
end if
```

Fig. 6

all other routers, when the router has the opportunity to get information from its neighbors, the probability of hitting the record becomes much higher. For $Router_i$, assume that the neighbors of $Router_i$ form the set S. The hitting probability to $Router_j$ is

$$p_{hit(i,j)} = p_{i,j} + \sum_{l \in S} p_{l,j} \tag{12}$$

In this way, the network would have better throughput and delay performance than the previous algorithms. In the simulation section, we will show the simulation results of the algorithm.

5.3 *Enhanced CMQ Algorithm*

The previous algorithm we have proposed do improve the overall performance of the network. It also reduces the burden of the controller. From equation (12) we know the reason is that we increase the hitting probability. However, there is still some drawbacks with this algorithm. If we want the packets to go through the shortest path, we still have to ask the controller for help. Another issue is that in some cases the controller figures out a long path that goes through some routers. When some of the routers in the middle of the long path have data to send out, it has to ask the controller again for path configuration. This is also repeated work for the controller. In order to avoid this issue, we proposed an enhanced CMQ algorithm. Based on Algorithm 1, when the packets go to a remote router, it can bring the path destination information to the routers within the path. As the packet goes from one router to another, the destination and path would be stored in the routers that the packet passed by. In this way, the routers in the same path all have the information to the destination. Meanwhile, the routers have higher probability to go through the shortest path than our previous proposed algorithm does. The hitting rate goes higher.

For the controller, it should have a function to ask for the routers for the paths. Compare to the time consumed in path computation and configuration, Asking routers to check if there is existed path costs far less time and energy. The processing detail is illustrated in the table Algorithm 2.

In the next section, we will provide the simulation results for the algorithms described above.

6 Numerical Simulation

In this section, simulation result are provided. We first illustrate the simulation parameters, and then show the figures.

6.1 Simulation Parameters

The simulation was designed according to the previous proposed scenario. Shown in Fig. 4, the network topology includes one controller, five mesh routers and 15 mesh clients. All the connections between the devices are wireless. The controller manages the whole topology. Each mesh router are marked by a number from 1 to 5. For router i, we simply add i clients behind it. Each client could contact any of the remaining in the whole network. The channel bandwidth between the Controller and Routers is set to 20Mbps. Generally, the data traffic between Controller and Routers is far less than that between routers and clients. The data link capacity between one router and another is 50Mbps. For simplicity, all the clients are both senders and receivers. One client could transmit data to another one randomly with equal probability. The size of one data packet is 1500Byets, which is the same as the size of IP packet in Internet. The processing time of the controller is 10 times the time of one packet transmission.

Three indicators are used for comparison, i.e. the average network throughput, average packet delay and the number of requests to the controller.

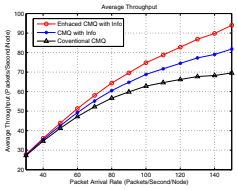


Fig. 7 Throughput Performance of Different Schemes

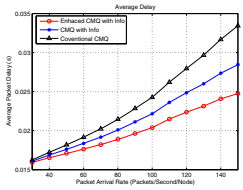


Fig. 8 Average Packet Delay of Different Schemes

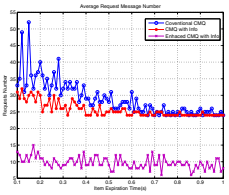


Fig. 9 Number of Requests sent to Controller

6.2 Simulation Results

Figure 7 shows the simulation result of the average network throughput under different schemes. It's obvious that the CMQ algorithm with information from neighbors outperforms that of the conventional CMQ algorithm. The enhanced CMQ algorithms with information from neighbors works even better. With the increasing data generation rate, the throughput is also rising. When the average input data arrival rate is larger than 130 packets/second, there is a ceiling for the conventional CMQ algorithm. However, with information from neighbors, the throughput could continue increasing.

Figure 8 presents the average packet delay in the network. As the data generate rate is increasing, the network is becoming crowded. Still, our proposed algorithms work better than the conventional CMQ algorithm. The packet delay of the proposed algorithms are lower. The delay of the conventional CMQ increases faster.

In Fig. 9, we simulate the number of requests to controller with respect to the expiration time of the source-destination pair records. As we can see from the figure, the number of requests tent to be a fixed value as the holding time of the items in flow table increases. This is because when the holding time is long, more flows are getting easily to hit the record in the flow table. Thus less requests the controller will receive. From the result we find that the number of our proposed algorithms decreases faster. In addition, when we apply the enhanced CMQ algorithm, the minimum number is less than other algorithm. The number changes slightly. In this way, we can set the expiration time to a low level to save more TCAMs in data plane devices.

7 Conclusion

In this paper, we first introduce an OpenFlow based Wireless Mesh Network system. Some of the major challenging issues are presented. An HDP model is introduced to sense and classify the channel for WMN. We also modified the wildcard rule to make it useful in wireless systems. Then, we analyze the problem and proposed an advanced CMQ algorithm to quench the number of requests to the controller. Simulation results show that our proposed algorithms work better compared to the conventional one.

References

1. McKeon, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: Enabling Innovation in Campus Networks. ACM SIGCOMM Computer Communication Review, April 2008
2. Iyer, A.S., Mann, V., Samineni, N.R.: SwitchReduce: reducing switch state and controller involvement in OpenFlow networks. In: IFIP Networking Conference, 2013, pp. 1–9, May 22–24, 2013
3. Luo, T., et al.: Enhancing responsiveness and scalability for OpenFlow networks via control-message quenching. In: 2012 International Conference on ICT Convergence (ICTC). IEEE (2012)
4. Huang, X., Hu, F., Wu, J.: Intelligent Cooperative Spectrum Sensing via Hierarchical Dirichlet Networks. IEEE Journal on Selected Areas in Communication
5. Curtis, A.R., Mogul, J.C., Tourrilhes, J., Yalagandula, P., Sharma, P., Banerjee, S.: DevoFlow: Scaling flow management for high-performance networks. ACM SIGCOMM (2011)
6. Open Networking Foundation. OpenFlow switch specification, April 16, 2012
7. Luo, T., Tan, H.-P., Quek, T.Q.S.: Sensor OpenFlow: Enabling software-defined wireless sensor networks. IEEE Communications Letters (2012, to appear)
8. Open Networking Foundation. Software-defined networking: The new norm for networks, April 2012 (white paper)
9. Yu, M., Rexford, J., Freedman, M.J., Wang, J.: Scalable flow-based networking with DIFANE. ACM SIGCOMM (2010)

10. Gude, N., Koponen, T., Pettit, J., Pfaff, B., Casado, M., McKeown, N., Shenker, S.: NOX: Towards an operating system for networks. SIGCOMM Comput. Commun. Rev. **38**(3), 105–110 (2008)
11. Koponen, T., Casado, M., Gude, N., Stribling, J., Poutievski, L., Zhu, M., Ramanathan, R., Iwata, Y., Inoue, H., Hama, T., Shenker, S.: Onix: a distributed control platform for largescale production networks. In: The 9th USENIX Conference on Operating Systems Design and Implementation (OSDI), pp. 1–6 (2010)
12. Tootoonchian, A., Ganjali, Y.: HyperFlow: a distributed control plane for openflow. In: INM/WREN. USENIX Association (2010)

SINR Maximization in Relay-Assisted Multi-user Wireless Networks

Umar Rashid, Faheem Gohar Awan and Muhammad Kamran

Abstract This paper considers throughput maximization in relays based multi-user wireless networks by enhancing the worst signal-to-interference-plus-noise ratio (SINR) among multiple users. Unlike the existing approaches that use semidefinite relaxation coupled with Gaussian randomization (SDR-G), we utilize the d.c. (difference of two convex functions) structure of the resulting objective function to develop efficient iterative algorithms of low complexity. Numerical results demonstrate that the proposed algorithm locates solutions that are close to the upper bound by a few iterations, and hence, shows better performance than the other methods.

Keywords Beamforming · Wireless relay networks · Interference

1 Introduction

Relay-assisted wireless communication is a very active research topic (see e.g. [1, 2]). Using the spatial diversity of relays, it is possible to expand the range of communication [3]. For multi-user communication, wireless relay networks provide assistance in the communication between multiple sources and destination nodes [4]. To strengthen the signal of interest for a user at the destination, and suppress interferences and noise, a beamforming vector is designed at relay nodes [5].

Most of the previous works on relay beamforming in multiuser systems considered minimization of power consumed by relay nodes under constraints on individual

Umar Rashid—Umar Rashid is also with Alkhwarizmi Institute of Computer Science (KICS) UET, Lahore.

U. Rashid(✉) · F.G. Awan · M. Kamran
Department of Electrical Engineering, University of Engineering and Technology, KSK Campus, Lahore, Pakistan
e-mail: {umar.rashid,fawan,mkamran}@uet.edu.pk,
http://www.uet.edu.pk

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_8

signal-to-interference-plus-noise ratio (SINR). This nonconvex optimization problem is then transformed into a relaxed semi-definite program (SDP) [5, 6]. However, due to the individual relay power constraints, the relaxed solution no longer remains a rank-one, and thus remains sub-optimal in terms of performance. In order to further improve the solution, a technique known as randomization [5, 6] is utilized to search for an optimal beamformer vector candidate with appropriate scaling.

One of the key characteristics of the state-of-the-art wireless networks is that they must provide better QoS at a high data rate, expressed via high SINR, under limited power resources. Such requirements lead to a beamformer vector that when applied to the relay nodes provide maximin SINR among the given users. However, such an optimization is a highly non-convex problem. The existing approach is primarily based on using a matrix variable to reformulate the original problem as an SDP that rather increases the size of the problem by many folds.

The major contribution of this correspondence is to use d.c. programming [7] to express the objective function as the difference of two convex functions. Secondly, the concave part of the objective function is convexified at some given point which results into an iteratively decreasing program. Under some given convergence criteria, a suboptimal solution is obtained after a few iterations. In one of our previous works, we have considered a similar problem but in a robust channel environment [8].

The rest of the paper is structured as follows. Section 2 describes the system model and the beamforming problem. Section 3 reformulates the beamforming optimization as a d.c. objective function with convex constraints. Subsequently, with effective convexification an iterative algorithm is suggested to obtain the solutions. Simulation results that demonstrate the effectiveness od the algorithm are given in Section 4. Concluding remarks are described in Section 5.

Notations: Boldfaced uppercase and lowercase characters are used to denote matrices and column vectors. A positive semi-definite matrix (Hermitian) is represented as $\boldsymbol{A} \geq 0$. We define $\langle \boldsymbol{a}, \boldsymbol{b} \rangle := \boldsymbol{a}^T \boldsymbol{b}$ and $||\boldsymbol{a}||^2 := \langle \bar{\boldsymbol{a}}, \boldsymbol{a} \rangle$. Notation for the Hadamard is $\boldsymbol{a} \odot \boldsymbol{b}$.

2 System Model and Problem Formulation

Consider a multiuser system consisting of M source-destination pairs communicating with the help of N relay nodes as shown in Figure 1. While operating in half-duplex mode, two time-slots are used by the relay and user nodes for communication. In the first time-slot, the source send their signal to the relays. Let $\boldsymbol{s} \in \mathcal{C}^M$ be the signals transmitted by M source nodes. The signal is zero mean with variance $\sigma_s = \mathbb{E}[|s_i|^2]$. For $m = 1, 2, ..., M$, we define

$$\boldsymbol{h}_m = (h_{m1}, h_{m2}, ..., h_{mN})^T \in \mathcal{C}^N, \tag{1}$$

as the backward channel vector between the m^{th} source and relay nodes. Similarly, for the forward channel between the relay network and he i^{th} destination we set

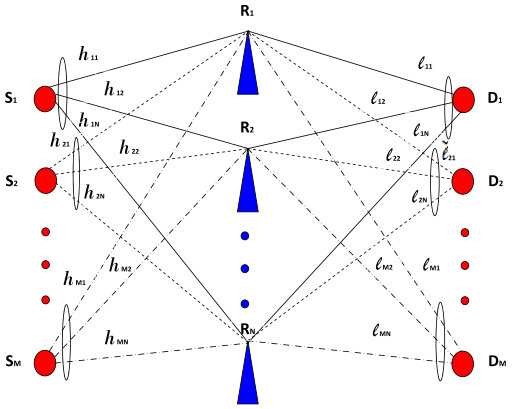


Fig. 1 A multi-user wireless relay network.

$$\boldsymbol{\ell}_i = (\ell_{i1}, \ell_{i2}, ..., \ell_{iN})^T \in \mathcal{C}^N, \ i = 1, 2, ..., M \tag{2}$$

The relays receive the following signal

$$\boldsymbol{y}_{up} = \sum_{m=1}^{M} \boldsymbol{h}_m s_m + \boldsymbol{n}_r, \tag{3}$$

where $\boldsymbol{n}_r \sim \mathcal{N}(0, \sigma_r^2 \boldsymbol{I}_N)$ denotes the additive white Gaussian noises at the relay receivers. Suppose $\boldsymbol{x} = (x_1, x_2, ..., x_N)^T$ are the beamforming weights. After multiplying with these weights, the signals forwarded by the relays to the destinations take the form

$$\boldsymbol{y}_{amp} = \boldsymbol{x} \odot \boldsymbol{y}_{up} = \sum_{m=1}^{M} \boldsymbol{x} \odot \boldsymbol{h}_m s_m + \boldsymbol{x} \odot \boldsymbol{n}_r. \tag{4}$$

Subsequently, the signal received by the destination node i is

$$\begin{aligned} \boldsymbol{y}_{di} &= \langle \boldsymbol{\ell}_i, \boldsymbol{y}_{amp} \rangle + n_{di} \\ &= \sum_{m=1}^{M} \langle \boldsymbol{c}_{mi}, \boldsymbol{x} \rangle s_m + \langle \boldsymbol{\ell}_i \odot \boldsymbol{n}_r, \boldsymbol{x} \rangle + n_{di}, \end{aligned} \tag{5}$$

where $n_{di} \sim \mathcal{N}(0, \sigma_d^2)$, and $\boldsymbol{c}_{mi} = \boldsymbol{\ell}_i \odot \boldsymbol{h}_m$ is the compounded channel vector between source m and destination i.

After defining system parameters, we can now express the signal-to-interference-plus-noise-ratio (SINR) at the i^{th} destination as

$$\boldsymbol{SINR}_i(\boldsymbol{x}) = \frac{\sigma_s^2 |\langle \boldsymbol{x}, \boldsymbol{c}_{ii} \rangle|^2}{\sigma_s^2 \sum_{m \neq i} |\langle \boldsymbol{x}, \boldsymbol{c}_{mi} \rangle|^2 + \boldsymbol{x}^H \boldsymbol{L}_i \boldsymbol{x} + \sigma_d^2}. \tag{6}$$

where $\boldsymbol{L}_i = \sigma_r^2 \text{diag}([|\ell_{i,1}|^2, \ldots, |\ell_{i,N}|^2])$. The beamforming power consumed by the n^{th} relay node is $\boldsymbol{P}_n(x_n) = r_n|x_n|^2$ where $r_n = \sigma_s^2 \sum_{m=1}^{M} |h_{mn}|^2 + \sigma_r^2, n = 1, 2, ..., N$.

Hence, SINR's threshold maximization is given as,

$$\begin{aligned} &\max_{\boldsymbol{x}\in\mathcal{C}^N} \min_{i=1,2,...,M} \boldsymbol{SINR}_i(\boldsymbol{x}) \\ &\text{s.t. } r_n|x_n|^2 \le \gamma_n,\ n = 1, 2, ..., N, \end{aligned} \tag{7}$$

which has a nonconvex objective function, and thus difficult to solve.

3 D.C. Programming Based Solution

This section reformulates the original beamforming design problem (7) as a d.c. optimization problem. It can be noticed that (7) is equivalent to

$$\max_{\boldsymbol{x}\in\mathcal{C}^N, \boldsymbol{y}\in R_+^M} \min_{i=1,2,...,M} \varphi_i(\boldsymbol{x}, y_i) := \frac{|\langle \boldsymbol{x}, \boldsymbol{c}_{ii}\rangle|^2}{y_i + \sigma_d^2/\sigma_s^2} \ : \tag{8a}$$

$$\sum_{m\neq i} |\langle \boldsymbol{x}, \boldsymbol{c}_{mi}\rangle|^2 + \frac{1}{\sigma_s^2}\boldsymbol{x}^H \boldsymbol{L}_i \boldsymbol{x} \le y_i,\ i = 1, 2, ..., M \tag{8b}$$

$$r_n|x_n|^2 \le \gamma_n,\ n = 1, 2, ..., N. \tag{8c}$$

It can be proved that each fractional function $|\langle \boldsymbol{x}, \boldsymbol{c}_{ii}\rangle|^2/(y_i + \sigma_{\text{D}}^2/\sigma_s^2)$ is convex (for proof see [9]). Now, although each $\varphi_i(\boldsymbol{x}, y_i)$ is convex, their minimum $\min_{i=1,2,...,M} \varphi(\boldsymbol{x}, y_i)$ is not concave and (8) is not a convex program. However, using decomposition [7]

$$\min_{i=1} \varphi_i(\boldsymbol{x}, y_i) = \sum_{i=1}^{M} \varphi_i(\boldsymbol{x}, y_i) - \max_{i=1,2,...,M} \sum_{j\neq i} \varphi_j(\boldsymbol{x}, y_j)$$

we see that (8) is the following d.c. program

$$-\min_{\boldsymbol{x}\in\mathcal{C}^N, \boldsymbol{y}\in R_+^M} [f_{01}(\boldsymbol{x}, \boldsymbol{y}) - f_{02}(\boldsymbol{x}, \boldsymbol{y})] \ : \ (8b) \tag{9}$$

with the convex functions

$$f_{01}(\boldsymbol{x}, \boldsymbol{y}) := \max_{i=1,2,...M} \sum_{j\neq i} \varphi_j(\boldsymbol{x}, y_j),\ \ f_{02}(\boldsymbol{x}, \boldsymbol{y}) := \sum_{i=1}^{M} \varphi_i(\boldsymbol{x}, y_i), \tag{10}$$

as maximum of convex function and as sum of convex function. Hence, we obtain the following iterative solution

$$\min_{\boldsymbol{x}\in\mathcal{C}^N,\boldsymbol{y}\in R_+^M}[f_{01}(\boldsymbol{x},\boldsymbol{y})-f_{02}(\boldsymbol{x}^{(\kappa)},\boldsymbol{y}^{(\kappa)})- \sum_{i=1}^{M}\langle\nabla\varphi_i(\boldsymbol{x}^{(\kappa)},y_i^{(\kappa)}),(\boldsymbol{x},y_i)-(\boldsymbol{x}^{(\kappa)},y_i^{(\kappa)})\rangle]:(8b) \tag{11}$$

where

$$\langle\nabla\varphi_i(\boldsymbol{x}^{(\kappa)},y_i^{(\kappa)}),(\boldsymbol{x},y_i)-(\boldsymbol{x}^{(\kappa)},y_i^{(\kappa)})\rangle = \frac{2\mathrm{Re}(\overline{\langle\boldsymbol{x}^{(\kappa)},\boldsymbol{c}_{ii}\rangle}.\langle\boldsymbol{c}_{ii},\boldsymbol{x}-\boldsymbol{x}^{(\kappa)}\rangle)}{y_i^{(\kappa)}+\sigma_d^2/\sigma_s^2}-\frac{|\langle\boldsymbol{x}^{(\kappa)},\boldsymbol{c}_{ii}\rangle|^2(y_i-y_i^{(\kappa)})}{(y_i^{(\kappa)}+\sigma_d^2/\sigma_s^2)^2}.$$

The detailed procedure of the above algorithm has been described in [8].

4 Numerical Results

For simulations, the powers of backward and forward noises are normalized to $\sigma_{\mathrm{R}}^2=\sigma_{\mathrm{D}}^2=1$. On the other hand, all the sources have the same signal power $\sigma_s^2=100$. Generation of the forward and backward channels follows circularly symmetric complex Gaussian distribution in the simulation settings. In order to execute the proposed DCI method, first we solve the following convex program for $n\in\{1,2,\ldots,N\}$ with any $10\log_{10}(\alpha_i)>0$ dB

$$\begin{aligned}&\min_{\boldsymbol{X}\in\mathcal{C}^{N\times N}}\max\boldsymbol{X}_{n,n}\quad\text{s.t.}\quad\boldsymbol{X}\succeq 0,\\&\sigma_s^2\langle\boldsymbol{C}_{ii},\boldsymbol{X}\rangle\geq\alpha_i\left(\sigma_s^2\sum_{m\neq i}^{M}\langle\boldsymbol{C}_{mi},\boldsymbol{X}\rangle+\sigma_{\mathrm{R}}^2\langle\boldsymbol{L}_i,\boldsymbol{X}\rangle+\sigma_{\mathrm{D}}^2\right)\end{aligned} \tag{12}$$

to obtain the initial $(\boldsymbol{x}^{(0)},\boldsymbol{y}^{(0)})$ feasible point as explained in [8].

Total relay power is assumed to be equally divided among the relay nodes. We have plotted the minimum SINR among all users achieved by different approaches. The upper bound for the SINR performance can be obtained by solving the relaxed SDPs as explained in [5, 6]. In every figure, the DCI method performs well enough to almost overlap the upper bound of the SDR curve which shows the effectiveness of our proposed method.

Figure 2 shows that with our proposed DCI approach achieves an SINR of 6 for each user when $P_T=10$ dB. On the other hand, the SINR obtained by SDP randomization only gives 3 dB for the same amount of total relay power. Moreover, the fair distribution of SINR among all users is also not achievable by the randomization technique.

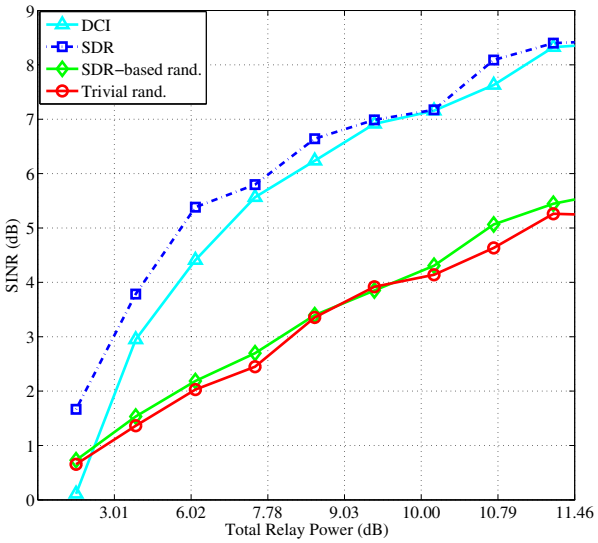


Fig. 2 Minimum SINR plotted against relay power when $M = 3$ and $N = 10$.

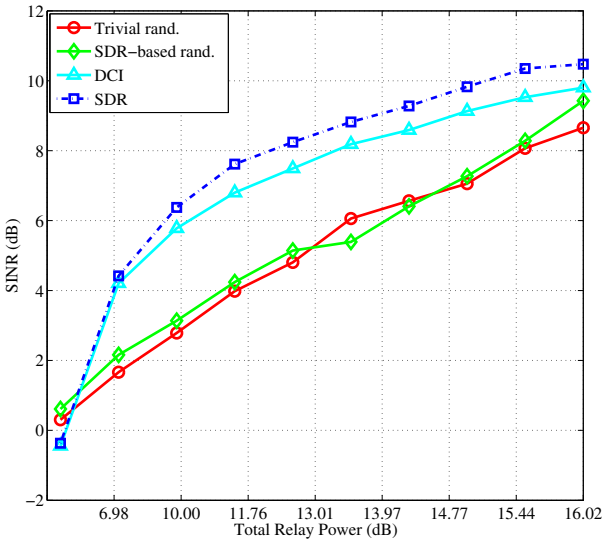


Fig. 3 Minimum SINR plotted against relay power when $M = 4$ and $N = 16$.

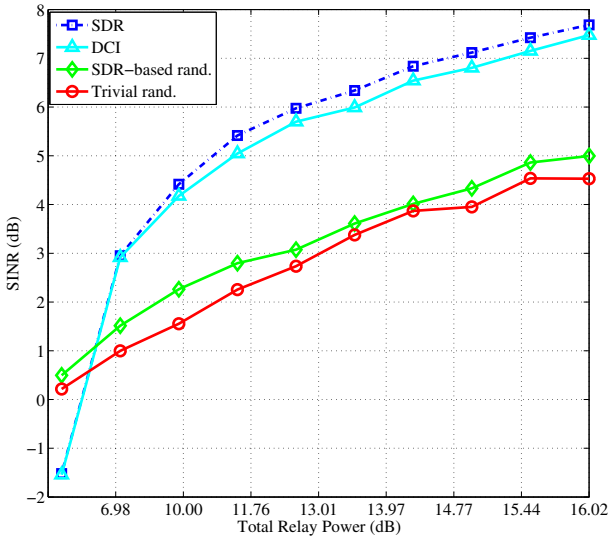


Fig. 4 Minimum SINR plotted against relay power when $M = 5$ and $N = 20$.

Table 1 Numbers of iterations for convergence

$(M, N) = (3, 10)$		$(M, N) = (4, 16)$		$(M, N) = (5, 20)$	
P_T (dB)	Iterations	P_T (dB)	Iterations	P_T (dB)	Iterations
0.04	13.57	0.04	21.03	-10	23.71
4.08	13.25	7.27	20.39	7.51	21.88
6.14	12.62	9.85	20.16	10.48	21.20
7.53	12.11	11.46	18.77	12.23	20.25
8.58	12.23	12.63	18.67	13.48	19.08
9.43	11.33	13.55	17.60	14.44	17.99
10.14	10.15	14.31	16.59	15.27	17.07
10.75	9.16	14.96	15.73	15.90	14.83
11.28	7.95	15.52	12.69	16.48	12.72
11.76	6.11	16.02	7.48	16.99	8.84

Figures 4 and 3 also analyze the impact of increasing number of relay nodes from $N = 12$ to $N = 16$ for a fixed number of users $M = 4$. As indicated, the individual SINR increases when more relays are utilized.

Table 1 presents the iterative convergence performance by the proposed DCI method. Table 2 gives the averaged value of the rank of the sub-optimal matrix-based solution $\boldsymbol{X}_{\text{opt}}$ of SDP [5, 6] at α_{opt}. As it can be seen that due to the high rank of the relaxed SDP based solution, it is not possible to guarantee the optimality. Furthermore, In terms of performance, this randomization is nearly same as a simple trivial randomization achieved by solving the SDP program for a feasible solution $\boldsymbol{X}$ of SDP [5, 6].

Table 2 Average rank i_{opt} of $\boldsymbol{X}_{\text{opt}}$ by SDP relaxation

$(M, N) = (3, 10)$		$(M, N) = (4, 16)$		$(M, N) = (5, 20)$	
P_T (dB)	Avg. i_{opt}	P_T (dB)	Avg. i_{opt}	P_T (dB)	Avg. i_{opt}
0.04	2.18	0.04	2.55	0.04	2.55
4.08	2.09	7.27	2.52	7.27	2.58
6.14	2.11	9.85	2.47	9.85	2.62
7.53	2.11	11.46	2.50	11.46	2.64
8.58	2.11	12.63	2.52	12.63	2.62
9.43	2.14	13.55	2.65	13.55	2.62
10.14	2.11	14.31	2.46	14.31	2.57
10.75	2.14	14.96	2.49	14.96	2.63
11.28	2.15	15.52	2.54	15.52	2.59
11.76	2.14	16.02	2.52	16.02	2.55

5 Conclusion

This paper has presented an efficient algorithm for the beamforming design problem while maximizing the worst SINR when multiple users communicate with the assistance of relay nodes. The proposed approach reformulates the original nonconvex problem as a low-dimension d.c. program and then develop an efficient iterative procedure to obtain a local optimal solution. Numerical results demonstrate that the developed algorithms perform much better than the existing relaxed SDP based solutions in terms of the throughput.

References

1. Fan, L., Lei, X., Duong, T.Q., Elkashlan, M., Karagiannidis, G.K.: Secure multiuser communications in multiple amplify-and-forward relay networks. IEEE Trans. on Comm. **62**, 3299–3310 (2014)
2. Chatzipanagiotis, N., Liu, Y., Petropulu, A., Zavlanos, M.: Distributed cooperative beamforming in multi-source multi-destination clustered systems. IEEE Trans. on Signal Processing. **62**, 6105–6117 (2014)
3. Jing, Y., Jafarkhani, H.: Single and Multiple Relay Selection Schemes and Their Achievable Diversity Orders. IEEE Trans. on Wireless Comm. **8**, 1414–1423 (2009)
4. Soysa, M., Suraweera, H., Tellambura, C., Garg, H.K.: Multiuser amplify-and-forward relaying with delayed feedback in Nakagami-m fading. In: IEEE Wireless Communications and Networking Conference (WCNC), pp. 1724–1729 (2011)
5. Fazeli-Dehkordy, S., Shahbazpanahi, S.: Multiple peer-to-peer communications using a network of relays. IEEE Trans. on Signal Processing. **57**, 3053–3062 (2009)
6. Hsieh, P., Lin, Y., Chen, S.G.: Robust distributed beamforming design in amplify-and-forward relay systems with multiple user pairs. In: 23rd IEEE International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 371–375 (2015)
7. Tuy, H.: Convex Analysis and Global Optimization. Kluwer Academic Publishers (1998)
8. Rashid, U., Tuan, H.D., Nguyen, H.: Relay Beamforming Designs in Multi-user Wireless Relay Networks based on Throughput Maximin Optimization. IEEE Trans. on Comm. **99**, 1739–1749 (2013)
9. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)

A Hybrid MAC for Long-Distance Mesh Network with Multi-beam Antennas

Xin Li, Fei Hu, Ji Qi and Sunil Kumar

Abstract In recent years, with the development of multi-beam smart antennas (MBSA), directional mesh networks is becoming more and more popular. With the popularity of UAVs and environment surveillance applications, airborne networks (ANs) have become important platforms for wireless transmissions in the sky. In this work we propose a hybrid MAC scheme for a hierarchical airborne network, which consists of high-speed, long-link, multi-beam aircraft nodes (in the higher level) and short-distance, high-density UAVs (in the lower level). Simulation results show that compared to existing 802.11 MAC schemes, our MAC has better performance in terms of network throughput and packet delay.

Keywords Airborne Mesh Networks · Multi-beam antennas · Heterogeneous MAC · Ku-band

1 Introduction

1.1 Two-Level Airborne Mesh Networks

In recent years, with the development of multi-beam smart antennas (MBSA), directional airborne network is becoming more and more popular in the sky. Aircrafts could be equipped with MBSAs instead of omni-directional antennas currently in use. With the help of MBSAs, each airplane in the airborne network would be able to communicate with its adjacent neighbors simultaneously in different beams. This would considerably improve the overall throughput of the network compared to

X. Li(✉) · F. Hu · J. Qi
Electrical and Computer Engineering, University of Alabama,Tuscaloosa, AL, USA
e-mail: xli120@crimson.ua.edu

S. Kumar
Electrical and Computer Engineering, SDSU, San Diego, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_9

that in networks with omni-directional antennas. In practice, a number of UAVs (unmanned aerial vehicles) usually work within a specific region-of-interest (RoI). The distance between UAVs is roughly within a range from 500m to 10km. The flying height of UAVs is much lower than that of aircrafts. Thus the UAVs form a *low level* airborne network which has a large number of nodes working in a small space. Typically an UAV is equipped with both omni-directional antenna and single-beam directional antenna. The omni-directional antenna is used to carrier-sense signals from surrounding nodes while the single-beam directional antenna is used to send out information. However, the aircrafts distribution in the *high level* is sparse. The average distance between these aircrafts may be longer than 100km. Each of these aircrafts is equipped with a MBSA. The high level airplanes are used to manage the RoIs. One of the UAVs in each RoI serves as the UAV head which could communicate with the corresponding aircraft. Figure 1 shows an illustration of hierarchical airborne mesh network. The figure shows that the airborne network has a typical wireless mesh network (WMN) architecture. The aircrafts in upper level form a wireless backbone network. The nodes in backbone network are named mesh routers (MR). In this case, the MR nodes are considerably powerful and sparse. While the UAVs in the lower level serve as mesh clients (MC).

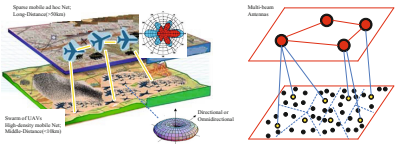


Fig. 1 An Illustration of Hierarchical Airborne Mesh Network.

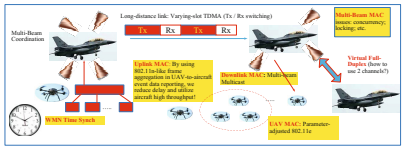


Fig. 2 Big Picture: Proposed heterogeneous MAC

Ku-band[1] spectrum, usually from 10GHz to 15GHz, was reserved for satellite communication. According to Shannon's equation, high frequency makes fast data rate. Ku-bands also provide better anti-interference capability, which means that the receiver node suffers less interference from other nodes. In addition, Ku-band signals are very sensitive to line-of-sight (LoS) blocking. Fortunately, airborne network nodes are set up in very high height with low blocking chance.

In this research, we target the MAC design in hierarchical airborne mesh network (HAMN). We propose an integrated, hybrid MAC for HAMN. There are three different types of links in a HAMN: Aircraft-to-aircraft (A2A), aircraft-to-UAVs (A2U) and UAV-to-UAV (U2U). For A2A links, we use a scheduled, TDMA-like multi-beam oriented MAC scheme to achieve high throughput transmissions. However, for U2U communications, we use the conventional CSMA/CA scheme with omni-directional antennas. The A2U links have a heterogeneous feature since these links connect the two network layers.

The rest of this paper is organized as follows: In Section 2, we first summarize the related work on MAC layer. In Section 3, we describe the system assumptions and then present the technical details of proposed heterogeneous MAC scheme.

Numerical simulation results will be shown in Section 4. Section 5 concludes the entire paper.

2 Literature Review

In the recent a few years, there has been plenty of research work on directional antenna oriented MAC for airborne networks [2]. Those MAC protocols are designed for real-time, high-throughput, low-delay transmissions.

The MAC used in long-distance aircraft communications is first mentioned in [3]. Most conventional 802.11 protocols are based on CSMA scheme. But the network based on 802.11 MAC has certain limitations. For example, it is only suitable to low-distance links (<500m). Now with the usage of new frequency such as Ku-band ($\sim$ 15GHz)[4],the long distance network becomes a reality. With Ku-band frequency, wireless communication links could have more reliability in transmissions. In practice the BER (bit error rate) is below 10^{-5}. Ku-band signals are less likely to scatter. Thus they are less likely to be interrupted by the noise.

There has been some work focusing on directional MAC protocols [5]. Unfortunately, the single-beam antenna model is considered in most of these studies. Some work did consider multi-beam MAC design. In [6], V. Jain proposed a hybrid MAC scheme in multi-beam networks. An enhanced point coordination function (PCF) is proposed in [7]. A distributed CSMA-based scheme is considered for multi-beam communication in [8].

In our work, we propose a new hierarchical multi-beam MAC protocol which is able to deal with the critical issues such as long distance, node deafness, sender polling, etc. The simulation results show that our proposed scheme could fully exploit the benefits of MBSA.

3 Heterogeneous MAC Design

Figure 2 shows the big picture of our MAC design. The distance between two nodes determines the performance boundary of CSMA-based and TDMA-like MAC. The random access scheme such as CSMA does not work well if the distance is too long. When the distance is larger than 50km, the round trip delay is 0.33ms. It is difficult to detect radio signal collision for such a long time. In addition, the long ACK timeout makes it difficult to achieve high throughput since it wastes too much time on ACK waiting. In this scenario, CSMA severely sacrifices the network throughput. While time-scheduled MAC protocols (TDMA-based) is more suitable to such a long propagation delay since it can use a dedicated time duration for data transmission.

In multi-beam antenna MAC design, our goal is to fully explore the multi-beam antennas to improve the network throughput. Theoretically, we could achieve N times

of throughput improvement over a single-link case if each antenna has N beams. Compared to one-beam transmission, multi-beam transmission can take advantage of neighbors′ relay. We can use more neighboring nodes to help forward data. Thus a high throughput is achieved via multiple path. Below we would describe MAC operations in A2A and U2U links.

3.1 A2A MAC Design

Our A2A MAC is based on TDMA-like scheme. Each node is entitled with a number of time slots to send or receive data packets. The node with the right to transmit data is named as 'star' node. The slots could be assigned according to deterministic or statistical rules. The entire aircraft is synchronized based on the clock synchronization scheme [9].

Since the upper level aircraft network is sparse, it it easy to manage the slot allocation through a satellite-based global management. In our work, A2A MAC is different from pure TDMA-MAC in the following two aspects:

(1) Each phase is longer than that in conventional TDMA. As we know, the time duration in conventional TDMA is set to only hundreds of microseconds (μs). In our A2A links the time in each phase could be hundreds of milliseconds (ms) in order to deal with a large number of data packets.

(2)While in conventional TDMA-based MAC, each node is allocated with a fixed-length of time slots, in our A2A MAC the length of time slots is variable. Since it may cause much delay if the multi-beam antenna frequently switches between Rx and Tx modes, it is better to finish a window of packets in one slot (i.e., one Tx/Rx phase). The window size should be proportional to the traffic load in the node.

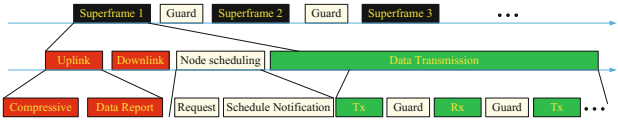


Fig. 3 Operation Phases in Each Aircraft

Figure 3 shows operation phases in each aircraft. As we can see, each superframe consists of three phases.

(1)Uplink/Downlink phase: In this phase, the node needs to determine the direction of the communication. For the uplink communication, the RoI nodes would first report their transmission durations to the aircraft through a compressive sensing based polling method. Then the center aircraft receives data packets from UAVs. For the downlink communication, the aircraft has to broadcast information to UAVs in its RoI.

(2)Node scheduling phase: In HAMN, if one aircraft wants to send messages to other nodes, it first sends request messages to a gateway node. which would determine the transmission order based on a hash function with node ID and times tamp:

$$Position(i) = Hash(ID(i), timestamp), i = 1, 2, ..., N \tag{1}$$

The value of the Hash function is normalized between $(0, 1]$. The variable $timestamp$ is the current time. A node with $ID = J$ wins the time slot only if the following equation holds true:

$$arg \max_{1 \le x \le b} Hash(ID(i), timestamp) = J \tag{2}$$

The Hash function is performed by each node. In this way, each node knows who is the current 'star' node and who will be the next. The Hash function is carefully chosen so that nodes have an equal chance to be the 'star' node.

(3)Data transmission phase: After the first two phases, the nodes enter scheduled communications. In this phase, we use a token-based [10] pipelined scheduling. Figure 4 shows an example of token handle process.

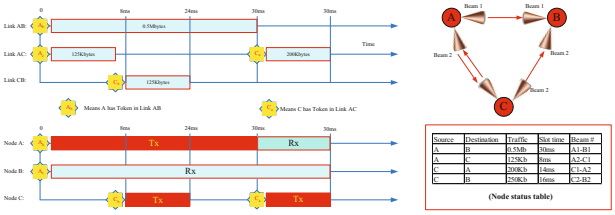


Fig. 4 Multi-beam Scheduled Transmission Process

As we can see from Fig. 4, although A cannot switch to Rx mode due to the longer Tx time in A-B link, it allows C to enter Tx mode after A-C transmission is done. Therefore, C can start to transmit data to B after 8ms. Such a scheme makes our MAC efficiently utilize each free link.

In addition, note that there is a node status table (shown in Fig. 4) which maintains the information on traffic amount in each beam of a node. In this example only two beams in each node are active. The beam status table is used to discover link quality and choose the proper beam to send out data packets. The gateway does not need to broadcast such a table to each node. It just needs to tell a node about its specific Tx /Rx timing information as well as the MAC address of its destination (when in Tx mode) or source (when in Rx mode).

3.2 U2U MAC Design

Compared to upper-level aircraft network, the UAV network has much higher node density. If we apply TMDA-based MAC in UAV network, we need a global coordinator as well as accurate timing synchronization among many UAVs. It is considerably difficult to manage such a global coordinator. Furthermore, using dedicated time slots could waste much bandwidth since the communications between UAVs are sparse. In most of the time, data transmission occurs between aircrafts and UAVs. Therefore, it is better to use a random access MAC scheme in UAV network. With a random access MAC, we do not need to consider any synchronization issues as well as global coordinators. Moreover, since the distance between UAVs is short (100m $\sim$ 10km, average 1km), it is less likely to have errors in carrier sensing as well as ACK timeout.

In this paper, the U2U MAC is designed based on existing IEEE 802.11 protocols. The latest 802.11 standards [11](such as 802.11e) recommend the use of point coordination function(PCF) for scheduling control of each neighbor′s transmissions.

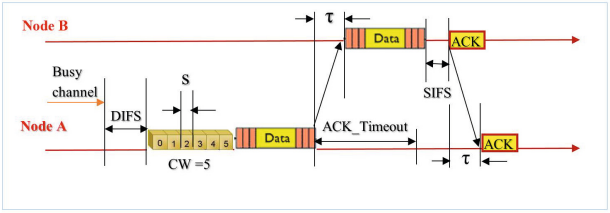


Fig. 5 Different CSMA time durations during DATA/ACK transmission

The timeline of the U2U MAC is shown in Fig. 5. Node A has a data packet for node B. According to CSMA scheme, node A first listens to the channel. If the channel is busy currently, it enters a backoff waiting phase. Once the channel is free, A would occupy the channel and sends out the packet. Here τ is the propagation delay. After B receives data packet, B waits for SIFS time and sends ACK to A. Note that the relationship between DIFS and SIFS is:

$$DIFS = 2 \times S + SIFS \tag{3}$$

In UAV network, we need to modify some parameters of 802.11e. The modifications are listed as follows:

(1) ACK Timeout: The latest 802.11 standard recommends that ACK timeout should include SIFS, S_{STD}(standard slot time) and PCLP (Physical layer convergence procedure). Here we adjust 802.11e $ACK_{Timeout}$ value as follows:

$$ACK_{Timeout} = SIFS + S_{STD} + PCLP + RTT \tag{4}$$

(2) Slot Time: The Slot Time (S) is defined as the maximum time needed to detect signal collision. In conventional 802.11, S must meet the following requirement:

$$S > (RTT/2) + CCA \tag{5}$$

Here CCA is the sum of all times except the light propagation time. In practical design, the constraint of S should consider the impact of the interference of a node on the ongoing link. In order to make sure that one node does not cause collisions to another one, we have:

$$S > (RTT/2) + 2 \times CCA \tag{6}$$

(3) DIFS: The purpose of setting SIFS and DIFS is to separate the transmission times between DATA, ACK, PCF control frames, and DCF data frames. After we adjust the slot time (S) based on the above formula, we can adjust the DIFS based on 802.11 recommendation:

$$DIFS = SIFS + 2 \times S \tag{7}$$

3.3 A2U/U2A MAC Design

In Fig. 3, we mentioned Uplink/Downlink phases. The links between UAVs and aircrafts are extremely important to a HAMN since the UAVs need to pass surveillance information data to aircrafts through uplink transmissions (U2A), and the aircrafts deliver command data via downlinks (A2U).

1. Uplink Transmission (U2A): The major issue in uplink transmission is the polling of each RoI UAVs about their data transmission requests. To quickly collect different RoI nodes requests, we propose to use compressive sensing (CS) concept to allow concurrent, uplink request transmission among a large amount of RoI nodes. CS scheme can simply ask all nodes to send out analog (instead of digital) signals in the air, and then the aircraft can use signal reconstruction to recover the original analog signal vector. Since we use analog signals to send out requests, those signals could simultaneously propagate in the air. And the aircraft can use CS signal reconstruction (again, this is analog operations) to recover the request of each RoI node.

2. Downlink Transmission (A2U): Unlike U2A transmission, in A2U we only need to broadcast the command messages. One challenging issue is multi-beam multicast (MBM) transmission. We cannot afford to lose any of those multicast messages since they are re-tasking commands. For A2U links, we propose to use link quality estimation. We require that each multicast UAV to piggyback their PER (packet error rate) and mobility information in their ACKs. After the aircraft collects the history link state parameters, it predicts the next link state via ARMA (Auto Regressive Moving Average) model.

$$\sum_{l=0}^{p} A_l y(t-l) = \sum_{l=0}^{q} M_l \epsilon(t-l) \tag{8}$$

Here A_0, A_1,..., M_0, M_1,... are all matrices of order $n \times n$, and $\epsilon(t)$ is a disturbance (noise) vector of n elements.

4 Performance Evaluation

In this section, we will show the simulation results of our proposed MAC protocol. Our results include network throughput performance, average packet delay and transmission time.

4.1 Simulation Setup

We consider an A2A network with 10 nodes. The nodes are uniformly distributed in a 300km-by-300km area. We use Ku-band signals with high data rate. The link capacity is set as 10Mb/s. As we mentioned before, in A2A network each node is equipped with a multi-beam antenna. In our simulation, each aircraft has an antenna with 4 beams, and each beam covers 90 degrees of area. Each packet contains up to 1500 Bytes of information. Each node has a buffer to store up to 30 packets. We conduct the whole simulation for 25 iterations. In each iteration, the simulation time is set to be 10s. Fist we test the overall throughput and delay performance of the A2A network. Then we simulate token scheme in Section 3.1. Finally, we test the A2U link and compressive sensing polling scheme.

4.2 Simulation Results

Figure 6 shows the throughput of the network with the average packet arrival rate. In Fig. 6 we compare our proposed A2A MAC protocol with traditional 802.11 DCF.

From the figure it is apparent that the overall network throughput with our proposed A2A MAC scheme is much better than that conventional IEEE 802.11. As we can see from the figure, when the average packet arrival rate is less than 100 packets/second/node, the two MAC schemes have similar performance. As the packet arrival rate goes up, conventional 802.11 network quickly saturates while our proposed MAC scheme could achieve almost twice throughput. The network starts to get congested when packet arrival rate is greater than 300.

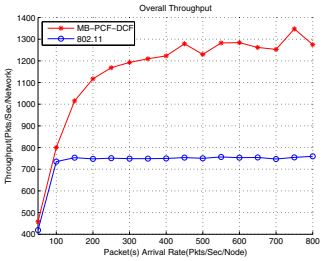


Fig. 6 A2A Throughput performance

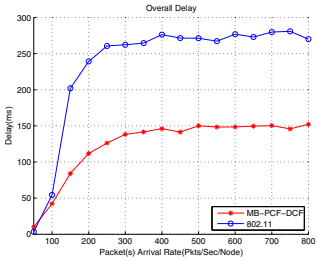


Fig. 7 A2A Delay performance

Figure 7 shows the average delay performance. It is also obvious that the proposed MAC has the lower average time delay, which could satisfy some of the QoS requirements (e.g, for video transmission the delay should be less than 200ms).

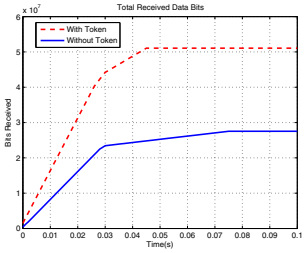


Fig. 8 Throughput of token-based scheme

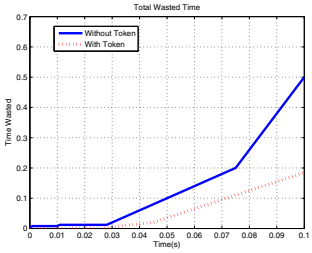


Fig. 9 Time Waste Performance

As we can see from Fig. 4, the network throughput is significantly increased with the token-based MAC scheme (almost doubled compared to non-token scheme). This is because any node can immediately switch to Tx (or Rx) mode after it finishes Rx (or Tx) phase, as long as it follows multi-beam antenna requirements (all beams should be in the same mode). Such a pipelined transmission also shortens the delay. As shown in Fig. 9, the delay is reduced for more than 50% after a certain time of communications.

5 Conclusion

In this paper, we proposed a two-layer hybrid MBSA-based MAC design for HAMN. In particular, we proposed TDMA-like A2A MAC for long-distance transmission. In UAV network, we modified the existing 802.11e to make it suitable for UAV communications. In addition, a compressive sensing based data polling scheme is used between the aircraft and its covered UAV nodes, in order to achieve fast

multi-beam multicast transmissions. The above MAC scheme has important applications in practical airborne networks. Simulation results showed that compared to existing 802.11 MAC scheme, our MAC has better performance in terms of network throughput and packet delay. Our future work includes MAC design with full-duplex transmission and anti-jamming algorithm.

References

1. Nagase, F., Mitsugi, J., Nakayama, M., Ueba, M.: Ku band mobile multimedia satellite communications system for trains. In: ICSSC, AIAA-paper-2003-2205 (2003)
2. Cheng, B.-N., Block, F.J., Hamilton, B.R., Ripplinger, D., Timmerman, C., Veytser, L., Narula-Tam, A.: Design considerations for next-generation airborne tactical networks. IEEE Communications Magazine **52**(5), 138–145 (2014)
3. Bekmezci, I., Sahingoz, O.K., Temel, Ş.: Flying ad-hoc networks (fanets): A survey. Ad Hoc Networks **11**(3), 1254–1270 (2013)
4. Lee, H.-C.: Ku-band link budget analysis of UAV with atmospheric losses. In: 25th Digital Avionics Systems Conference, 2006 IEEE/AIAA, pp. 1–8. IEEE (2006)
5. Vilzmann, R., Bettstetter, C.: A survey on MAC protocols for ad hoc networks with directional antennas. In: EUNICE 2005: Networks and Applications Towards a Ubiquitously Connected World, pp. 187–200. Springer (2006)
6. Jain, V., Gupta, A., Agrawal, D.P.: On-demand Medium Access in Multihop Wireless Nnetworks with Multiple Beam Smart Antennas. IEEE Transactions on Parallel and Distributed Systems **19**(4), 489–502 (2008)
7. Chou, Z.-T., Huang, C.-Q., Chang, J.: QoS Provisioning for Wireless LANs With Multi-Beam Access Point. IEEE Transactions on Mobile Computing **13**(9), 2113–2127 (2014)
8. Bao, L., Garcia-Luna-Aceves, J.: Transmission scheduling in ad hoc networks with directional antennas. In: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking, pp. 48–58. ACM (2002)
9. Hu, F., Li, X., Kumar, S.: Intelligent Multi-Beam Medium Access Control in Ku-band for Mission-oriented Mobile Mesh Networks. IEEE Transactions on (Submitted to) Mobile Computing (2015)
10. Nedevschi, S., Patra, R.K., Surana, S., Ratnasamy, S., Subramanian, L., Brewer, E.: An adaptive, high performance MAC for long-distance multihop wireless networks. In: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, ser. MobiCom 2008, pp. 259–270. ACM, New York (2008)
11. IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007), pp. 1–2793, March 2012

Future Approach to Find Business Model Orientation for Technological Businesses

Sepehr Ghazinoory, Fatemeh Saghafi, Maryam Mirzaei, Mohammadali Baradaran Ghahfarokhi and Parvin Baradaran Ghahfarokhi

Abstract Nowadays, the competition is not only among technological businesses but among business models as well. Previous researches have pursued the concept of future business model through future methodologies with respect to specific cases but the literature does not suggest the successful orientation of business model components under future plausible scenarios with the macro analysis level. This paper suggests an approach for making the best orientation of the new technological firms' business model components to be clear for the future. In this regard, after reviewing related literature, the industry ecosystem uncertainties should be explored in order to introduce four scenarios based on expert ideas and secondary data. Subsequently, the priority of elements of business model components will be analyzed based on each scenario. Business model components are reduced to value proposition and revenue model in this paper since the core part of the business model is proposing the value and capturing it via revenue model. This study is a quantitative research supplemented with quantitative measures with regard to its data collection method. The successful orientation of future business model is induced through a prioritizing method (GAHP) to attain the priorities of business model attributes for each element of business model components in each scenario. Entrepreneurs can map the priorities of the business model elements on designing the specific revenue model and value proposition of their own company.

Keywords Scenario development · Business model · Software industry

S. Ghazinoory
Tarbiat Modares University Tehran, Tehran, Iran
e-mail: ghazinoory@yahoo.com

F. Saghafi(✉) · M. Mirzaei · M.B. Ghahfarokhi
Faculty of Management, University of Tehran, Tehran, Iran
e-mail: fsaghafi@ut.ac.ir, mirzaeimarya@yahoo.com, mali.baradaran@gmail.com

P.B. Ghahfarokhi
PhD of Technological Entrepreneurship, University of Tehran, Tehran, Iran
e-mail: p.baradaran.g@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_10

1 Introduction

A business is in interaction with its environment of forces that shape opportunities and pose threats to the company. According to the institutional theory, the institutional environment intensely impacts the formation of official structures in an organization [1]. However, the very changing and unstable current business conditions [2] lead to complexity and increases the future uncertainty [3]. In such an altering atmosphere, businesses ought to make changes in their business models regularly in order to maintain the competitiveness and adjust or shift the business form [4, 5]. Indeed, the competition is not only among businesses but among business models as well [6].

On the other hand, futures study is capable of sketching the alternative futures of the industry environment [7] and proposing several approaches for insight acquisition about the state of the future [8] among which scenario planning has been considered by organizational leaders as an effective method for understanding future uncertainties, readiness against undetermined future, mental model change, decision test and performance enhancement [9].

One of the fundamental essentials for changing business models is to collect, process and analyze the industry situation from the political, economic, social and technological perspectives over time [10]. This challenge is complicated when this information should sketch futures which enable businesses to design or revise their business models in accordance with them due to institutional complexity from opacity in decision making in governmental institutions and regulatory to management and culture.

Previous researches have pursued the concept of future business model through future methodologies with respect to specific cases but the literature does not suggest the better state of the business model elements. “A good business model is a profitable investment for owners, a beneficial service for customer, a rewarding and knowhow-favorable work organization for employees and a competitive business strategy compared to competitors” [11]. However, the need for understanding the successful orientation of business model components under alternative futures still remains unfolded. Therefore, this paper suggests an approach for making the best orientation of the firms’ business model components to be clear for the future. In this regard, after reviewing related literature, the firm ecosystem uncertainties will be explored in order to introduce four scenarios based on expert ideas and secondary data. Subsequently, the priority of elements of business model components will be analyzed based on each scenario. Business model components are reduced to value proposition and revenue model in this paper since the core part of the business model is proposing the value and capturing it via revenue model [12].

2 Literature Review

The secret for the success of a company lays in the future orientation of the company, paired with strong foresight capability based on flexible and adaptable systems [13]. One of the systems [14] that the company should have the capability to adapt it to the future is its business model(s).

2.1 Futures Studies and Business Model

There are researches that have approached to the concept of business model via future scenarios. The first part points to how external factors often force companies to rethink their business models. Based on scenario analyses, it can be determined whether the business models fit generic trends, one specific scenario or all scenarios. Depending on the results, it can be decided whether or not to pursue the current business model. They believe that scenarios help reducing uncertainty and direct design issues. Bouwman, Faber [15] , Bouwman, De Vos [16] and Gnatzy and Moser [10] are classified in this category.

In the second perspective which is called, the robustness of specific business models would be tested through what-if scenarios. Haaker and Van Buuren's [17] and Klemettinen [18] are vategorized in this section.

All the above works lack a universal view on future business models orientation in the mind set of business persons of a specific industry. However, current paper argues that under the developed future states, the main orientation of any successful and acceptable business model can be explored by prioritizing the elements of the core part.

2.2 Scenario Development

"Scenario planning is a process of positing several informed, plausible and imagined alternative future environments in which decisions about the future may be played out, for the purpose of changing current thinking, improving decision making, enhancing human and organization learning and improving performance" [9]. In this line, policy makers and strategists in various industries use scenarios to develop and test their strategy robustness in alternative futures [19]. As a matter of fact, commercial success of firms and innovation development is in a close relationship with the market information and application of those information. Access to this information enables businesses to recognize environment change in the very early stages to be proactively ready for new conditions [20, 21]. Van der Heijden [22] suggests that scenarios are the best language for strategic issues since they represent different futures besides creating common perception of future for individuals to make better decisions. Scenario planning approach reduces the risk of focusing on a single expected future instead of considering all plausible futures [23]. Furthermore, they should be created from all the stakeholders' points of view. The aim of scenarios is to recognize the change components which have

intense effect on industry despite of their undetermined consequences. Generally, scenarios are built in a 2 dimension matrix which makes scenarios different from each other [22].

At the next level, there are different approaches for developing scenarios which are classified by Bradfield (2005) to “The intuitive logics school”, “The probabilistic modified trends school” and “The La prospective school”. Huss and Honton [24] also categorized scenario planning methodology approaches to intuitive logics, trend impact analysis and cross-impact analysis, believing that since ‘the intuitive logic approach is not tied to any mathematical algorithm, it can, with careful tailoring, adjust to the particular needs and political environment of the company. The approach basic presumption is that the decisions of the company is based on bundle of complicated relations among economic, political, technological, social and environmental factors; and the changes and effects of them should be recognized.

Abundance of various methods of Intuitive Logic depicts that the number of these methods are as much as the researchers applying them [25]. Researchers believe that there is no theoretically, operationally and publicly acceptable methodology for scenario development in this approach [26, 27].

One of the techniques [23] of the mentioned approach is GBN which starts with listing future variables and trends and classifying them to (1) trends and variables with certain effects on the subject, (2) trends and variables with no certain effects on the subject, and (3) trends and variables with uncertain effects on the subject and using those as the basis for alternative futures

In this paper GBN technique is used due to its inherit uncertainty in predictive forecasting which means information is not sufficient and human systems are in chaos and emergent state and related human behavior theories are not as determined as mathematical ones. The technique is the focal subject of this paper and the paper concerns with it as a default scenario technique [23].

2.3 Business Model

Business models are commonly defined as the logic of the enterprise, the way it operates, and how it creates and captures value for its stakeholders [28]. Being studied for various purposes, business model components or taxonomies are developed [29]. In accordance with the purpose of this article, the business model component approach is applicable to clarify the underlying priorities of business model elements under each scenario. Among several proposed models, business model canvas [30] is chosen due to three reasons. First, it is a popular tool that makes it simple for practitioners to design business models in a creative session [29]. As a result, many business reviewer of this article may use this tool, thus findings of this paper should be compatible with their used tool. Second, since this papers does not focus on a single case and has a general point of view, it picks the canvas with a vivid classification of every aspects of a business model which can be generally applicable for an industry. Finally, its ontological approach which

creates a hierarchical classification of business model components, their elements and their attributes in three levels makes an appropriate structure for the prioritizing purpose of this paper.

The canvas introduces nine components including value proposition and revenue model. Revenue model is the ability of a firm to translate the value that it offers its customers into money and incoming revenue streams; and value proposition is described as defining how items of value, such as products and services as well as complementary value-added services are packaged and offered to satisfy the customer needs [30]. The hierarchical structure of such business model is illustrated in Figure (1). The meaning of criteria and attributes can be found in (Osterwalder, 2004).

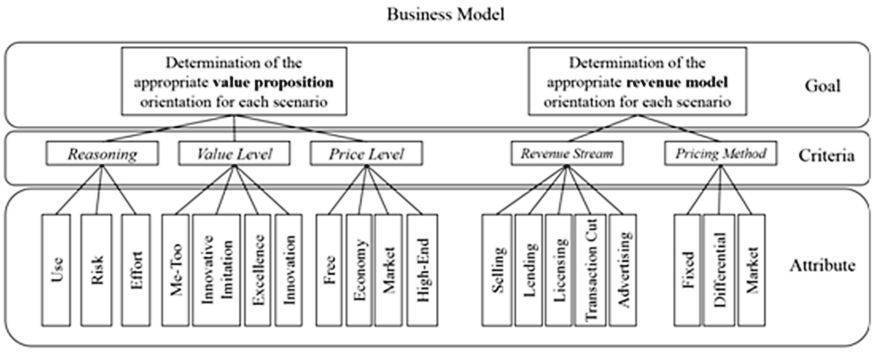


Fig. 1 (1) hierarchical structure of business model

3 Methodology of the Approach

This article aims at investigating the future orientation of the business models of a certain industry. Thus in terms of the objective, this is an applied study. Moreover, this approach is a quantitative research supplemented with quantitative measures with regard to data collection method and the final results. The proposed scenarios of this approach are explorative [31]. The paper includes two general parts: developing the scenarios of the software industry; and prioritizing the business model. The former includes three steps in its own.

3.1 Part One: Scenario Development

At the first step, the environmental data are specified based on the institutional theory [32] using PEST framework and Porter's market forces [33] in the environment of the industry through studying the available literature and documents of the journals, newspapers and articles. Then the conventional content analysis is used [34] to discover a classification of driving forces of the industry.

At the second step, semi-structured face-to-face interviews with open-ended questionnaire should be used in order to determine the key uncertainties [35]. This part deals with asking the level of importance and ambiguity of the key factors for discovering the key uncertainties. Each of the Likert scales are converted to the numerical measures so that the value 1 implies the high importance and high vagueness; the value 0.5 implies the average vagueness or importance, and the value 0.1 implies the low importance or vagueness, while their multiplication in a 3 × 3 matrix is shown in the following table [36]. Then the importance level of the uncertainty is calculated by equation (1) as explained in detail by[37]. Equation (1) $r_{ij} = \alpha_{ij}\beta_{ij}$

Where r_{ij} is the uncertainty score assessed by respondent j for the factor i; i = ordinal number of risk, i∈(1,m); m = total number of uncertainty; j = ordinal number of expert i, j∈(1,n); n = total number of experts i; α_{ij}= vagueness level of the factor i, assessed by expert j; β_{ij} = importance level of the factor i, assessed by expert j.

Table 1 Matrix for calculating the uncertainty score

i person	β		
α	Higher importance	Average importance	Lower importance
Highly vague	1	0.5	0.1
Probable	0.5	0.25	0.05
Very unlikely	0.1	0.05	0.01

The average score for each uncertainty is calculated through Eq. (2) which is used for ranking uncertainties. R_i equals the significance index score for uncertainty i.

Equation (2) $$R_i = \frac{\sum_{j=1}^{n} r_{ij}}{n} = \frac{1}{n}\sum_{j=1}^{n} \alpha_{ij}\beta_{ij}$$

The uncertainties are ranked based on their scores, and then the two first factors are selected as the most important ones and as the logic for the formation of the scenarios. The two factors with highest value are selected as the key uncertainties and are selected as the framework for formation of the scenario matrix.

In The last step, we have to enrich and present the scenarios. For formulating the scenarios we have to consider the logic of the scenarios as the framework of each scenario. Then the scenarios are formulated by investigating the effects of the key uncertainties on the several factors as identified in previous steps and their dynamics in each scenario. Hence four scenarios should be formulated and presented to the experts of the field and their corrective advices should be applied in a panel of experts.

3.2 Part Two: Prioritizing the Business Model

In order to understand the successful orientation of future software business model in a normative approach, a prioritizing method (MADM) is used to attain the priorities of business model attributes for each element of business model components. Therefore, Analytic Hierarchy Process (GAHP) method identifies the weight of each attribute sets per scenarios. For implementing GAHP separately in form of five pairs, wise comparison questionnaires under each scenario are used. Three of them pertain to the value proposition and two of them refer to the revenue model component.

AHP is a powerful method in decision making, firstly introduced by [38] to pattern subjective decision-making procedure based on multiple features on a hierarchical system.

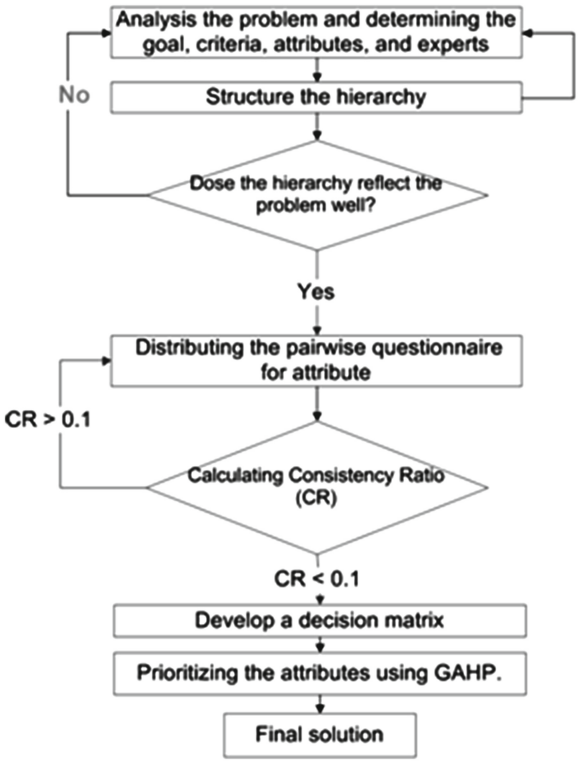


Fig. 2 (2) AHP methodology chart

4 Conclusion and Discussion

Technological businesses have to pay attention to the way of their interaction with the environment and their facing changes and challenges. The future successfulness of the business models depends on the proactive approach to such changes. Thus if the companies manage to be aware of the successful orientation of such business model components in the future scenarios, then they would be able to benefit from the mentioned general model in designing the business models of their company.

Moreover, they can consider the results of this article as the starting point of their business model roadmap and pay attention to the planning of needed technologies and attaining the needed resources.

By and large, each scenario of the future of the industry has some threats and opportunities for the active companies and future entrepreneurs. Companies and entrepreneurs can survive in such an environment by choosing the best design of the business model for their services and products. Thus this research is practically useful both for entrepreneurs, and the policy-makers who are supposed to support the future of the industry.

References

1. Peng, M.W., Wang, D.Y., Jiang, Y.: An institution-based view of international business strategy: A focus on emerging economies. Journal of International Business Studies **39**(5), 920–936 (2008)
2. Botha, A., Kourie, D., Snyman, R.: Coping with continuous change in the business environment: knowledge management and knowledge management technology. Elsevier (2014)
3. Chesbrough, H.: Business model innovation: opportunities and barriers. Long Range Planning **43**(2), 354–363 (2010)
4. Casadesus-Masanell, R., Zhu, F.: Business Model Innovation and Competitive Imitation. Citeseer (2010)
5. Andries, P., Debackere, K.: Adaptation in new technology-based ventures: Insights at the company level. International Journal of Management Reviews **8**(2), 91–112 (2006)
6. Morris, L.: Business model warfare: The strategy of business breakthroughs. Journal of Business Models **1**(1), 13–37 (2013)
7. Olsmats, C., Kaivo-oja, J.: European packaging industry foresight study—identifying global drivers and driven packaging industry implications of the global megatrends. European Journal of Futures Research **2**(1), 1–10 (2014)
8. Bouman, H., Haaker, T., De Reuver, M.: Some reflections on the high expectations as formulated in the Internet Bubble era. Futures **44**(5), 420–430 (2012)
9. Chermack, T., Lynham, S.A., Ruona, W.: A review of scenario planning literature. Futures Research Quarterly **17**(2) (2001)
10. Räsänen, K.: kehittyvä liiketoiminta: Haaste tulevaisuuden osaajille. Espoo, Weilin & Göös (1997)

11. Boons, F., Lüdeke-Freund, F.: Business models for sustainable innovation: state-of-the-art and steps towards a research agenda. Journal of Cleaner Production **45**, 9–19 (2013)
12. Battistella, C.: The organisation of Corporate Foresight: A multiple case study in the telecommunication industry. Technological Forecasting and Social Change **87**, 60–79 (2014)
13. Zott, C., Amit, R.: Business model design: an activity system perspective. Long Range Planning **43**(2), 216–226 (2010)
14. Bouwman, H., Faber, E., Van der Spek, J.: Connecting future scenarios to business models of insurance intermediaries. BLED 2005 Proceedings, p. 16 (2005)
15. Bouwman, H., De Vos, H., Haaker, T.: Mobile service innovation and business models. Springer Science & Business Media (2008)
16. Gnatzy, T., Moser, R.: Scenario development for an evolving health insurance industry in rural India: INPUT for business model innovation. Technological Forecasting and Social Change **79**(4), 688–699 (2012)
17. Haaker, T., Buuren, R.V.: Future scenarios approach for aiding the design of innovative systems. In: 16th European ITS Regional Conference, Porto, Portugal (2005)
18. Klemettinen, M.: Enabling technologies for mobile services: the MobiLife book. John Wiley & Sons (2007)
19. Ringland, G., SCHWARTZ, P.P.: Scenario planning: managing for the future. John Wiley & Sons (1998)
20. Frishammar, J.: Managing information in new product development: A literature review. International Journal of Innovation and Technology Management **2**(03), 259–275 (2005)
21. Galliers, R.: Information analysis: selected readings. Addison-Wesley Longman Publishing Co., Inc. (1987)
22. Van der Heijden, K.: Scenarios: the art of strategic conversation. John Wiley & Sons (2011)
23. Bishop, P., Hines, A., Collins, T.: The current state of scenario development: an overview of techniques. Foresight **9**(1), 5–25 (2007)
24. Huss, W.R., Honton, E.J.: Scenario planning—what style should you use? Long Range Planning **20**(4), 21–29 (1987)
25. Bradfield, R., El-Sayed, H.: Four scenarios for the future of the pharmaceutical industry. Technology Analysis & Strategic Management **21**(2), 195–212 (2009)
26. Bunn, D.W., Salo, A.A.: Forecasting with scenarios. European Journal of Operational Research **68**(3), 291–303 (1993)
27. Jungermann, H.: Inferential processes in the construction of scenarios. Journal of Forecasting **4**(4), 321–327 (1985)
28. Brea-Solís, H., Casadesus-Masanell, R., Grifell-Tatjé, E.: Business model evaluation: quantifying Walmart's sources of advantage. Strategic Entrepreneurship Journal **9**(1), 12–33 (2015)
29. De Reuver, M., Bouwman, H., Maclnnes, I.: Business model dynamics: a case survey. Journal of Theoretical and Applied Electronic Commerce Research **4**(1), 1–11 (2009)
30. Osterwalder, A.: The business model ontology: A proposition in a design science approach. University of Lausanne, Ecole des Hautes Etdes Commerciales (2004)
31. Börjeson, L., et al.: Scenario types and techniques: towards a user's guide. Futures **38**(7), 723–739 (2006)
32. Meyer, K.E., et al.: Institutions, resources, and entry strategies in emerging economies. Strategic Management Journal **30**(1), 61–80 (2009)

33. Porter, M.E.: The five competitive forces that shape strategy. Harvard Business Review **86**(1), 78–93 (2008)
34. Hsieh, H.-F., Shannon, S.E.: Three approaches to qualitative content analysis. Qualitative Health Research **15**(9), 1277–1288 (2005)
35. Gabzdylova, B., Raffensperger, J.F., Castka, P.: Sustainability in the New Zealand wine industry: drivers, stakeholders and practices. Journal of Cleaner Production **17**(11), 992–998 (2009)
36. Zou, P.X., Zhang, G., Wang, J.: Understanding the key risks in construction projects in China. International Journal of Project Management **25**(6), 601–614 (2007)
37. Zou, P.X., Zhang, G., Wang, J.-Y.: Identifying key risks in construction projects: life cycle and stakeholder perspectives. In: Pacific Rim Real Estate Society Conference (2006)
38. Saaty, T.L., Vargas, L.G.: The Logic of Priorities

Social Media Coverage of Public Health Issues in China: A Content Analysis of Weibo News Posts

Jiayin Pei, Guang Yu and Peng Shan

Abstract Content analysis is a useful tool for better understanding how different portrays of public health issues may affect news dissemination. In this research, we conducted a content analysis on health-related messages published by opinion-leading news outlets. We examined generic content attributes, including number of words, sentences, images, links, and published time, as well as content valence features. Our findings show that there are no significant differences in the amount words, sentences and hyperlinks used between the highly forwarded and lowly forwarded news posts when covering a public health issue. Whereas the topic and title length as well as content valence of highly forwarded new posts is different from the lowly forwarded ones. We also discuss ways to improve popularity of health-related news.

Keywords Content analysis · Feature analysis · Microblogging · Public management · Social media

1 Introduction

Mass media determine, to a considerable extent, the amount and kind of health-related information that reaches the public. Such information provides feedback on the characteristics and nature of risk [1, 2], and may influence public perceptions of and affect public attitudes towards a public health issue [3, 4].

J. Pei(✉) · G. Yu
School of Management, Harbin Institute of Technology, Harbin 150001, China
e-mail: {peijiayin,yug}@hit.edu.cn

P. Shan
Business School, Jiangnan University, Wuxi 214122, China
e-mail: 51060681@qq.com

This work is supported by National Natural Science Foundation of China under Grant No.71171068.

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_11

Content analysis of health related information has been the focus of some newspaper content analysis research work, which were mostly done through survey questionnaire [5, 6]. Social media news coverage of public health issues [7] differ significantly from traditional news outlets in language style, audience reach as well as the way public opinions are recorded. In addition, computer assisted content analysis can provide more objective information towards a public health issue. For example, instead of the subjective coding process by two or more coders, a large number of forwards of a news post can simply represent mass public concern about a specific issue. Despite the advantages, there are no in-depth studies on analyzing the content features of health-related microblogging news posts, even though this clearly is an interesting and challenging domain.

In addition, the researches that try to discover media's role in communicating health risks have been limited in scope and mostly drawn upon the context of developed nations [8-11]. So far, content analysis of social media coverage on particularly public health issues in developing countries, such as China, remained almost unconsidered.

For this reason, this study makes use of quantitative content analysis methods to examine the content features of health-related messages on Weibo, a Chinese microblogging site. The objective of this research is to find out main characteristics of the news content about public health issues. In addition, we observed that while some new posts disseminate widely, other posts are paid less attention to. Thus a content analysis for different levels of public dissemination was performed to discover what public health related news characteristics are more attractive to the public.

To answer these questions, we first assess the content features of the public health related news in section 2.1 to identify aspects that might be linked to public perceptions of health related issues. Secondly, the content differences between these two categories of posts need to be determined with a method introduced in Section 2.2. We then report the results and finally conclude.

2 Methodology

2.1 Content Feature Analysis

The Content features can provide descriptive characteristics for discriminating news posts that will be widely forwarded from those that will not be widely forwarded. The selection of variables demonstrating latent content is particularly important [12]. For this research, we employed a deductive method to manifest content [13], that is, selecting content variables which are observable and countable. These attributes can include some generic content characteristics (such as title length, topic length, number of sentences, and publish time, etc.) and content valence.

The generic content characteristics affect news presentation [14-17]. For example, a news information will be more appealing if it clearly presents a news issue or if it contains one or more visual elements such as graphics and hyperlinks. Besides, title length and number of sentences affect readability of the health-related news. If a post is not long, no doubt the breadth/depth of discussion regarding public health issues will be limited to the mention of the issue only, and thus will not be attractive to the public. In addition, a news post is published in business hour may be read by fewer people because a number of people are at work. To examine the influence of generic content characteristics on the spreading of health-related news on Chinese microblogging site. We test the following hypotheses.

Hypothesis 1: The generic content characteristics of the health-related news will differ between the highly and lowly forwarded news posts. Specifically:

Hypothesis 1a: The amount words, sentences and hyperlinks in the health-related news will differ between the highly and lowly forwarded news posts.

Hypothesis 1b: The length of the health-related news as well as its title length and topic length will differ between the highly and lowly forwarded news posts.

Hypothesis 1c: The time the health-related news is published will differ between the highly forwarded and lowly forwarded news post.

Looking beyond the generic content features, the valence of a story may also be subject to individuals' news evaluations [18-21]. Analysis of the valence of a news post is mainly based on identifying positive or negative words and emotion expressions, and to processing text with the purpose of classifying its emotion as positive or negative [22]. Traditional persuasion research have discovered that content valance is an important factor when persuading citizens [23]. It is hypothesized in this article that:

Hypothesis 2: Highly forwarded news will show different valence towards a public health issue with lowly forwarded ones.

Overall, the content of health-related news posts extracted fifteen variables to further specify the news content. These elements were to be calculated through computer assisted quantitative content analysis (described in the following section) for each news message in the dataset, resulting in structured data set to be analyzed.

2.2 Computer Assisted Quantitative Content Analysis

Content analysis is a research method that generally involves a "systematic and replicable" analysis of messages [24]. It grew out of quantitative newspaper analysis in the USA in the 1920s [12] and was traditionally utilized by communication researchers. Whereas in recent years, scholars from other disciplines began using the method, and thus it has gain its popularity as a useful methodology for studying news messages [25, 26].

Many of the early studies focused on the quality of the news presented [13]. Findings from these traditional content analysis did offer important information. This focus has remained central in online media analysis [27, 28]. However, in the big data environment, as for the detection or decision making tasks based on enormous microblog posts, methodologies designed for measuring traditional media content (e.g. through coding process processed by two coders) do not satisfy online news content. Instead, researches analyzing online social media messages have to make use of computer assisted quantitative content analysis methods. This research follows traditional quantitative content analysis of news, but attempts to find appropriate content variables that can be automatically processed to characterize the categories of features.

To evaluate the content of information on public health issues available to the Chinese public, we undertook a content analysis of different news sources. We first selected 12 Chinese mainstream news outlets as the primary focus of the analysis. We then collected all the news messages posted by these news outlets by the end of April 2014 and manually selected news posts that are reporting public health issues. The final data set was comprised of 863 health-related news posts published in the four months period from 1 January 2014 until 30 April 2014[29].

Next, we conducted a quantitative content analysis of news posts published by the 12 mainstream news outlets to examine content related to public health issues. The coding variables corresponded with all the content features.

Coding of the articles was performed in two stages. The first stage involved using a java language that was designed by us to process all the related features, including the generic features and the valence of the post. This script scanned the articles and identified the number and length of specific content features, which was then automatically entered into a spreadsheet. Fig.1 shows a snip of Java code to select and count the content features. The coding process was conducted automatically that could minimize human error and allow automatic detection of alert signals. In the second stage, posts were coded for valence.

```
public int GetFrequency(String msg, String regExp)
{
    int counter=0;
    try{
    Pattern pattern=Pattern.compile(regExp);
    Matcher m= pattern.matcher(msg);
    while(m.find())
    {
        counter++;
    }
    return counter;
    }
    catch(Exception e)
    {//System.out.println(regExp+"---"+msg);
    return counter;}
}
```

Fig. 1 An example of a snip of Java code to select and count the content features

These steps resulted in a structured data set for statistical analysis. The final data set contain 15 variables corresponding to the previously discussed 15 features, including 10 features that represent the generic content of a post and 5 features that describe the affective content feature. Wilcoxon test was then performed on the dataset to test differences of the content features between news groups with different levels of public dissemination.

3 Results

3.1 Main Characters

Fig. 2 shows a distribution of all the 14 features that has a value (the news messages show no positive emotion). In our data set, the analyzed news outlets published 625 lowly forwarded news posts about public health issues in the first four months in 2014, which almost double the 238 highly forwarded news posts published by these news sources.

Public health issue is not a topic that can be covered in a short post unless it is accompanied with a link to other source information. Our result shows that posts about public health issues are not very short, although it varies in a wide range (standard deviation 17.37). Among the data sample, the shortest length of a news post is 16 and the highest is 208. The mean length of the analyzed health related news is approximately 143 words.

Overall, most news posts about public health issues are with picture (789 out of 863 posts). This represents that most of the news posts will repost a public health issue with a picture accompanied. The use of links are not as popular as the use of

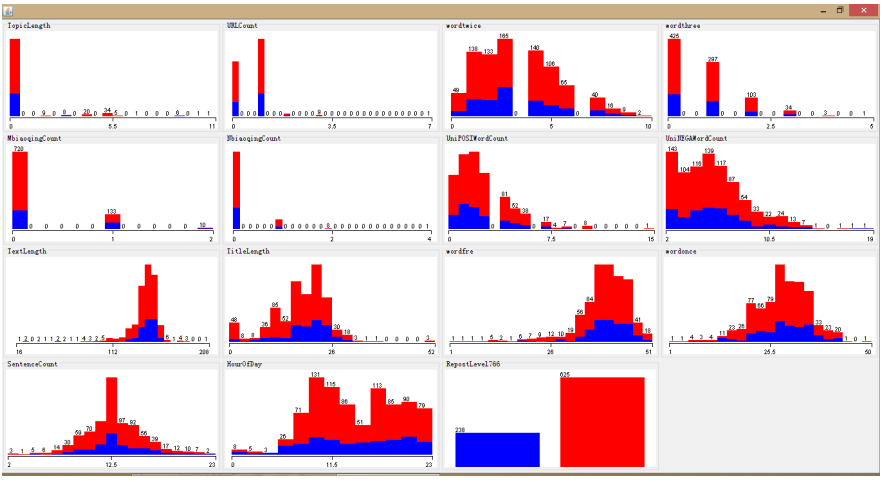


Fig. 2 Distribution of the 14 content features

graphs among the selected news outlets, but still more posts share a link when covering a public health issue (515 out of 863 posts). Similarly, most news posts post on a public health issue with title (815 out of 863 posts). While, surprisingly, only few posts contain a keyword hashtag when broadcasting public health issues. In addition, very few news are posted in the early morning (42 out of 863 posts). While news media publish more posts later in the day (567 out of 863 posts). Our results differ from previous researches [30, 31] in the time peak activities happens. We suppose one reason might be the behaviors of twitter users and Weibo users are different. Since we are focusing on analyzing the publish behavior of news outlets, the other reason for such a peak hour difference might be the publishing behavior of news outlets are different from the main public.

3.2 Feature Analysis

As is shown in tab.1, contrary to hypothesis 1a, there is no significant difference (Pr>Chi-Square over 0.05) in the amount words, sentences and hyperlinks used between the highly and lowly forwarded health-related news posts.

The mean length of highly forwarded new posts' title are different from the lowly forwarded ones when covering public health issues. The title of highly forwarded news tend to be not too long and not too short. Approximately 90.34% of the news posts in this category are no longer than 17 words and 1.26% of the news messages contain fewer than 3 words. Surprisingly all highly forwarded news posts appeared to have a title (no zero as title length). However, the lowly forwarded news posts contain either very long title or very few words. Approximately 91.89% of the news posts' title in this category contain more than ten words and 4.75% of the posts include no title.

Table 1 The generic content features and valence of media coverage of public health issues on Chinese microblogging site

	Minimum	Maximum	Mean	StdDev	Pr>Chi-Square
Generic content features					
number of links shared	0	7	0.643	0.615	0.5580
the length of the news post	16	208	143.417	17.37	0.0578
the length of the title	0	52	18.117	7.021	<0.0001
the length of the topic	0	11	0.468	1.533	0.0180
number of sentences	2	23	12.749	3.025	0.1011
number of words	1	51	39.148	6.447	0.2224
number of words appeared once	1	50	28.557	6.505	0.1563
number of words appeared twice	0	10	3.346	2.05	0.9627
number of words appeared three times	0	5	0.721	0.861	0.8851
published hour of day	0	23	14.256	5.018	0.8640
Valence					
number of unique positive words	0	15	2.45	2.068	0.1933
number of positive emotions	0	0	0	0	
the number of neutral emotion	0	2	0.177	0.411	<0.0001
number of unique negative words	2	19	6.242	2.797	0.3396
the number of negative emotions	0	4	0.134	0.386	0.6265

The mean length of highly forwarded new posts' topic is also different from the lowly forwarded ones when covering a public health issue. All the highly forwarded news posts contain topic, while 88.64% of the lowly forwarded news posts indicate no topic. We provide a plausible explanation to this phenomenon. A topic appeared in the news helps information seekers locate articles of interest more easily. Besides, it reminds readers that the public health issues reported are in hot discussion and thus posts that has attracted much public concern tend to be highly forwarded.

Hypothesis 1b is partly supported. The title and topic length of the health-related news will differ between the highly and lowly forwarded news posts is supported (Pr>Chi-Square less than 0.0001 and equals 0.0180 respectively). Whereas the length of the health-related news will differ between the highly and lowly forwarded news posts is not supported (Pr>Chi-Square equals 0.0578).

Contrary to hypothesis 1c, there is no significant difference (Pr>Chi-Square 0.8640) in the time the health-related news is published between the highly and lowly forwarded news posts.

The news outlets did write about public health issues with different valence as previously hypothesized in the methods section (Hypothesis 2). The difference lies in the uses of neutral emotion expressions like surprised. Highly forwarded news was more likely to contain one (24.37%) or two (1.68%) neutral emotion expressions than the lowly forwarded ones, with one (12%) and two (0.96%) neutral face included.

4 Conclusions

Media influence public awareness by selecting information about how the issue is interpreted [32]. People's perceptions of and concerns about health risks are thus shaped by how public health issues are portrayed in the news. An examination of how certain features of interpretation may result to widely dissemination is of great interest. Content analysis of social media data like the one described in this article can reveal features of high user interest about public health issues. It can contribute to the examination of how different characteristics may result in varying popularity of news posts, which may ultimately affect public concern about public health issues.

In this paper, results of a computer assisted quantitative content analysis of different health-related microblog messages have been presented. The study finds significant differences in the title and topic length of the highly forwarded news posts and the lowly forwarded news posts. All the highly forwarded news posts contain title and topic. Besides, the title of highly forwarded news tend to be in a moderate range from the average. The two groups of health-related news also differed in their valence, with highly forwarded news more likely to contain neutral emotion expressions than the lowly forwarded ones. Thus in order to make a health-related news popular, it is advised to include a title that is not too long and

not to short. In addition, a topic if appropriate and a neutral emotion expression is encouraged to be included in the content of a health-related news post.

It turns out that the analyzed content features did not vary significantly in this sample content of health-related information. We remark that only three content features checked show difference among the highly and lowly forwarded news posts. Some features like the subjective or objective attitude conveyed in the news posts that may affect news dissemination are to be taken into consideration. In the future, we plan to expand our analysis to other features, by making comparisons of different categories of public health issues and more news posts from additional news sources, in order to further support the results presented in this paper. Besides, the role of news sources should be checked. In future work, we will focus on these issues in more detail.

References

1. Claassen, L., Smid, T., Woudenberg, F., Timmermans, D.R.M.: Media coverage on electromagnetic fields and health: Content analysis of Dutch newspaper articles and websites. Health, Risk & Society. **14**(7–8), 681–696 (2012)
2. Boin, A., Renaud, C.: Orchestrating joint sensemaking across government levels: challenges and requirements for crisis leadership. Journal of Leadership Studies **7**(3), 41–46 (2013)
3. McCallum, D.B., Hammond, S.L., Covello, V.T.: Communicating about environmental risks: how the public uses and perceives information sources. Health Education Quarterly **18**(3), 349–361 (1991)
4. Powell, M., Dunwoody, S., Griffin, R., Neuwirth, K.: Exploring lay uncertainty about an environmental health risk. Public Understanding of Science **16**(3), 323–343 (2007)
5. Tanzler, D., Feil, M., Kromker, D., Eierdanz, F.: The challenge of validating vulnerability estimates: the option of media content analysis for identifying drought-related crises. Reg Environ Change **8**(4), 187–195 (2008)
6. Moshrefzadeh, A., Rice, W., Pederson, A., Okoli, C.T.C.: A content analysis of media coverage of the introduction of a smoke-free bylaw in vancouver parks and beaches. Int J Env Res Pub He. **10**(9), 4444–4453 (2013)
7. Glik, D.C.: Risk communication for public health emergencies. Annual Review Of Public Health **28**(1), 33–54 (2007)
8. Freberg, K., Palenchar, M.J., Veil, S.R.: Managing and sharing H1N1 crisis information using social media bookmarking services. Public Relations Review **39**(3), 178–184 (2013)
9. Rutsaert, P., Pieniak, Z., Regan, Á., McConnon, Á., Kuttschreuter, M., Lores, M., et al.: Social media as a useful tool in food risk and benefit communication? A strategic orientation approach. Food Policy **46**, 84–93 (2014)
10. Gaspar, R., Gorjão, S., Seibt, B., Lima, L., Barnett, J., Moss, A., et al.: Tweeting during food crises: a psychosocial analysis of threat coping expressions in Spain, during the 2011 European EHEC outbreak. International Journal of Human-Computer Studies **72**(2), 239–254 (2014)

11. Shepherd, A., Sanders, C., Doyle, M., Shaw, J.: Using social media for support and feedback by mental health service users: thematic analysis of a twitter conversation. BMC psychiatry **15**, 29 (2015)
12. Sjovaag, H., Stavelin, E.: Web media and the quantitative content analysis: methodological challenges in measuring online news content. Convergence-US **18**(2), 215–229 (2012)
13. Neuendorf, K.A.: The content analysis guidebook. Sage (2002)
14. Martinez-Romo, J., Araujo, L.: Detecting malicious tweets in trending topics using a statistical analysis of language. Expert Systems with Applications **40**(8), 2992–3000 (2013)
15. Peñalver-Martinez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Rodríguez-García, M.Á., Moreno, V., Fraga, A., et al.: Feature-based opinion mining through ontologies. Expert Systems with Applications **41**(13), 5995–6008 (2014)
16. Lu, X., Yu, Z., Guo, B., Zhou, X.: Predicting the content dissemination trends by repost behavior modeling in mobile social networks. Journal of Network and Computer Applications **42**, 197–207 (2014)
17. Zhang, K.Z.K., Zhao, S.J., Cheung, C.M.K., Lee, M.K.O.: Examining the influence of online reviews on consumers' decision-making: a heuristic-systematic model. Decision Support Systems **67**, 78–89 (2014)
18. De Vreese, C., Boomgaarden, H.: Valenced news frames and public support for the EU. Communications **28**(4), 361–381 (2003)
19. Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D.C.L.: Text mining of news-headlines for FOREX market prediction: a multi-layer dimension reduction algorithm with semantics and sentiment. Expert Systems with Applications **42**(1), 306–324 (2015)
20. Flanagin, A.J., Metzger, M.J.: Trusting expert- versus user-generated ratings online: the role of information volume, valence, and consumer characteristics. Computers in Human Behavior **29**(4), 1626–1634 (2013)
21. Kim, K., Lee, J.: Sentiment visualization and classification via semi-supervised nonlinear dimensionality reduction. Pattern Recognition **47**(2), 758–768 (2014)
22. Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D.C.L.: Text mining for market prediction: a systematic review. Expert Systems with Applications **41**(16), 7653–7670 (2014)
23. O'Keefe, D.J.: Trends and prospects in persuasion theory and research. Canadian Journal of Communication (1991)
24. Riffe, D., Lacy, S., Fico, F.G.: Analyzing media messages: using quantitative content analysis in research. Lawrence Erlbaum Associates, Mahwah-New Jersey (2005)
25. Denecke, K., Nejdl, W.: How valuable is medical social media data? content analysis of the medical web. Information Sciences. **179**(12), 1870–1880 (2009)
26. Bjerre, L., Paterson, N.R., McGowan, J., Hogg, W., Campbell, C.: Do Continuing Medical Education (CME) events cover the content physicians want to know? a content analysis of CME offerings. The Journal of continuing education in the health professions 35(1), 27--37 (2015)
27. Ormond, D., Warkentin, M.: Is this a joke? The impact of message manipulations on risk perceptions. Journal of Computer Information Systems **55**(2), 9–19 (2015)
28. Myneni, S., Cobb, N.K., Cohen, T.: Finding meaning in social media: content-based social network analysis of quitnet to identify new opportunities for health promotion. Stud Health Technol. **192**, 807–811 (2013)

29. Pei, J., Yu, G., Tian, X., Donnelley, M.R.: A new method for early detection of mass concern about public health issues. Journal of Risk Research, 1–17 (2015)
30. Wimpory, D., Nicholas, B., Nash, S.: Social timing, clock genes and autism: a new hypothesis. Journal of Intellectual Disability Research **46**(4), 352–358 (2002)
31. Alwagait, E., Shahzad, B. (eds.) Maximization of Tweet's viewership with respect to time. IEEE (2014)
32. Coombs, W.T.: Ongoing Crisis Communication: Planning, Managing, and Responding. 3rd revised edn. SAGE Publications, Thousand Oaks (2011)

Advertisement Through the Most Popular Twitter Users Based on Followers in Saudi Arabia

Abeer A. AlSanad and Abdulrahman A. Mirza

Abstract Advertising is a major commercial activity in the Internet. Nowadays, online social networks present a new way of disseminating advertisements. This new way is considered the fastest way to reach large groups of customers. Even so, there has been little research focus on advertising through users in social networks. In this research paper, a focus was made on one of today's most popular social networks which is Twitter. Our intention is to assess the extent of the utilization of the most followed Saudi Arabian Twitter users in making advertisements. This study employs a survey method that measures the paper goal.

Keywords Social networks · Advertisement · Twitter

1 Introduction

At the present time, social media has spread dramatically and got a large number of users with a long time spent by each user. So it is a rich field for businesses to advertise their products and services through social media [1]. Recently, the turnout on these social networks is increasing every day [2]. The eBusiness Knowledgebase listed the top 15 most popular social networking sites in February 2015 [3]. It shows that Twitter ranked second among other popular social

A.A. AlSanad · A.A. Mirza
College of Computer and Information Sciences, King Saud University,
Riyadh, Saudi Arabia
e-mail: amirza@ksu.edu.sa

A.A. AlSanad(✉)
College of Computer and Information Sciences,
Al Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia
e-mail: Abeer.AlSanad@ccis.imamu.edu.sa

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_12

networks worldwide. In Saudi Arabia the most popular social networks used are WhatsApp, Facebook and Twitter [4]. WhatsApp is considered the most used social network in Saudi Arabia. However, Twitter is the fastest growing social networking platform, representing the fastest-growing Twitter market in the world with a growth rate of more than 3,000% from 2011 to 2012 [5]. This refers to the expected increase in the trend of using Twitter in Saudi Arabia in the coming years. As a result, we have decided to focus our study specifically on Twitter usage in Saudi Arabia.

Twitter is a micro-blogging site that emerged in 2006 and has become a social phenomenon [6]. Twitter users exceeded 284 million users worldwide [7]. In 2014, Saudi Arabia had around 2.4 million active users [8]. According to an official definition from Twitter, an "active user" is someone who logs in, but does not necessarily tweet, once a month [9]. Besides the proliferation of using social networks, online marketing has dramatically increased over the past few years. One form of online advertisements relies on observing users behavior by checking which ads have mostly attracted them in the past. Then predictions for users' interest can be made to recommend ads for them. Social networks shorten the link by allowing users to declare their interests through social connections [10]. Advertisement through social networks has widely opened a new channel of commerce and for reaching buyers. Advertising through social networks has two advantages. The first is that items are more trusted if recommended by friends, relatives or neighbors. Usually, friends tend to the same products since their interests are similar. The second advantage is the ability to easily acquire the interests of social network users via observing their activities [11]. According to Faraj [12], 75% of interviewed businesses claimed that they use social media mainly for promoting and advertising.

There are some websites that are mainly focus on advertising through Twitter like Wingsplay. Wingsplay website connects viral video advertisers with influential social media users. Wingsplay pays their users for sharing video ads through social networks, Twitter or Facebook. Their users are paid each time someone from their social network clicks on the user's link and watches the video. But what is interesting with Wingsplay is that the payment is differing based on the user's influence in the social network. The more influential you are the more money you get [13]. Also, advertising companies like Sponsored Tweets and Magpie are using Twitter users' tweets to spread their advertisements. Users are paid for every ad they make a tweet for. The payment depends on the number of users' followers. The more the followers, the more money a user can gain per tweet [14]. Recently, social media has become a very serious business tool in Saudi Arabia. Since very little research has focused on advertising through Twitter users, we are motivated to focus on this field. This study draws upon the investigation in the case of advertising through Saudi Twitter users with high number of followers. In this work, a survey was conducted to gather information and results were analyzed to determine the level and effectiveness of such activities.

2 Related Works

First, there has been much work in online advertising in general, and advertising in social networks in particular. However, to the best of our knowledge there has not been much work focusing on discussing advertisements through Twitter users that have large number of followers.

Michaelidou et al. in [15]measured the effectiveness of Social Networking Sites (SNS) as a marketing tool. They found that 27% of the business-to-business (B2B) in Small and medium-sized enterprises (SMEs) were using SNS for marketing purposes. From this percent, around 55% are using Twitter as a platform for advertising.

Cha et al. [16] focused their research on measuring user influence in Twitter. They determined that a user's influence is measured by three factors which are in degree, retweets, and mentions. In degree influence is the number of user followers. Retweet influence is the number of retweets a user makes. Mention influence is the number of comments a user made on other tweets. The more you have of these factors, the more you have influence in Twitter. The more you are influential, the best you are in disseminating the advertisement.

Alqahtani et al. [17] studied the influence of Twitter advertisements on university students in Saudi Arabia. They used data collection of 451 university students in Saudi Arabia. Results show that Twitter advertisements had highly significant impact on students' purchase intention. Also, it was fund that there was no difference between genders on the influence factor on the purchase intentions.

3 Methodology

This section outlines the research methodology for conducting this study. The objective of this study is assessing the extent of making advertisements through the most popular Twitter users in Saudi Arabia. Data was obtained through a structured survey that was completed by the most popular Twitter users in Saudi Arabia based on the number of followers. This study examined whether Saudi Twitter users are getting offers to make advertisements, or not? Also, how many offers do they receive? Moreover, are they paid for advertising? If yes, how much? We determined the most popular Twitter users in Saudi Arabia according to Twitaholic [18] and asked them to participate in answering the questionnaire.

Participants:
Fifty persons of the most popular 100 Twitter users in Saudi Arabia have participated in answering the survey questionnaires. The sample contains both genders male and female. In addition, the sample participants come from different backgrounds. These backgrounds are religious, political, cultural, athletic, art and others. Moreover, they differ in age ranging from 18 to above 55. All of them have 100,000 followers or more.

Procedure:
Since the sample taken in this study is most popular Saudi Twitter's users, we contacted them through Twitter and asked them to fill out the survey.

4 Results

This section presents the data collection of the survey distributed to the most popular Saudi Twitter users based on the number of followers. The results in Figure 1 show that from the 50 responses gained; 40% were 25 to 34 year-old, 30% from 18 to 24, 16% from 35 to 44, 10% from 45 to 54, 4% above 55, and none less than 18. Moreover, 55% of the participants were females (Fig. 2). The participants also represented different categories of the society with different confessed backgrounds. 28% of the sample participants in the study have indicated that they consider themselves to be of a religious background, while 22%indicated

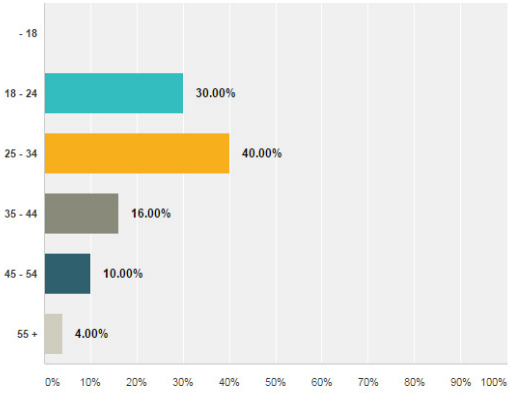


Fig. 1 The distribution of the sample age.

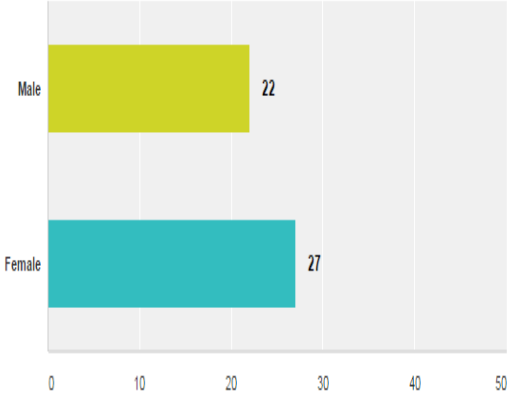


Fig. 2 Sample gender distribution

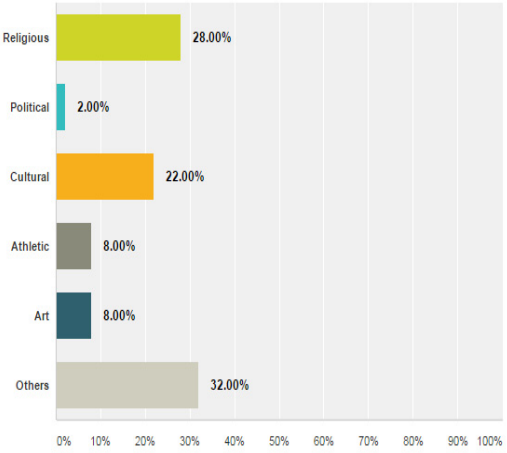


Fig. 3 The sample's backgrounds

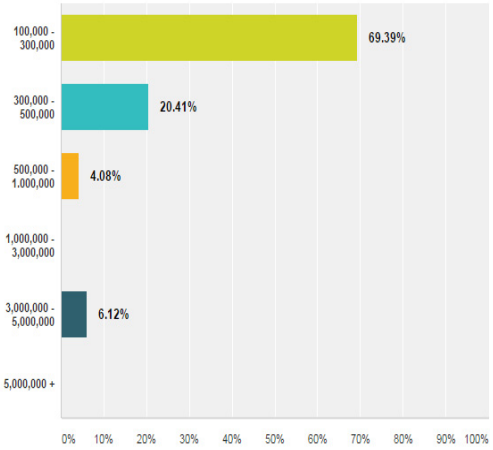


Fig. 4 Distribution of the number of followers

that they have a cultural background. 8% of the participants claimed to be artistic, and the same percentage considered themselves to be athletic. 32% referred to "Other" when asked about their backgrounds (Fig. 3). Furthermore, up to 69.4% of the participating people had followers ranging from 100,000 to 300,000, and 20% of the participants had 300,000 to 500,000. Those who had followers from 500,000 to 1 million and from 3 Million to 5 Million were 4% and 6% respectively (Fig. 4).

Figure 5 represents the percentage of the participants who got offers to advertise. The data shows that almost 49% of the participants got such offers. Out of those 49%, nearly 55% participants had 1 to 20 offers. 20 to 50 advertisements were offered to around 25% of those contacted for advertising purposes. Those receiving 50-100 advertisement requests were near 17%. Finally, only about 4% received between 100 and 500 requests (Fig. 6). Types of products offered for advertisement varied as presented in Figure7. Food represented 24% of the overall products, followed by tourism ads at 12%. Both electronic and sporting goods were equally advertised at 8% each. This was followed by cars and fashion,

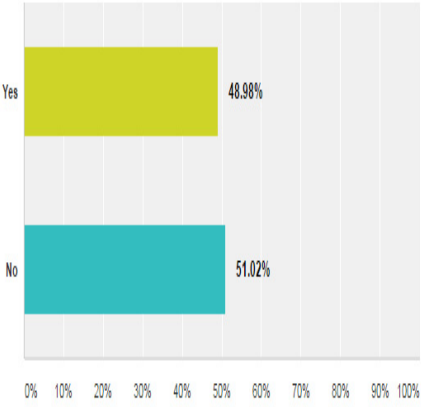


Fig. 5 Whether the participant got offers for advertising or not.

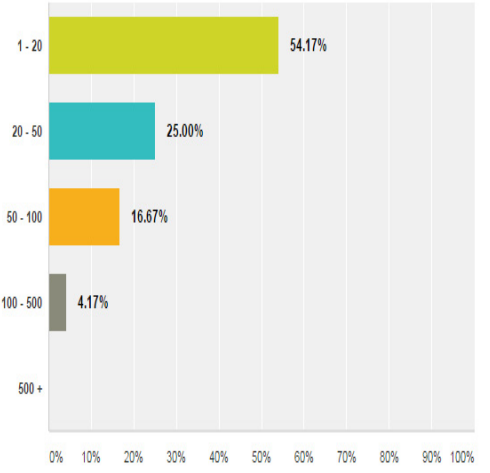


Fig. 6 Number of offers they got.

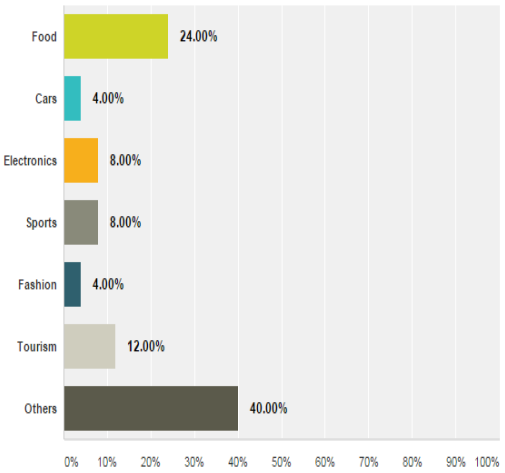


Fig. 7 Type of Offers

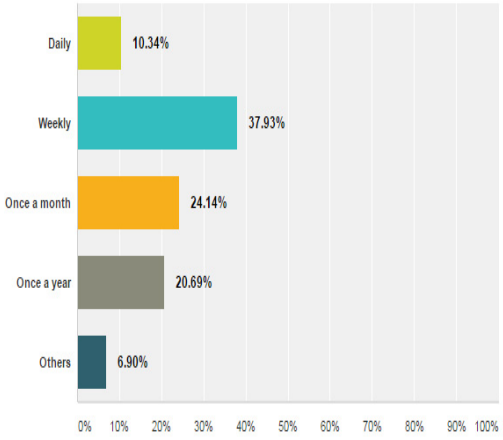


Fig. 8 Frequency of advertisement offers

each at 4%. Lastly, 40% of the advertised products were not among the above listed categories. Almost 38% of those receiving advertisement requests received them on a weekly basis. Twenty four percent received an offer once a month, near 21% received requests once a year, about 10% on a daily basis, and nearly 7% did not fit into any of the above categories. (Fig. 8). Figure 9 shows that 60% from those who accepted advertisement did not gain any financial rewards in return. Among the other 40% who gain money for advertising, 25% received over SR 2000, while another 25% received between SR 1000 and SR 2000 (Fig. 10).

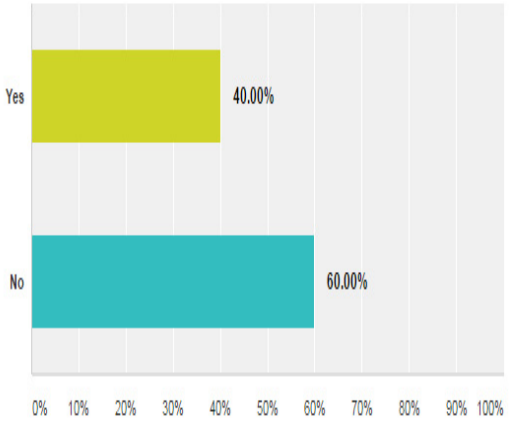


Fig. 9 Payment for advertising

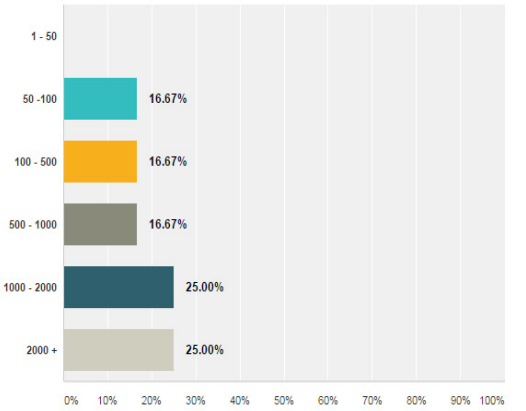


Fig. 10 Amount gained from advertising

5 Discussion

Table 1 has been designed to illustrate the relationship between the variables among those who received advertisement offers. The total number of participants who got offers was 24 (48%), most of them were males (70.8%), and 66.6% were below 35 year of age. The majority of them (91.6%) had less than 500,000 followers; however, this may not reflect the reality due to the small number of participants with followers over 500,000, which were only 5 participants. Looking to the background of the studied group, we noticed that 37.5% of them referred to religious backgrounds, while 20.8% referred to "Other" to express their background culture. The remaining 42% were almost evenly distributed among

athletic, artistic, and cultural backgrounds. Most offer types were from food products companies (33.3%), while tourism constituted 12.5% of the offer types. Electronics, sports, and fashion offers were the least in term of percentages. One third (33.3%) of the participants had offers from "Other" types of products. We still can observe a trend of getting offers to those who had religious backgrounds. The amount of money and number of offers were not significantly affected by age, gender, number of followers, or the religious background; while the chance of getting an advertisement offer is higher among male gender with religious backgrounds.

Table 1 The relationship between the variables among those who received advertisement offers. *Number and percentage of participants who got offers are shown in the table.

Age		Gender		# of followers		Backgrounds						Offers Type							
< 35	> 35	Male	Female	< 500,000	> 500,000	R	P	C	Ath	A	O	F	C	E	S	Fa	T	A	O
16 (66.6%)	8 (33.3%)	17 (70.8%)	7 (29.2%)	22 (91.6%)	2 (8.4%)	9 (37.5%)	1 (4.1%)	4 (16.6%)	3 (12.5%)	2 (8.3%)	5 (20.8%)	8 (33.3%)	0	2 (8.3%)	2 (8.3%)	1 (4.1%)	3 (12.5%)	0	8 (33.3%)

***R:** Religious, **P:** Political, **C:** Cultural, **Ath:** Athlete, **A:** Art, **O:** Others, **F:** food, **C**: Cars, **E:** Electronics, **S**: Sport, **Fa:** fashion, **T:** Tourism.

6 Conclusion and Future Work

In conclusion, a key goal of this study is to determine to which extent are advertisements made through the most popular Saudi Twitter users based on the number of followers. In this paper, a survey method was used to collect the data and the results drawn from this survey were presented and analyzed. The results show that the chance of getting an advertisement offer is higher among male gender with religious backgrounds.

References

1. Mao, E., Zhang, J.: What drives consumers to click on social media ads? The roles of content, media, and individual factors. In: 48th Hawaii International Conference on System Sciences (2015)
2. Hari, K.K.K., Rajan, S.P.: A clubbing of e-commerce and social networking sites. In: International Conference on Ubiquitous Computing and Multimedia Applications (2011)

3. E. Rank. http://www.ebizmba.com/articles/social-networking-websites
4. Statista, Penetration of leading social networks in Saudi Arabia. http://www.statista.com/statistics/284451/saudi-arabia-social-network-penetration/
5. Bennett, S.: The State Of Social Media In Saudi Arabia. http://www.adweek.com/socialtimes/social-saudi-arabia/475292
6. Inouye, D., Kalita, J.K.: Comparing twitter summarization algorithms for multiple post summaries. In: International Conference on Social Computing, Boston, MA (2011)
7. Statista. www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users
8. Abdurabb, K.T.: Saudi Arabia has highest number of active Twitter users in the Arab world. http://www.arabnews.com/news/592901
9. Arab Social Media Report, Dubai School of Government, vol. 1(2) (2013)
10. Parekh, R., Bagherjeiran, A.: Combining behavioral and social network data for online advertising. In: IEEE International Conference on Data Mining Workshops (2008)
11. Karimzadehgan, M., Agrawal, M, Zhai, C.: Towards advertising on social networks. In: 8th Americas Conference on Information Systems, USA, Boston (2009)
12. Faraj, A.: Desining a social media plan for food business in Saudi Arabia in Ball State University, Muncie, Indiana (2014)
13. Wingsplay. http://wingsplay.com/
14. Agarwal, A.: How Much Money Do People Make on Twitter. http://www.labnol.org/tech/money-from-twitter-advertisements/12770/
15. Michaelidou, N., Siamagka, N.T., Christodoulides, G.: Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands. Industrial Marketing Management **40**(7), 1153–1159 (2011)
16. Cha, M., Benevenutoz, F., Haddadi, H., Gummadi, K.P.: Measuring user influence in twitter: the million follower fallacy. In: International AAAI Conference on Weblogs and Social (2012)
17. Alqahtani, S.S., Al-Homoud, H., Al-Otaibi, M.: The impact of twitter advertisement on university students purchase intentions. Pezzottaie Journals **3**(4), 1298–1311 (2014)
18. The Twitaholic.com Top 100 Twitterholics based on Followers in Saudi Arabia. http://twitaholic.com/top100/followers/bylocation/saudi+Arabia/

Sentiment Analysis for Arabic Reviews in Social Networks Using Machine Learning

Mustafa Hammad and Mouhammd Al-awadi

Abstract In this emerging age of social media, social networks become growing resources of user-generated material on the internet. These types of information resources, which are an expansive platform of humans' emotions, opinions, feedback, and reviews, are considered powerful informants for big industries, markets, news, and many more. The great importance of these platforms, in conjunction with the increasingly high number of users generating contents in Arabic language, makes maiming the Arabic reviews in social networks necessary. This paper applies four automatic classification techniques; these techniques are Support vector Machine (SVM) and Back-Propagation Neural Networks (BPNN), Naïve Bayes, and Decision Tree. The main goal of this paper is to find a lightweight sentiment analysis approach for social networks' reviews written in Arabic language. Results show that the SVM classifier achieves the highest accuracy rate, with 96.06% compared with other classifiers.

Keywords Polarity · Sentiment analysis · Text classification · Data mining · Arabic language · Social media

1 Introduction

Finding an automatic way to analyze and classify users' reviews in social networks is very important. This is because it is the most empirical way to get the direct feedback or information from people. For example, finding which reviews are giving positive or negative polarities helps in dealing with customers' behaviors. The process of classifying texts or documents according to their polarity is known as Sentiment Analysis (SA) [1,2].

M. Hammad(✉) · M. Al-awadi
Department of Information Technology, Mutah University, Al Karak, Jordan
e-mail: hammad@mutah.edu.jo

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_13

Sentiment analysis, which is part of Natural Language Processing (NLP), is set to extract the meanings from a text in a way to find the polarity of the text. During sentiment analysis process, it is important to try to maintain data accuracy despite the need for features selection methods. These features are then classified to which polarity this text belongs.

A study done on 2014 [3] estimated that the percentage of websites containing Arabic language contents constitutes around 0.8 % of all the contents on the internet. Users of social networks on the web create around 30% of the Arabic content. However, Arabic language has fuzzy and vague semantics, which makes the process of analyzing Arabic texts hard to be developed in a systematic way. Moreover, Arabic users use informal Arabic language, which is deferent than the formal Modern Standard Arabic (MSA), and can be vary from one to another. Informal Arabic language is the most using language of speaking and communication on social media outlets. Moreover, this spoken Arabic varies, in terms of vocabulary and sentence structure, by country. Unlike English language, the research interest on processing Arabic language texts is started recently.

This paper performs SA on Arabic reviews and comments by using four machine learning algorithms, which are Support Vector Machine (SVM) [4], Back-Propagation Neural Networks (BPNN) [5], Naïve Bayes [6], and Decision Tree [7]. We collected 2000 Arabic reviews from social media in order to evaluate the different machine learning algorithms based on sentiment analysis. In this paper, a particular attention is paid on the preprocessing and features selection for Arabic reviews to optimize the classification process. Our motivation was to evaluate the strengths and weaknesses of these commonly used machine learning methodologies in the field of informal Arabic text sentiment analysis.

The rest of this paper is organized as follows; related works are presented in Section 2. Section 3 describes the used machine learning algorithms. The proposed methodology is presented in Section 4. Section 6 presents the experimental results, followed by the conclusion.

2 Related Work

Sentiment analysis is a research space in which a lot of difficult problems are to be tackled. A number of different approaches have been taken to sentiment analysis. In general, these approaches have relied on one of two techniques: either supervised or unsupervised machine learning. These techniques for sentiment analysis have been principally focused on Indo-European languages, especially English. However, to date, there are few approaches to sentiment analysis social media's texts that have focused on Arabic language.

Santidhanyaroj et al. [8] presented a system that uses analysis of social network data to evaluate public sentiment of issues and actions affecting the general social conscience. They made use of two algorithms: Naïve Bayes and SVM classifiers. Within the model for sentiment analysis, a pre-categorized data set was compiled

from Twitter. The SVM outperformed the Naïve Bayes classifier in analysis and provided more consistent, reliable results.

A sentimental analysis for Arabic language using two popular tools for social analysis: Senti-Strength and Social Mention was proposed by Khasawneh et al. [9]. The data set for this study was taken from Facebook and Twitter posts in Arabic. Based on the contents of the posts, the tools were used to identify polarity of the comments. A comparison of the two tools in ability to guess polarity revealed that the Naïve Bayes gave 91.83% for the Social Mention tool and 95.59% accuracy for Senti-Strength.

Kamal et al. [10] merged the rule-based and machine-learning methods to propose a new sentiment analysis method for identifying polarity. This approach's novelty stems from its use of NLP features and statistical sets, which allow for taxonomy of sentiment at the word-level.

In [11], Abdul-Mageed and Diab exhibited AWATIF a multi-type corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis (MSA SSA). This approach is based on by looking to demonstrate how annotation examines inside subjectivity and assessment dissection (SSA) can both be roused by existing phonetic hypothesis and cater for type subtleties.

Other works, such as [12], used three different datasets pulled from data on Twitter to classify the sentiment of Semitic features using three different methods of analysis: replacement, augmentation, and interpolation. Alaa El-Halees [13] presented another approach for extracting opinions from Arabic text. His combined approach was tri-fold. It consisted of a lexicon-based method for text classification, a training set of the classified texts for Maximum Entropy model, and finally a K-nearest neighbor for classification of the remaining texts.

M. Rushdi-Saleh et al. presented an Opinion Corpus for Arabic (OCA) in [14]. The corpus consists of Arabic text pulled from web pages specifically focused on movies and films. They produced different classifiers using SVM and Naïve Bayes algorithms. A. Al-Subaihin et al. [15] proposed an Arabic language sentiment analysis tool that relies on human-based computing. The tool is beneficial as it aids in the construction, dynamic development, and maintenance of the lexicon.

3 Machine Learning Algorithms

In sentiment analysis, the appropriate type of machine learning to use is text classification, which is known as "supervised learning". This section briefly describes back-propagation neural network, naïve bayes, decision tree, and support vector machine classification methodologies, which are employed in this paper.

3.1 Back-Propagation Neural Networks

Artificial Neural Networks (ANN) simulates human neurons based on building three layers; input layer, hidden layers and output layer. Input layer consists of vectors after applying feature dimensionality reduction, hidden layers is the activity region in

categorization. Output layer represents the categories. There are two types of training the neural networks, which are feed-forward and feedback. Feedback algorithm is designed to reduce the mean square error between the real output of a networks and the desired output. Within the back-propagation algorithm, there is both a forward pass, which gets the activation value, and a backward pass, which adjusts weights biases relative to desired and actual outputs. Until the network converges, the backward and forward passes iterate continuously.

3.2 Naïve Bayes

Naïve Bayes classifier is frequently used for text classification problems, which is a simple probabilistic on the so-called Bayesian theorem. Bayes' classifier is particularly suited when the dimensionality of the inputs is high. It assumes that a document's features are conditionally independent of one another. Given a document T_i; the probability that this text is belong to a class C_j, noted as $P(C_j| T_i)$, is calculated as: is calculated as:

$$P(Cj|Ti) = \frac{P(\{Ti|Cj\})*P(Cj)}{P(Ti)} \quad (1)$$

where $P(T_i |C_j)$, $P(C_j)$, and $P(T_i)$ is the probability of text Ti is in class C_j, class C_j occurrence, text Ti occurrence respectively in the training set. Bayes' classifier classifies the target text to the class that has the highest probability value among all others classes.

3.3 Decision Trees

Decision tree [23] is one of the most common techniques for classification. The decision tree is composed of nodes that are connected with branches. The leaf nodes are the final decision that categorized the target input to a specific class. All inner nodes are set to split the classes into subsets of classes according to a set of selected features checks.

Designing the decision tree is done by analyzing the training data to find the discriminatory features. The most significant features are placed at the top of the tree, and recursively add the next features checks in the following nodes levels until we reach the tree's leafs, which labels the target text to a category.

3.4 Support Vector Machine (SVM)

SVM is based on finding the best hyper plane that divide the data into two categories with the largest possible margin. In this linear classification, the hyper plane $y = w\ f(x) + b$ Where w is the weights vector, b is the bias, and f(x) denotes the mapping function from input feature space with maximum margin. SVM can be used to create boundaries among categories in case of nonlinear classifications. In this case, the data is mapping into another space, which SVMs perform the linear algorithm over it.

4 Sentiment Analysis Model

The proposed sentiment analysis model consists of two basic phases; training the classifier with a set of reviews, and then testing it with another set. However, data sets in both phases need to be preprocessed before the classification process. The preprocessing step, as in any text classification, has a huge impact in the classification results. Figure 1 describes the proposed approach.

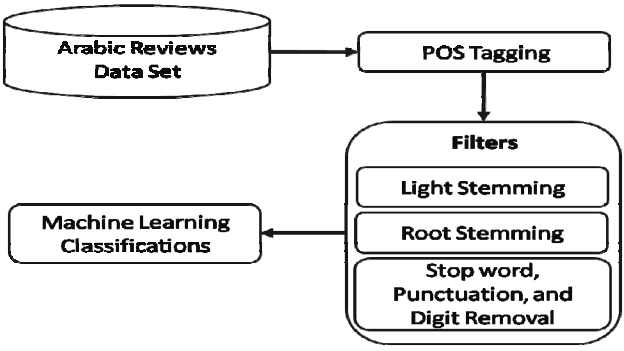


Fig. 1 Overview of the proposed sentiment analysis approach.

The first step is to categorize a review's words to recognize its structure. This is known as parts of speech (POS) tagging. The main goal for POS tagging is to identify all the structural features of a sentence, such as verbs, nouns, adjectives, and adverbs. We used well defined Java code for tagging the various parts of speech in Arabic reviews.

The second step in the proposed model is set to filter unimportant features in the data. Stemming is a morphological analysis technique for text processing. Light stemming removes some common affixes, which include definite article prefix and pluralizing suffix. On the other hand, root stemming reduces the words to only their stems.

In the final step, before applying machine learning algorithms, stop words that do not affect the polarity of a text are removed. In addition to that, punctuations and digits are removed by using their Unicode representations.

5 Experimental Results

To evaluate Machine learning algorithms on Arabic reviews sentiment analysis we used the RapidMiner tool[1], which is an open source java-based environment for machine learning and data mining. RapidMiner creates XML files that contain the user-defined operations to be applied to the data. There are many operators on the RapidMiner tool for different machine learning processes, such as preprocessing

[1] http://rapidminer.com

and visualizing data. We found that the RapidMiner tool can accommodate the Arabic language easily.

To evaluate the proposed approach, the experiments is built based on dividing the labeled data into two groups; the first group is set for tanning the machine learning algorithms, which is 70% of the collected data set, and the second group, which is the remaining 30% of the data set, is set to evaluate and test the classifiers.

5.1 Collected Data Set

In order to apply machine learning classifications, a set Arabic reviews should be collected to train and test the classifiers. The proposed data set is collected from Jordanian hotels' customers' reviews on the internet. The collected data set is a combination of Arabic reviews and comments from Facebook, Twitter, and YouTube. The total number of collected reviews is 2,000. Table 1 shows the total number of reviews for each category of the three polarities; positive, negative, and neutral. The collected data set is has a lot of informal Arabic and vernaculars.

Table 1 Number of reviews by polarity.

Category	Number of Arabic Reviews
Positive	1003
Negative	593
Neutral	404

5.2 Performance Measures and Results

In our experiments, we have used the most common performance measures that include recall (R), precision (P), and F-measure (F). Given a test set of reviews S1 that are labeled to the polarity k, and a prediction set S2 that is labeled with polarity k by the machine learning algorithm, the recall (R) and precision (P) measures are defined by:

$$R = \frac{S1 \cap S2}{S1} \tag{2}$$

and

$$P = \frac{S1 \cap S2}{S2} \tag{3}$$

respectively [16]. To compare the different algorithms with a single rate, we used the F-measure, which combine recall and precision. F-measure that is used in our experiments is defined as:

$$F - measure = \frac{2RP}{R+P} \tag{4}$$

The experiments' results for applying sentiment analysis for the proposed data set using SVM, BPNN, Naïve Bayes, and Decision Tree classifications algorithms are shown in Table 2 in term in terms of Precision, Recall and F-measure.

As shown in Table 2, SVM achieved the highest accuracy, which is 96.06%, followed by Naïve Bayes with average accuracy of 88.38%, and Decision Tree with average accuracy of 85.82%. BPNN was the lowest average accuracy of 69.77%. Moreover, SVM has the highest precision and recall values.

Table 2 Performance results for each classifier.

Classifier	Precision	Recall	F-measure
SVM	95.80%	96.40%	96.06%
BPNN	72.14%	67.61%	69.77%
Naïve Bayes	92.62%	84.99%	88.38%
Decision Tree	83.98%	87.99%	85.82%

In order to test the learning abilities we setup five different experiments with 300, 600, 900, 1200, and 1500 training reviews set. Figure 2 shows the F-measure value for each classifier with these five training sets. As shown in Figure 2, the accuracy of SVM, Decision tree, and Naïve Bayes classifiers increase as long as we increase the training data. However, BPNN does not learn reasonably well. We believe this is because the ambiguity of the data set, which affect the behavior of the neural network.

Fig. 2 F-measure value versus number of training reviews.

Figure 3 shows the average training time in seconds for each classifier. As shown in the figure, Decision tree classifier needs the largest amount of training time compared to other three classifiers. SVM has the smallest training time, which is 6.45 seconds. Naïve Bayes has also a small training time with 11.41 seconds.

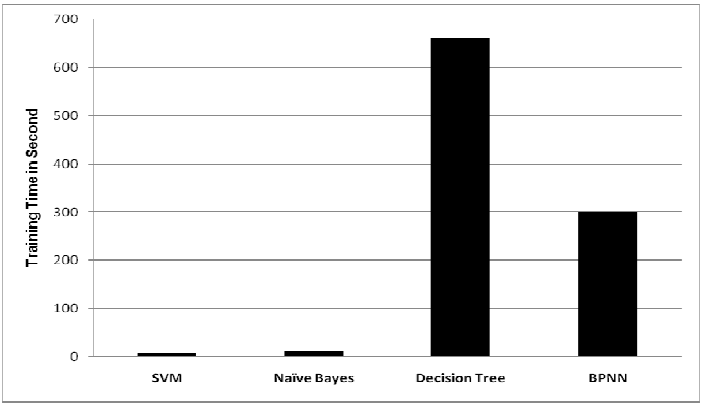


Fig. 3 Average training time for each classifier.

6 Conclusion

This paper applies sentiment analysis on Arabic reviews, which are collected from social media using machine learning algorithms. Experimental results showed that the proposed preprocessing technique for Arabic reviews allow different machine learning algorithms to be applied in the extracted feature for the target data set. In this paper, we used four different classification algorithms to analyze the polarity of sentiments for Arabic reviews. Results showed that the most accurate machine-learning algorithm, based on our approach, was SVM with F-measure equal to 96.06%.

Sentiment analysis of Arabic language requires a lot more research especially in the preprocessing phase. As future works, larger and more diverse data sets can be used. It is also important to increase the size of the dictionaries and remedy the memory problem in RapidMiner.

References

1. Nasukawa T., Jeonghee, Y.: Sentiment analysis: capturing favorability using natural language processing. In: Proceedings of the 2nd International Conference on Knowledge Capture. ACM (2003)
2. Dave, K., Lawrence, S., Pennock, D.M.: Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In: Proceedings of the 12th International Conference on World Wide Web. ACM (2003)
3. Usage of content languages for websites. http://w3techs.com/technologies/overview/content_language/all
4. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning **20**, 273–297 (1995)
5. Goh, A.T.C.: Back-propagation neural networks for modeling complex systems. Artificial Intelligence in Engineering **9**(3), 143–151 (1995)

6. Chen, J., Huang, H., Tian, S., Qu, Y.: Feature selection for text classification with naïve bayes. Expert Systems with Applications **36**, 5432–5435 (2009)
7. Quinlan, J.R.: Induction of decision trees. Machine Learning **1**, 81–106 (1986)
8. Santidhanyaroj, P., Khan, T., Gelowitz, C.M., Benedicenti, L.: A sentiment analysis prototype system for social network data, pp. 1–5 (2014)
9. Khasawneh, R.T., Wahsheh, H.A., AI-Kabi, M.N., Aismadi, I.M.: Sentiment Analysis of Arabic Social Media Content: A Comparative Study, pp. 101–106 (2013)
10. Kamal, A., Abulaish, M.: SMIEEE.: Statistical Features Identification for Sentiment Analysis using Machine Learning Techniques, pp. 178–181 (2013)
11. Abdul-Mageed, M., Diab, M.: AWATIF: a multi-genre corpus for modern standard arabic subjectivity and sentiment analysis. In: Proceedings of LREC, Istanbul, Turkey (2012)
12. Elhawary, M., Elfeky, M.: Mining arabic business reviews. In: 2010 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 1108–1113. IEEE (2010)
13. El-Halees, A.: Arabic opinion mining using combined classification approach. In: Proceedings of the International Arab Conference on Information Technology, ACIT (2011)
14. Rushdi-Saleh, M., Martín-Valdivia, M., Ureña-López, L., Perea-Ortega, J.M.: Bilingual Experiments with an Arabic-English Corpus for Opinion Mining (2011)
15. Al-Subaihin, A., Al-Khalifa, H., Al-Salman, A-M.: A proposed sentiment analysis tool for modern arabic using human-based computing. In: Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services. ACM (2011)
16. Rijsbergen, C.J.V.: Information Retrieval: Butterworth-Heinemann (1979)

Part II
Security and Privacy in Next Generation Networks

An Interactive Model for Creating Awareness and Consequences of Cyber-crime in People with Limited Technology Skills

Sheeraz Akram, Muhammad Ramzan, Muhammad Haneef and Muhammad Ishtiaq

Abstract Technology is being used every day by the people irrespective of their technology skills. People with limited technology skills are not aware of what actions are legal and what actions are illegal according to the laws. In this paper, an interactive model for creating awareness and consequences of cyber-crime in people with the limited technology skills is proposed.

Keywords Regulator · Media · Technology skills · Cyber-companies · Social websites

1 Introduction

The 21st century is the era of information technology. In past people have different mechanism to share the information which were very slow, expensive, and unreliable and not enough to fulfill needs of all users. The people use to sit physically in their social circles for the social activities. For education purposed, they physically gathered at some location like schools and colleges. They use to have different games and it was necessary for all the participants to be present at one location for playing the game. The messages were sent by writing them on the

S. Akram(✉) · M. Haneef · M. Ishtiaq
Foundation University Rawalpindi Campus, Foundation University Islamabad, Islamabad, Pakistan
e-mail: {sheeraz,muhammadhaneef,ishtiaq}@fui.edu.pk

M. Ramzan
College of Computing and Information Technology,
Saudi Electronic University, Riyadh, Saudi Arabia
e-mail: m.ramzan@seu.edu.sa

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_14

pages and later sent through different means like postman; to be a carrier. The money was kept in the form currency notes at the homes and banks.

Currently, the internet is medium to pass any kind of information from one location to another location. The social websites are available for people to create their social circles and friendships. The messages are sent from one place to another place by emails and different chat applications. Now people can play games by sitting on different locations and can compete with each other. The money is transferred online from one location to another location. All this is achieved by internet which is the backbone of current technology.

The world of internet is known as cyberspace. This world has its methods of communication, financial transaction, sports and developing social circles. In the cyber space, some actions are legal and some actions are not legal. The illegal actions in the cyberspace are known as cybercrime. There are people who are aware of what they are doing in the cyberspace, so they are well aware of the status of their actions whether legal or illegal. But there are great number of people who are illiterate with reference to the usage and working of cyberspace and they do illegal actions unknowingly. There is need to create awareness about cybercrime and it consequences especially for the people with limited literacy and knowledge regarding rules and working in cyber world. The related work is given below.

In [1], the current status of cybercrime related to financial based crime, non-financial matters, the threat to public and business are given and strategies again mentioned crimes are discussed. The conventional crimes, cyber-crimes, different models of committing cyber-crimes and cyber-crime prevention strategies are given in [2]. The basic measures required to curb the cybercrimes and spamming activities in Kenya are discussed in [3]. In [4], the current concepts and strategies of cybercrime and cybersecurity, possible elements of cybercrime policies and strategies are discussed. The motivations behind the cybercrimes, analysis of behavior of cybercrimes and the impact of cybercrime on the society are discussed in [5]. In [6], the phenomenon of cybercrime, topology of cybercrime, challenges of fighting cybercrime and anti-cybercrime strategies are discussed. The awareness level about cybercrimes and cyber law in Bangladesh is given in [7]. In [8], the impact of cybercrime on business and losses due to these crimes are presented. The cybercrime and what should do when you become victim of cybercrime is described in [9]. In [10], the cybercrime measurements, legislation and framework to handle cybercrime, investigation of cybercrime and electronic evidence required for handling cybercrime are discussed. In [11], the components of malicious activities are briefly described. The cybercrime related to identity theft, the nature and type of cybercrime and consequences of cybercrime in tertiary institution are discussed in [12]. The cybercrime in banks, the actors involved in cybercrime and its impact on banking finance is described in [13]. In [14], the existence of law against cybercrime and impact of cybercrime on the society in Nigeria is presented.

There is very less work related to creating awareness about cybercrime and consequences of cybercrime for the people who are not familiar with the technology.

So in this paper, an interactive model for creating and increasing awareness about cyber-crimes and its consequences for people with limited technology skills and literacy is proposed.

2 Proposed Model

The proposed model is to create awareness of cyber-crime and its consequences in the common public especially with low skills in technology. The components of proposed model are Government, Regulators, Media, Cyber Companies and User with Limited Technology skills. All the components collaborate with each other and focus is on User with Limited Technology Skills as given in Fig. 1. The purpose is to create awareness about cyber-crime and it consequences for the user having limited technology skills.

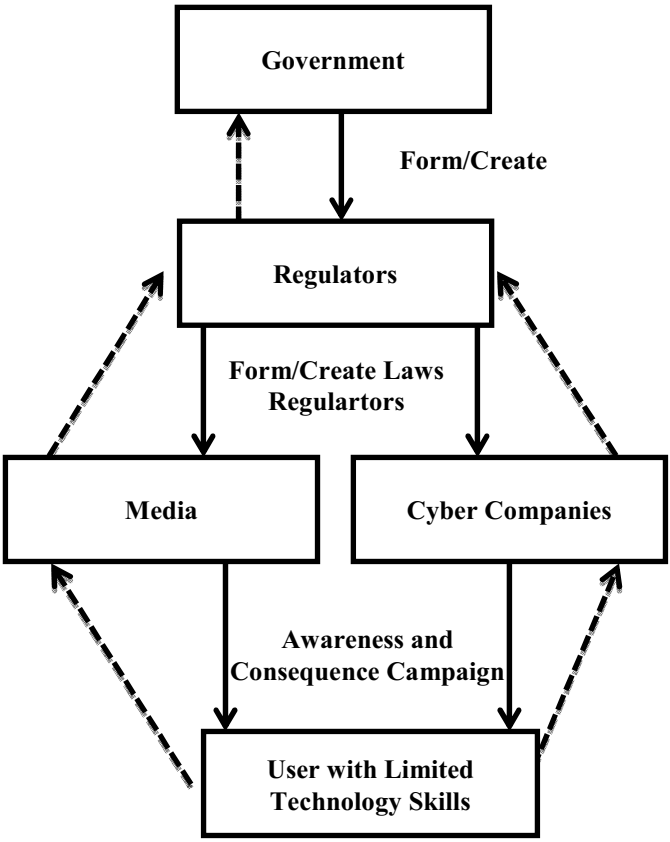


Fig. 1 Proposed Interactive Model

3 Roles of Components of Proposed Model

The components in the proposed models interact with each other. Each component provides feedback to the top component.

3.1 Government

The government is most important component of this model. The policy of the government depict about the rules and regulation for usage of any technology. Government is not a technical entity which is expert about everything; rather the role of government is produce/form certain institutions to look after different matters. The government work under set of rules and same rules allow government to form different agencies/regulators for the support of government. The agencies role is to help and make sure the proper implementation of orders given by the government. The Regulators are which takes response from the government as well as user too, so the regulator hold important place in proposed model. The Regulator provides feedback to the Government. The policies are modified according to feedback provided by the Regulators.

3.2 Regulator

The regulator is formed by the government using power, inherits at the time of formation. The regulator works according to policy given by the government. In our model, the component Regulator is working to control Media (Electronic & Print) and another Regulator control Cyber companies, for example the internet services provider, social websites etc as given in Fig. 2. The Regulator works on policy given by Government and makes rules for Media and Cyber companies. The Regulator advises them to prepare and launch campaign for creating and increasing awareness of cyber-crimes and its consequences in the people with limited technology skills.

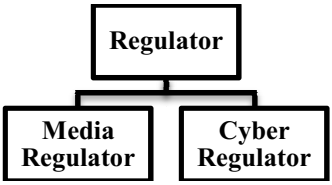


Fig. 2 Types of Regulator

The Regulator also takes the feedback from the Media and Cyber companies and presents it to the Government. The Government modifies the policy for user with limited technology skills based on feedback it receives from the Regulators. There is separate Regulator for Media and Cyber companies. The print media exist since long but electronic media got fame in last two to three decades.

At the same time the Cyber companies are more complicated and provide greater facilities and functionalities. So, the Regulator for the Cyber companies must be equipped with more power and vision.

3.3 Media

The Media is an important component in our proposed model. The types of media are given in Fig. 3. The print media is working since years in the form of newspaper, magazine etc. This is basic mean to convey any news and information to user who can read little despite having limited access to the technology. The print media produce and print such advertisements in the newspaper and magazine which guide the reader regarding the usage of technology. In the proposed model, the role of print media is very important as the impact of print media is long lasting. In case of any miss-use of the technology, where the user has to contact, should also be mentioned clearly in those advertisements. In proposed model, the news of cyber-crime is reported in way which mention that the person who has done crime is a technology skilled person or non-skilled. Also the crime should be reported in way which could be understood by person with limited technology skills.

The electronic media works through Television. The electronic media use the audio for any news along with visual aid. In the proposed model, the electronic media is instructed to show such programs which make people comfortable in using technology, especially people with limited technology skills. The cyber-crime news must be reported in way that the audio is understandable to the person with limited technology skill. The consequences of cyber-crime must be reported with examples so that the person can understand about cyber-crime and its consequences.

The media receives the feedback of the news and advertisements regarding increasing awareness of cyber-crime and its consequences and pass it to Regulators. The Regulator changes the rules for Media and also passes the feedback to the Government in order to make appropriate changes in the policy.

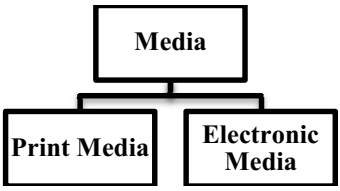


Fig. 3 Media Types

3.4 Cyber Companies

In the proposed model, the Cyber Companies are referring to internet service providers and other websites like email service providers and social websites. The internet service providers keep the track of each internet user through IP address.

In proposed model, the cyber companies must have separate groups of IP address for the technology skilled people and limited technology skilled people. The material must be sent to limited technology skills people in a language that is easily understandable for them. The illegal action from IP address in group of limited technology skill must be given immediate warning and provided guidance toward proper usage of the technology.

The email service providers allow user to send message electronically from one place to another place. In the proposed model, the email service providers collect the data at the time of account creation regarding the level of technology skills and later different ads appears along with emails to increase awareness of cyber-crimes and its consequences.

The social websites are very common now days. The people are using social websites for maintaining their social circle. The user uploaded different kind of material in the form of text, images and video. This material can be offensive to any single person or a community. In the proposed model, only that social website should be accessible which follow the instructions to display material for guiding user about cyber-crime and its consequences.

3.5 User with Limited Technology Skills

The User with Limited Technology Skills is most important component of our proposed model as all the components in model works for the awareness of this component. In the proposed model, the user with limited technology skills must read the advertisements on the media and ads in the email with care and should act on those instructions.

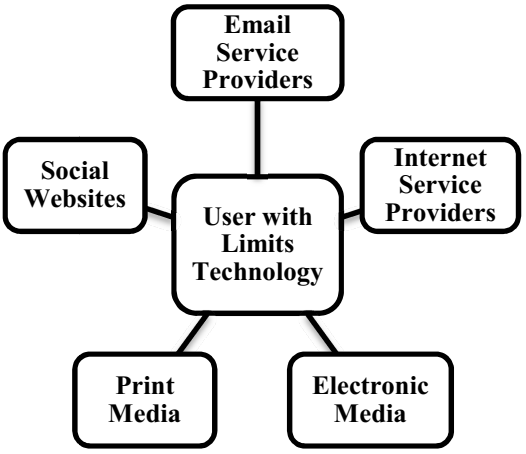


Fig. 4 Interaction of User with Different Components

The user must not perform any action which is not known to him. In the proposed model, the use must provide proper feedback to the Media and Cyber-companies so that they can convey to the Regulators to modify the policies. The limited Technology skilled people keep on learning with proper feedback mechanism in our proposed model. The user held responsible for committing a cyber-crime faces the consequences. The proposed model also suggests the design of consequences which should be developed in way that the user can become more skilled about the technology in the future.

4 Conclusion

The technology is vital in the current time. The people use the technology in their daily activities. The people do actions which are not suitable for the others and sometimes they cross the legal boundaries. This paper proposed an interactive model which shows the interaction between different components of the proposed model. The proposed model focuses on creating the awareness and consequences of cyber-crime in the people with the limited technology skills.

References

1. Department, S.o.S.f.H.: Cyber Crime Strategy. Surrey (2010)
2. Dashora, K.: Cyber crime in the society: problems and preventions. Journal of Alternative Perspectives in the Social Sciences **3**(1), 240–259 (2011)
3. Magutu, P.O., Ondimu, G.M., Ipu, C.J.: Effects of cybercrime on state security: types, impact and mitigations with the fiber optic deployment in Kenya. Journal of Information Assurance & Cybersecurity **2011**(1), 1–20 (2011)
4. Crime, G.P.o.C.: Cyber Crime Strategies (2011)
5. Saini, H., Rao, Y.S., Panda, T.C.: Cyber-crimes and their impact: a review. International Journal of Engineering Research and Applications (IJERA) **2**(2), 202–209 (2012)
6. Sector, T.D.: Understanding CyberCrime: Phenomena. Challenges and Legal Response (2012)
7. Khan, K.A.: Awareness towards cybercrime & cyber law. Insight Bangladesh Foundation (2013)
8. Study of the impact of cyber crime on business in canada. International Cyber Security Alliance (2013)
9. Cruz, A.: Cyber Crime and how it effect you. Cyber Security Tips **7**(1), 2 (2013)
10. Crime, U.N.O.o.D.a.: Comprehensive Study on cybercrime (2013)
11. The economic impact of cybercrime and cyber espionage. Centre for strategic and international studies (2013)
12. Okeshola, F.B., Adeta, A.K.: The nature, causes and consequences of cyber crime in tertiary institutions in Zaria-Kaduna state, Nigeria. American International Journal of Contemporary Research **3**(9), 98–114 (2013)
13. Raghavan, A.R., Parthiban, L.: The effect of cybercrime on a bank's finances. International Journal of Current Research and Academic Review **2**(2), 173–178 (2014)
14. Olayemi, O.J.: A socio-technological analysis of cybercrime and cyber security in Nigeria. International Journal of Sociology and Anthropology **6**(3), 116–125 (2014)

Power Analysis Attack and Its Countermeasure for a Lightweight Block Cipher Simon

Masaya Yoshikawa and Yusuke Nozaki

Abstract This study proposes a power analysis attack and a countermeasure for a lightweight cipher Simon. Simon can be embedded in the smallest area among lightweight block ciphers. In the proposed power analysis method, an analysis based on conventional power analysis attacks is applied to Simon. In the proposed countermeasure, random masks are applied to data resisters. Experiments revealed the vulnerability of the normal implementation method and verified the validity of the proposed countermeasure.

Keywords Power analysis attack · Countermeasure · Lightweight block cipher · Simon

1 Introduction

In built-in apparatuses, the scale of a circuit is limited. Therefore, the development of a lightweight block cipher that can be embedded in a small-scale circuit has been highly anticipated. Simon [1] is a lightweight block cipher disclosed by the United States National Security Agency (NSA) [2], and in comparison to other typical lightweight ciphers [3,4,5], it can be embedded in a very small area [1].

The threat of side-channel attacks against circuits, in which a cipher whose safety is computationally secured has been embedded, is pointed out [6,7,8,9]. Side-channel attacks illegally obtain confidential information using physical information, such as power consumption and electromagnetic waves generated during encryption processing. In particular, since power analysis attacks [6], [7] using power consumption can easily analyze confidential information, they are considered to be the most dangerous types of attacks.

At present, many studies have reported on power analysis attacks against the advanced encryption standard (AES) [10], but few studies have reported on these attacks against lightweight ciphers. In particular, as far as we know, studies presenting measures to prevent power analysis attacks against Simon have not been published yet. The present study proposes a power analysis method against Simon,

M. Yoshikawa(✉) · Y. Nozaki
Department of Information Engineering, Meijo University, Nagoya, Japan
e-mail: dpa_cpa@yahoo.co.jp

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_15

a lightweight block cipher, and it also proposes a countermeasure to prevent power analysis attacks. In the proposed power analysis method, an analysis based on conventional power analysis attacks is applied to Simon. In the proposed countermeasure, random masks are applied to data resisters. The present study also verifies the validity of the proposed power analysis method by performing an evaluation experiment using a field-programmable gate array (FPGA).

2 Preliminaries

2.1 Simon

Simon is a lightweight block cipher that has a Feistel structure. Its block length can be 32, 48, 64, 96, or 128 bits, and its key length can be 64, 72, 96, 128, 148, 196, or 256 bits. The number of rounds changes depending on the selected block length and key length.

When the intermediate value at the *r*th round is expressed as x^r, its left half is processed as x_L^r and its right half is processed as x_R^r, separately. By repeating the left rotation processing, the AND operation, and the XOR operation in a bit unit, encryption is performed. The intermediate value can be calculated using formula (1).

$$\begin{cases} x_L^{r+1} = \left(S^1\left(x_L^r\right) \& S^8\left(x_L^r\right)\right) \oplus x_R^r \oplus S^2\left(x_L^r\right) \oplus RK^r \\ x_R^{r+1} = x_L^r \end{cases} \tag{1}$$

In formula (1), S^1, S^8, and S^2 express the left rotation processing at 1, 8, and 2 bits, respectively; & expresses the AND operation; $\oplus$ expresses the XOR operation; and *RK* expresses a partial key generated form the key-scheduling section.

2.2 Power Analysis

Power analysis is a method that obtains confidential information using the power consumption that is generated during the operation of a cryptographic circuit. Differential power analysis (DPA) is a typical type of power analysis attack.

DPA consists of a hamming weight (HW)-type DPA and a hamming distance (HD)-type DPA. HW-type DPA uses the difference in power consumption that is generated due to the difference in the transition probability when certain input and output values (hamming weight) are noticed in the nonlinear element inside a cryptographic circuit. HD-type DPA assumes that a linear relationship exists between the intra-data resister hamming distance of data and the power consumption, and it uses this linear relationship.

In HW-type DPA, power consumption waveforms are divided into two groups based on whether the hamming weight with certain input and output values is 0 or 1. In HD-type DPA, power consumption waveforms are divided into two groups based on whether the hamming distance between the data resisters is larger or smaller than a predetermined value.

The hamming weight and the hamming distance are obtained using an already known cryptogram, and they are calculated using the predicted value of a partial key that is used during encryption processing. Subsequently, the difference in the average of each group is obtained. The predicted value of a key with the largest differential power value is estimated to be the correct key.

3 Proposed Method

3.1 Analytical Method Using Correction Processing

In the first step of this proposed approach, a basic power analysis method is explained in which HW-type DPA is applied to Simon. As shown in Fig. 1, the hamming weight of input value *A* of the AND operation is actually used.

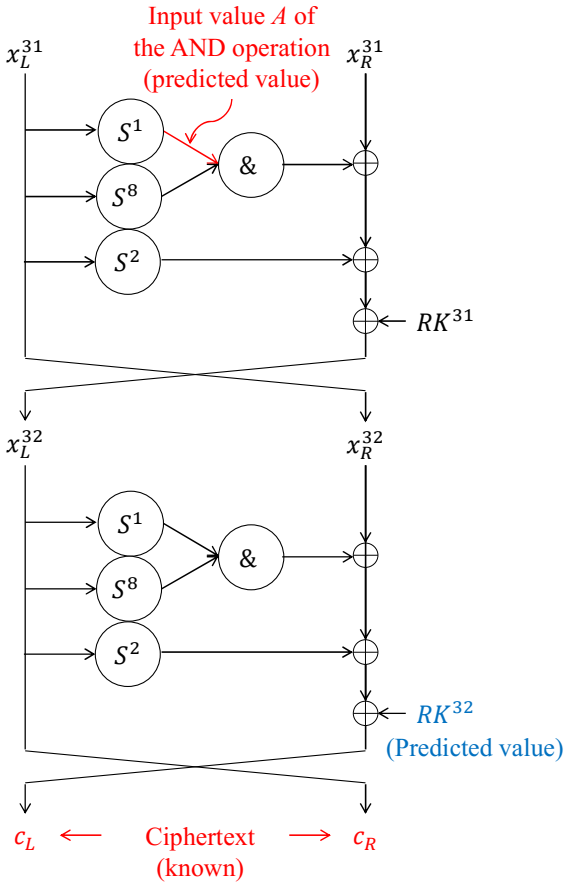


Fig. 1 Attack point

The input value A of the AND operation can be calculated using formula (2).

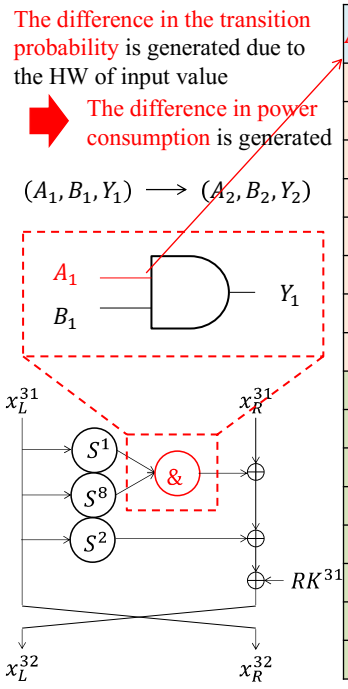
$$A = S^1\left(x_R^{32}\right) \tag{2}$$

In formula (2), x_R^{32} can be calculated as the already-known cryptogram c_R and the predicted value RK^{32} of a partial key. This calculation is expressed as formula (3).

$$x_R^{32} = \left(S^1\left(c_R\right) \& S^8\left(c_R\right)\right) \oplus c_L \oplus S^2\left(c_R\right) \oplus RK^{32} \tag{3}$$

The difference in the transition probability generated in the AND operation of Simon is explained in Fig. 2, which shows the true table of input and output values when a two-input AND gate is used. When the input value of certain time 1 (A_1) is 0, the transition probability of output Y is 1/4. When input value A_1 is 1, the transition probability of output Y is 1/2. Therefore, the difference in the transition probability is generated due to the hamming weight of input value A_1.

In the analysis, power consumption waveforms are divided into two groups based on whether the hamming weight of input value A is 0 or 1. The difference in the average of the power consumption waveforms in each group is obtained, the differential power value is calculated, and the predicted value of a key with the largest differential power value is estimated to be the correct key.



A_1	B_1	A_2	B_2	Y_1	Y_2	Transition	probability
0	0	0	0	0	0	×	
0	0	0	1	0	0	×	
0	0	1	0	0	0	×	
0	0	1	1	0	1	✓	
0	1	0	0	0	0	×	1/4
0	1	0	1	0	0	×	
0	1	1	0	0	0	×	
0	1	1	1	0	1	✓	difference
1	0	0	0	0	0	×	
1	0	0	1	0	0	×	
1	0	1	0	0	0	×	
1	0	1	1	0	1	✓	
1	1	0	0	1	0	✓	1/2
1	1	0	1	1	0	✓	
1	1	1	0	1	0	✓	
1	1	1	1	1	1	×	

Fig. 2 Example of the transition probability generated in the AND operation of Simon

Next, a method to apply HD-type DPA is explained. In the proposed power analysis method, the hamming distance of the right half of the intermediate value is used. As shown in Fig. 3, when the final round is targeted, the hamming distance between cryptogram c_R and intermediate value x_R^{32} is used, which is calculated using the cryptogram, the predicted value of a partial key, and formula (3).

In the proposed power analysis method, the hamming distance is used one bit by one bit, and a partial key is predicted one bit by one bit. Therefore, to estimate a 16-bit partial key, the computational complexity becomes 2^1 x 16 = 32 ways. To analyze a 64-bit partial key, the computational complexity becomes 32 x 4 = 128 ways.

Finally, correction processing is explained. As shown in Fig. 4, the voltage value is slightly shifted in the amplitude direction due to measurement errors in the power consumption waveforms that are obtained using an oscilloscope. Because a lightweight cipher consumes little power, measurement errors have a significant effect on it. Therefore, before applying the proposed power analysis method, the obtained power consumption waveforms are processed so that the effect of the measurement errors can be minimized.

In this correction processing, the average of the voltage values in a region where no voltage variation occur, as shown in Fig. 4, is obtained in each power consumption waveform. Subsequently, all of the sample points of each power consumption waveform are shifted so that the obtained average in each power consumption waveform is the same. This is the correction processing.

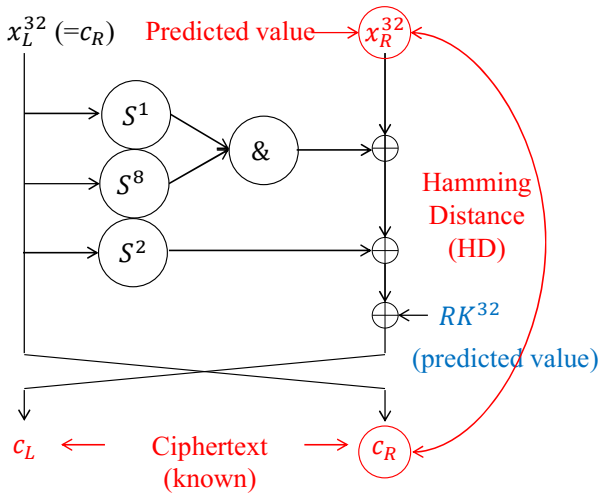


Fig. 3 Example of the hamming distance between cryptogram c_R and intermediate value x_R^{32}

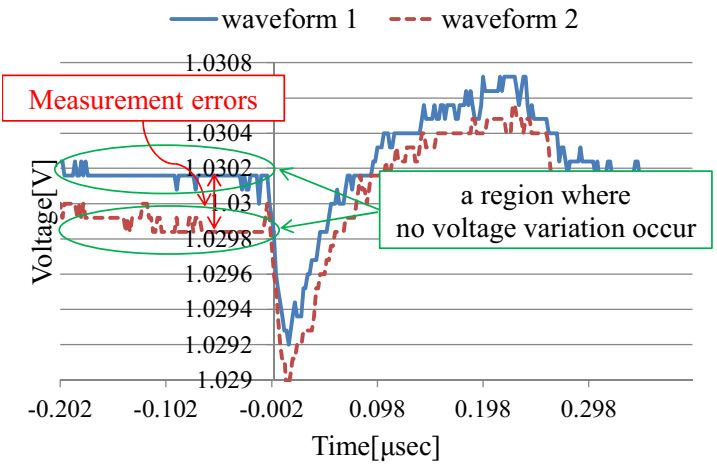


Fig. 4 Example of measurement errors

3.2 Method to Prevent Power Analysis Attacks using Random Masks

In power analysis attacks, the linear correlation between the hamming distance and power consumption is used. The hamming distance is the data transition between data resisters. The proposed method to prevent power analysis attacks pays attention to this linear correlation and ensures tamper resistance by eliminating this linear correlation.

As shown in Fig. 5, mask processing (i.e., the XOR operation) is actually performed for the data resisters used for side-channel attacks. By performing mask processing using random numbers, the values stored in the data resisters differ from the normal intermediate values. Thus, the linear correlation between the hamming distance and power consumption can be eliminated. For the values called out from the data resisters, unmask processing is performed using the same random numbers.

Thus, mask processing is performed immediately before data are stored in the data resisters, and unmask processing is performed immediately after data are read out from the data resisters. Consequently, the correlation between the intra-data register hamming distance and power consumption can be eliminated and correct encryption results can be obtained.

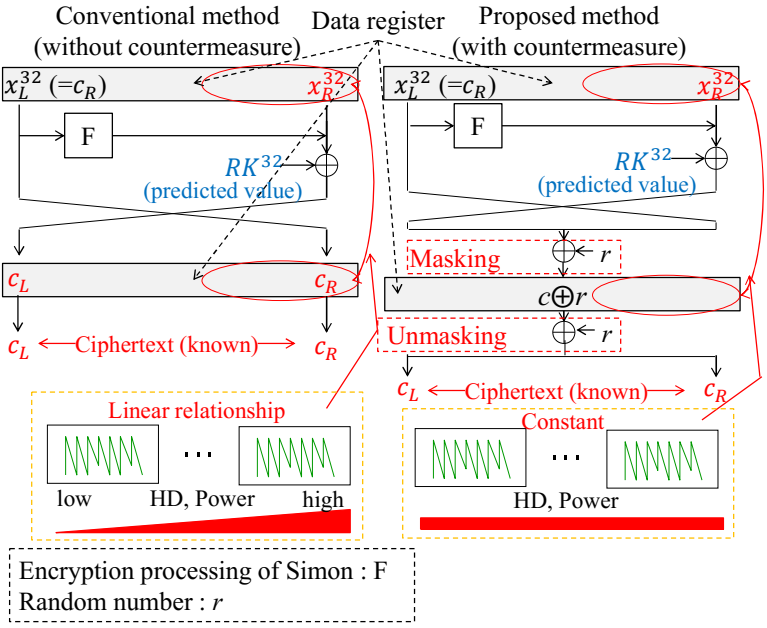


Fig. 5 Mask processing

4 Evaluation Experiments

In the evaluation experiment, Simon with a 32-bit block length and a 64-key length was embedded in an FPGA. Table 1 and Fig. 6 show the experimental environment. Under this experimental environment, encryption processing of Simon was performed 50,000 times, and 50,000 cryptograms and power consumption waveforms were obtained. In the evaluation experiment, a 16-bit partial key RK^{32} at the final round (1-bit partial keys are expressed as RK_1^{32}, RK_2^{32}, and RK_{16}^{32}) was analyzed.

Table 1 Experimental conditions

Encryption algorithm	Simon
Block size [bit]	32
Key size [bit]	64
Evaluation board	SASEBO-GII
FPGA	Virtex-5 XC5VLX30
Implementation tool	Xilinx ISE Design Suite 14.1
Oscilloscope	Agilent DSO1024A
Sampling rate [Gsa/sec]	2
Power supply	USB power supply from PC

Table 2 shows the analytical results obtained using 50,000 waveforms by applying HW-type DPA to Simon. As shown, not all of the partial keys could be analyzed. However, of the 16-bit partial keys, 12-bit partial keys could be analyzed. Therefore, HW-type DPA is effective for Simon. It is considered that when the number of waveforms used for the analysis is increased, the remaining partial key can be analyzed.

Next, the conventional method in Fig. 7 shows the analytical results obtained using 20,000 waveforms by applying HD-type DPA to Simon. The horizontal axis represents the number of waveforms used and the vertical axis represents the number of correct keys, which is expressed by the number of bits of each successfully analyzed partial key. In this experiment, since partial key RK^{32} at the final round was analyzed, the number of the correct keys is 16 [bits], at the maximum. As shown, all of the partial keys could be analyzed using 6,000 waveforms. Therefore, the proposed power analysis method was demonstrated to be effective.

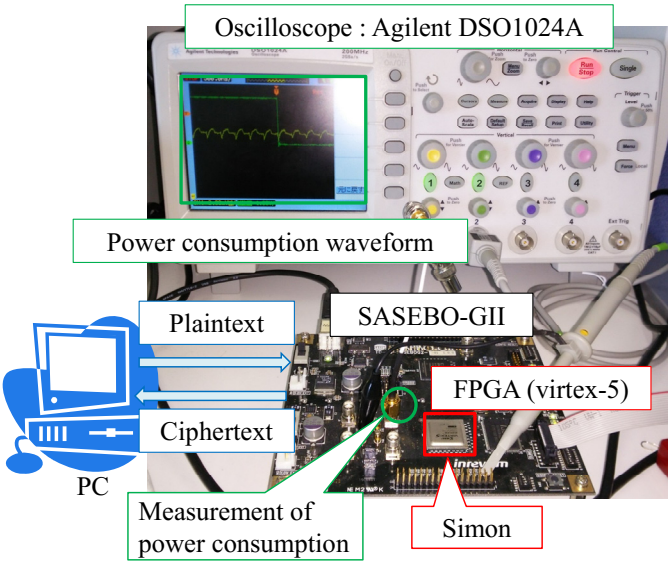


Fig. 6 Experimental environment

Table 2 Analytical results (50,000 waveforms by applying HW-type DPA)

	Success	Failure
Key	RK_1^{32}, RK_3^{32}, RK_4^{32}, RK_5^{32}, RK_5^{32}, RK_6^{32}, RK_7^{32}, RK_{10}^{32}, RK_{11}^{32}, RK_{12}^{32}, RK_{13}^{32}, RK_{14}^{32}, RK_{16}^{32}	RK_2^{32}, RK_8^{32}, RK_9^{32}, RK_{15}^{32}

Finally, power analysis attacks were performed against Simon. In one of the experiments, the proposed method to prevent those attacks was embedded in Simon, and in the other experiment, no countermeasures were taken. When the proposed method to prevent power analysis attacks was embedded, only eight partial keys were successfully analyzed, although 20,000 waveforms were used. Because all of the partial keys consisted of 16 bits, these results are the same as the results obtained by predicting the value one bit by one bit, randomly. Therefore, the proposed method to prevent power analysis attacks is considered to be resistant to power analysis attacks.

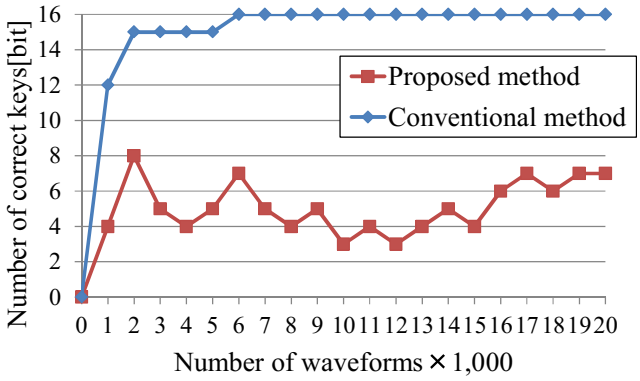


Fig. 7 Comparison results of countermeasure and normal implementation

5 Conclusion

The present study presented two types of power analysis attacks against Simon and proposed a method to prevent those attacks. The evaluation experiment using an FPGA revealed that the proposed power analysis method could attack Simon. Moreover, the proposed method to prevent power analysis attacks was resistant to power analysis attacks. In the future, we will examine whether the proposed method to prevent power analysis attacks is effective against electromagnetic analysis attacks by performing an experiment.

References

1. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: The SIMON and SPECK Families of Lightweight Block Ciphers. Cryptography ePrint Archive, Report 2013/404 (2013) http://eprint.iacr.org/
2. National Security Agency. https://www.nsa.gov/

3. Bogdanav, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B, Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Proceedings of 9th International Workshop Cryptographic Hardware and Embedded Systems (CHES 2007). LNCS, vol. 4727, pp. 450–466. Springer (2007)
4. Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T., Shirai, T.: Piccolo: an ultra-lightweight blockcipher. In: Proceedings of 13th International Workshop Cryptographic Hardware and Embedded Systems (CHES 2011). LNCS, vol. 6917, pp. 342–357. Springer (2011)
5. Suzaki, T., Minematsu, K., Morioka, S., Kobayashi, E.: TWINE: a lightweight, versatile blockcipher. In: Proceedings of ECRYPT Workshop on Lightweight Cryptography (LC11), pp. 146–149 (2011)
6. Kocher, P., Jaffe, J., Jun, B.: Differential power analysis. In: Proceedings of International Cryptology Conference (CRYPTO 1999). LNCS, vol. 1666, pp. 388–397 (1999)
7. Brier, E., Clavier, C., Olivier, F.: Correlation power analysis with a leakage model. In: Proceedings of 6th International Workshop on Cryptographic Hardware and Embedded Systems (CHES 2004). LNCS, vol. 3156, pp. 16–29. Springer (2004)
8. Gandolfi, K., Mourtel, C., Olivier, F.: Electromagnetic analysis: concrete results. In: Proceedings of 3rd International Workshop on Cryptographic Hardware and Embedded Systems (CHES 2001). LNCS, vol. 2162, pp. 251–261. Springer (2001)
9. Meynard, O., Guilley, S., Danger, J.-L., Sauvage, L.: Far correlation-based EMA with a precharacterized leakage model. In: Proceedings of Design, Automation and Test in Europe (DATE 2010), pp. 977–980 (2010)
10. Federal Information Processing Standards (FIPS) Publication 197: Advanced Encryption Standard (AES), U. S. Department of Commerce/National Institute of Standard and Technology (2001)

Perpetuating Biometrics for Authentication

Introducing the Durable True-Neighbor Template

Fawaz E. Alsaadi and Terrance E. Boult

Abstract The number of biometrically-enhanced authentication applications has outstripped our limited number of biometric features and controversy surrounds the actual availability of biometric information, fueling skepticism regarding their effective use for authentication. These concerns emphasize the imperative of addressing the singular nature of biometrics, motivating the need for durable biometric credentials that allow perpetual use of biometric information for authentication, even in the face of compromise. Toward this end, this paper introduces the Durable True-Neighbor Template (DTNT), a novel enhancement of our existing True-Neighbor Template (TNT) approach to fingerprint-based authentication that overcomes the singular nature of fingerprint biometrics through diversification using standard encryption. The results of a standard benchmark experiment, conducted using the FVC protocol with the FVC2006 and FVC2002 databases, are used to evaluate the effectiveness of DTNT for authentication. DTNT shows respectable authentication performance compared with TNT with a generally-negligible loss in fingerprint distinguishability.

Keywords Authentication · Biometrics · Durable · Encryption · Fingerprints · Security · Singular

1 Introduction

The growing interest in enhancing the security of authentication systems using credentials based on biometric features has faced certain challenges. Biometric features are singular, meaning that each of us has only one of any given biometric feature. Once the number of authentication applications outstrips the number of

F.E. Alsaadi(✉) · T.E. Boult
Department of Computer Science, University of Colorado at Colorado Springs, Colorado Springs, CO, USA
e-mail: {falsaadi,tboult}@uccs.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_16

biometric features we possess, an unavoidable overlap in usage will occur based on the pigeonhole principle. Without overcoming the singular nature of biometric features, a biometrically-enhanced authentication system provides an opportunity and incentive for the permanent compromise of the underlying biometric feature information for authentication purposes. This biometric dilemma is not the only challenge that biometrically-enhanced authentication systems face. The security of an authentication system depends upon maintaining the confidentiality of its credentials, which, for biometrically-enhanced authentication systems, implies limiting the availability of biometric feature information. The actual availability of biometric feature information is controversial with recent events bringing into question the assumption of confidentiality.

In 2008, the German hacker group Chaos Computer Club published what it said is the fingerprint of German Interior Minister Wolfgang Schauble, which was lifted from a water glass the minister was using while participating in a discussion [1]. In 2014, a hacker was able to reconstruct the fingerprint of German Defense Minister Ursula von der Leyen using commercial fingerprint software VeriFinger [2] and several close-range photos, including one the hacker took himself from three meters away and another from a press release issued by the minister's own office [3]. Events like these have rightfully fueled skepticism regarding the confidentiality of biometrics and engendered distrust in their use for authentication.

To address these challenges, biometric credentials need to not only provide strong security, but also be durable by being diversifiable and reissuable, which respectively enable use in multiple authentication systems and recovery from compromise. In this regard, we discuss several existing approaches, propose the Durable True-Neighbor Template (DTNT), an enhancement to our existing True-Neighbor-Template (TNT) [4] approach to fingerprint-based authentication, and provide a benchmark evaluation to demonstrate its effectiveness.

2 Existing Approaches

Many existing approaches focus on attempting to inseparably combine singular biometric feature information with a plural key to assure the confidentiality of underlying biometric feature information. An assurance of confidentiality does not necessarily assure durability for authentication purposes. On the other hand, an approach can be durable and yet not assure confidentiality. The ramifications of this must be carefully considered based on the risks and challenges of the scenario. Several notable existing approaches are discussed herein.

2.1 Fuzzy Commitment

Juels and Wattenberg [5] introduced the Fuzzy Commitment protocol to combine biometric feature information and a random correctable codeword as a key into what is known as the helper data. The combined information can later be separated

by applying a new sample of the biometric information and correcting the result as a corrupted codeword. An important restriction is that the underlying biometric templates must be structured such that they are fixed-length and order-invariant. The authors stress that the helper data should not reveal any significant information about the underlying biometric data. For authentication purposes, a hash of the codeword key is stored on the server alongside the helper data. During authentication, the helper data is transmitted to the client who then responds with a hash of the reconstructed codeword key, which is then compared by the server to the stored hash to render an up-or-down authentication verdict.

The Fuzzy Commitment approach has numerous problems. The formatting requirements for the underlying biometric template are extremely restrictive and ill-suited for strong biometric representations [6]. The helper data is freely transmitted to a potentially untrustworthy client. There are a very finite number of codewords from which to select a key, which weakens the security of the system against brute-force attacks. It should be noted that the Fuzzy Commitment approach is not suitable for authentication purposes if absolute confidentiality of the underlying biometric template is a priority. This is due to the need for storing a hash of the key alongside the helper data on the server. With an extremely limited number of codewords from which to select the key, it is possible for an attacker who gains access to the server to perform a relatively-efficient brute-force attack to discover the underlying codeword key and therefore compromise the confidentiality of the underlying biometric template. Bringer, Chabanne, and Kindarji [7] noted a decodability attack wherein two pieces of helper data produced from the same biometric feature information could be applied together to produce a third, damaged codeword. If successful, the assumption is that the two pieces of helper data originated from the same underlying biometric template. To address this attack, Kelkboom, Breebaart, Kevenaar, Buhan, and Veldhuis [8] proposed passing the feature vectors through a record-specific, public permutation process. However, this countermeasure enabled another attack analyzed by Simoens, Tuyls, and Preneel [9] and demonstrated by Tams [10] that can allow an adversary to fully break two related records.

Kaizhi and Aiqun [11] proposed a practical implementation of Fuzzy Commitment to counter these weaknesses. They utilized the FingerCode [12] fingerprint representation to fulfill the order-invariance and fixed-length requirements and BioHashing [13] to non-invertibly transform the representation to provide further resistance to decodability, replay, and some other types of attacks and prevent latent leakage of the underlying biometric template. The authors indicated that their results were affected by information loss from BioHashing as identified in [14,15]. In order to produce a consistent FingerCode, the authors also require that the fingerprint minutiae and core point be stored in the database, which inherently defeats the confidentiality the method attempts to provide.

It should be noted that Fuzzy Commitment is not suitable for authentication purposes because it is not durable. The compromise of underlying biometric feature information in any way results in Fuzzy Commitment being unusable in the future with the compromised biometric feature.

2.2 Fuzzy Vault

Juels and Sudan [6] introduced the Fuzzy Vault protocol to combine biometric feature information and a fixed-length list of random numbers as a key into what is known as the vault. The Fuzzy Vault protocol was created as an attempt to overcome some of the limitations in Fuzzy Commitment. The Fuzzy Vault approach requires that biometric feature information be divided into a set of values that can be evaluated by a polynomial function whose coefficients are the elements of the key. The biometric feature values and their corresponding function values are then recorded as the vault along with chaff biometric feature values and chaff function values based on a chaff polynomial function. The key can later be reconstructed by using a newly-sampled biometric feature to identify enough real biometric feature values and their corresponding function values to solve for the coefficients of the polynomial function. The chaff feature values must be specifically tailored to not correspond to the genuine feature values. Likewise is also true of the chaff function values when constructing the chaff polynomial function. For authentication purposes, a hash of the key is stored on the server alongside the vault. During authentication, the vault is transmitted to the client who then responds with a hash of the reconstructed key, which is then compared by the server to the stored hash to render an up-or-down authentication verdict.

The Fuzzy Vault approach has numerous problems. The formatting requirements for the underlying biometric template can prove very limiting for constructing strong biometric representations. The vault is freely transmitted to a potentially untrustworthy client. It should be noted that the Fuzzy Vault approach is not suitable for authentication purposes if absolute confidentiality of the underlying biometric template is a priority. This is due to the need for storing a hash of the key alongside the vault on the server. It is possible for an attacker who gains access to the server to perform a brute-force attack to discover the underlying key and therefore compromise the confidentiality of the underlying biometric template. Scheirer and Boult [16] noted numerous attacks involving record multiplicity, stolen key inversion, and blended substitution. Nagar, Nandakumar, and Jain [17,18] proposed a hybrid implementation of the Fuzzy Vault and Fuzzy Commitment approaches to counter some of the weaknesses of each and combine their strengths. The formatting requirements for the underlying biometric template are still limiting, but an improvement in security is made to resist brute-force attacks on the vault.

Several practical implementations of Fuzzy Vault have been proposed with varying degrees of security [19,20].

It should be noted that Fuzzy Vault is not suitable for authentication purposes because it is not durable. The compromise of underlying biometric feature information in any way results in Fuzzy Vault being unusable in the future with the compromised biometric feature.

2.3 Two-Factor Protected Minutiae Cylinder-Code (2PMCC)

In response to a claim of confidentiality weakness [21] in their Minutiae Cylinder-Code (MCC) representation [22], Ferrara, Maltoni, and Cappelli [23] proposed a protocol enhancement in the form of Protected Minutiae Cylinder-Code (PMCC) that performs binary KL-projections [24] upon MCC templates, resulting in shortened templates that distill the distinguishability of the original templates to some degree. The authors performed security and confidentiality analyses on their enhanced protocol and found that shorter representations provided greater levels of confidentiality when compared with their longer counterparts with a corresponding cost in security. In order to address the fact that PMCC is not diversifiable or reissuable, Ferrara, Maltoni, and Cappelli [25] proposed Two-Factor Protected Minutiae Cylinder-Code (2PMCC), which permutes the cylinders in PMCC templates based upon a key s and shortens them based on a constant c.

3 Durable True-Neighbor Template (DTNT)

Herein we propose DTNT, a novel enhancement to our existing True-Neighbor Template (TNT) fingerprint representation and matching approach. The DTNT generation process uses our existing, consistent neighbor-selection process to construct central-minutia structures then diversifies those structures with a set of plural encryption keys. The thorough DTNT matching process, which focuses on definitively eliminating false matches to facilitate effective security, utilizes the same fundamental matching process as TNT with slight modification to utilize encrypted values.

3.1 DTNT Generation Process

The DTNT generation process improves on the TNT generation process by overcoming the singular nature of fingerprints by bounding and encrypting intra-user variants in fingerprint information. To recap the heart of the TNT generation process, for any two minutiae with indices i and j in a fingerprint sample, their extracted locations ((x_i, y_i) and (x_j, y_j)) and directions (θ_i and θ_j) can be used to form a relationship between them that includes the Euclidean distance d_{ij} between them and two invariant angles α_{ij} and β_{ij}, as illustrated in Fig. 1. Minutiae i and j are considered neighbors if j is closer to i or i is closer to j than another minutia k. To mitigate the effects of relative distortion and spurious minutiae on neighbor-selection, a factor $F > 1$ is used to scale the distance between i and j, meaning that i and j are neighbors if:

$$\mathrm{d_{ij}} / F \leq \mathrm{d_{ik}} \text{ or } \mathrm{d_{ji}} / F \leq \mathrm{d_{jk}} \tag{1}$$

A TNT then consists of a neighborhood for each minutia i where each neighborhood contains a neighbor relationship of the form (i, j, d_{ij}, α_{ij}, β_{ij}) for each

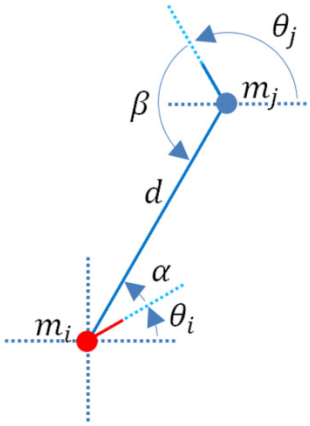


Fig. 1. Relative fingerprint biometrics d, α, and β

qualifying neighbor j of i. Enrollment and authentication TNTs are constructed in the same manner. Although a TNT captures the complexity of minutia relationships within a fingerprint, only a single, differentiable TNT can be constructed for any given finger. Here a piece of information is considered differentiable if versions of that information from two different credentials can be used to distinguish those credentials.

In order to enhance multiple authentication systems using fingerprints from a single finger, it is necessary to be able to produce many differentiable credentials enhanced by fingerprints from that finger, herein known as issuances for that finger. This is a bit of a paradox as the goal of producing fingerprint-enhanced credentials is in producing credentials that are not differentiable. Solving this paradox involves overcoming the singular nature of fingerprints through diversification by specializing fingerprint-enhanced credentials for each authentication application such that each issuance is differentiable.

In order to understand how to diversify a TNT, it is first necessary to understand what information is differentiable within a TNT. In this regard we consider a top-down approach. The number of neighborhoods in a TNT is not differentiable because TNTs originating from fingerprints from different fingers are likely to have similar numbers of neighborhoods. The same principle applies to the number of neighbors within neighborhoods. Thus, there is no need to diversify these quantities. For any given neighbor relationship, i and j are arbitrary index values assigned solely on the basis of the order of the extracted minutiae presented to the generation process. The values of i and j are not differentiable because TNTs originating from fingerprints from different fingers contain approximately the same range of values for i and j. Therefore, there is no need to diversify these values. Moreover, attempts at differentiation based on the values of i and j can be practically thwarted by obfuscating the values of i and j every time a template is generated by randomly permuting the order of the extracted minutiae prior to their presentation to the generation process. The values of d, α, and β in individual

neighbor relationships are not alone or collectively differentiable because they may be quite similar between TNTs originating from fingerprints from different fingers. However, neighborhoods as collections of neighbor relationships are differentiable and therefore need to be diversified.

Because the order of neighborhoods will vary with the randomly-permuted order of extracted minutiae presented to the generation process, all neighborhoods must be diversified in the same manner. Such diversification must be reproducible, meaning that it must be possible to perform the same diversification process at a later point in time with a different fingerprint sample. Moreover, it must be possible to match two related, diversified neighborhoods from different fingerprint samples from the same finger that originated from the same physical minutia. Diversification of neighborhoods can be accomplished by diversifying the values of d, α, and β for all neighbor relationships within each neighborhood, which is the heart of the DTNT enhancement to the TNT representation. Such diversification can be accomplished through encryption with each issuance utilizing a different set of keys. However, if encryption were performed directly on the values of d, α, and β, the resulting ciphertext values would vary unpredictably across fingerprint samples for any given issuance, inhibiting the matching process. In order to use encryption to effectively diversify d, α, and β, the design of the DTNT representation approaches the authentication process in reverse by first looking at what the TNT matching process needs to determine the similarity of two neighbors from two different TNTs.

For a neighbor p from a neighborhood in an authentication TNT and a neighbor g from a neighborhood in an enrollment TNT, their neighbor similarity score (NSS_{pg}) in step 2 of the TNT matching process is based on their spatial-difference Δs_{pg} as shown in (2) and (3):

$$NSS_{pg} = e^{-\Delta s_{pg}^2 / 2\sigma^2} \tag{2}$$

$$\Delta s_{pg}^2 = d_p^2 + d_g^2 - 2\, d_p\, d_g \cos(\alpha_p - \alpha_g) \tag{3}$$

In step 2 of the TNT matching process, if $|\Delta s_{pg}| > T_s$, where T_s is an integer equal to 3σ and σ is a constant, or β_p and β_g differ by more than T_β, where T_β is an integral constant, NSS_{pg} will be assigned 0.

In deriving the DTNT generation process, the polar form of Euclidean distance presented in (3) is refactored into its Cartesian form in (4) with values for x and y being rounded to integers:

$$\Delta s_{pg}^2 = \Delta x^2 + \Delta y^2 \tag{4}$$

$$\Delta x = x_p - x_g \tag{5}$$

$$\Delta y = y_p - y_g \tag{6}$$

$$x_p = \text{round}(d_p \cos(\alpha_p)) \tag{7}$$

$$y_p = \text{round}(d_p \sin(\alpha_p)) \tag{8}$$

$$x_g = \text{round}(d_g \cos(\alpha_g)) \tag{9}$$

$$y_g = \text{round}(d_g \sin(\alpha_g)) \tag{10}$$

Based on (4), values for Δx and Δy that will result in positive values for NSS_{pg} are constrained by (11), resulting in constraints for Δx and Δy in (12) and (13):

$$\Delta s_{pg}^2 = \Delta x^2 + \Delta y^2 \leq T_s^2 \tag{11}$$

$$|\Delta x| \leq T_s \text{ when } \Delta y = 0 \tag{12}$$

$$|\Delta y| \leq T_s \text{ when } \Delta x = 0 \tag{13}$$

Rearranging (5) into (14) and (6) into (15), it can be seen that the values of x_p and y_p are respectively considered in the matching process based on (x_g, Δx) and (y_g, Δy):

$$x_p = x_g + \Delta x \tag{14}$$

$$y_p = y_g + \Delta y \tag{15}$$

Applying (12) to (14) and (13) to (15) results in constraints for x_p and y_p in (16) and (17):

$$x_g - T_s \leq x_p \leq x_g + T_s \tag{16}$$

$$y_g - T_s \leq y_p \leq y_g + T_s \tag{17}$$

From (16) and (17) all possible values for x_p and y_p that will result in a positive NSS_{pg} can be determined based solely on x_g, y_g, and T_s. Because the values for x_p and y_p that can appear in an authentication template will be integers per (7) and (8), the ranges of possible values for x_p and y_p in (16) and (17) are quite finite. These respective ranges for x_p and y_p, herein termed X_g and Y_g, when arranged according to $|\Delta|$ values, are shown in (19) and (21) with their corresponding $|\Delta|$ values shown in (18) and (20):

$$|\Delta x| = [\ 0 \quad 1 \quad 1 \quad 2 \quad 2 \quad \ldots \quad T_s \quad T_s\ ] \tag{18}$$

$$X_g = [\ x_g{+}0 \quad x_g{+}1 \quad x_g{-}1 \quad x_g{+}2 \quad x_g{-}2 \quad \ldots \quad x_g{+}T_s \quad x_g{-}T_s\ ] \tag{19}$$

$$|\Delta y| = [\ 0 \quad 1 \quad 1 \quad 2 \quad 2 \quad \ldots \quad T_s \quad T_s\ ] \tag{20}$$

$$Y_g = [\ y_g{+}0 \quad y_g{+}1 \quad y_g{-}1 \quad y_g{+}2 \quad y_g{-}2 \quad \ldots \quad y_g{+}T_s \quad y_g{-}T_s\ ] \tag{21}$$

The ranges shown in (19) and (21) can be determined for each neighbor relationship when generating a DTNT enrollment template to know all of the possible x_p and y_p values that might be found in a neighbor relationship in a matching DTNT authentication template. The ranges of corresponding $|\Delta|$ values shown in (18) and (20), which are identical, only require knowledge of T_s and therefore can be generated during the matching process to respectively correlate the x_p and y_p

values found in an authentication DTNT to the $|\Delta x|$ and $|\Delta y|$ values needed to determine Δs_{pg}^2 in (4), which in turn allows determination of NSS_{pg} in (2).

Because every pair of columns past the first columns in X_g and Y_g share the same $|\Delta|$ values, these pairs of columns can be randomly permuted every time an authentication DTNT is generated to obfuscate their order without affecting the matching process, resulting in the example ranges shown in (22) and (23):

$$X_g = [\, x_g{+}0 \quad x_g{+}1 \quad x_g{-}1 \quad x_g{-}2 \quad x_g{+}2 \quad \ldots \quad x_g{-}T_s \quad x_g{+}T_s \,] \tag{22}$$

$$Y_g = [\, y_g{+}0 \quad y_g{-}1 \quad y_g{+}1 \quad y_g{+}2 \quad y_g{-}2 \quad \ldots \quad y_g{-}T_s \quad y_g{+}T_s \,] \tag{23}$$

The same logical process can be used to diversify β for pg, resulting in the example range shown in (25) with corresponding $|\Delta|$ values shown in (24):

$$|\Delta\beta| = [\, 0 \quad 1 \quad 1 \quad 2 \quad 2 \quad \ldots \quad T_\beta \quad T_\beta \,] \tag{24}$$

$$B_g = [\, \beta_g{+}0 \quad \beta_g{-}1 \quad \beta_g{+}1 \quad \beta_g{+}2 \quad \beta_g{-}2 \quad \ldots \quad \beta_g{+}T_\beta \quad \beta_g{-}T_\beta \,] \tag{25}$$

Because the matching process need only determine the positions of x_p, y_p, β_p values found in authentication DTNTs in ranges of values like (22), (23), and (25) found in enrollment DTNTs in order to determine the necessary $|\Delta x|$ and $|\Delta y|$ values, the X_g, Y_g, B_g values for enrollment DTNTs and x_p, y_p, β_p values for authentication DTNTs can be encrypted using almost any modern encryption scheme using respective encryption keys k_x, k_y, k_β that correspond to the particular issuance for which templates are being generated.

An enrollment DTNT consists of a neighborhood for each minutia i where each neighborhood contains a neighbor relationship of the form (i, j, E(X_g, k_x), E(Y_g, k_y), E(B_g, k_β)) for each qualifying neighbor j of i. An authentication DTNT consists of a neighborhood for each minutia i where each neighborhood contains a neighbor relationship of the form (i, j, E(x_p, k_x), E(y_p, k_y), E(β_p, k_β)) for each qualifying neighbor j of i.

3.2 DTNT Matching Process

The DTNT matching process required modification of only steps 1 and 2 of the TNT matching process to incorporate comparison of encrypted neighbor information.

The DTNT matching process is as follows:

Given:
An enrollment template with rows containing (i, j, E(X_g, k_x), E(Y_g, k_y), E(B_g, k_β))
An authentication template with rows containing (i, j, E(x_p, k_x), E(y_p, k_y), E(β_p, k_β))
Constants T_β, $T_s = 3\sigma$, T_{cs} defined per Table 1
$|\Delta x|$, $|\Delta y|$, $|\Delta\beta|$ vectors respectively defined per (18), (20), (24)

1. Organize templates into neighborhoods.

For each template, every neighbor row containing the same i is associated with a neighborhood for that template.

2. Determine neighborhood similarities.

For every neighborhood combination ab, where neighborhoods a and b are respectively from the authentication and enrollment templates, a neighborhood similarity score ($NHSS_{ab}$) is determined. For every neighbor combination pg, where neighbors p and g are respectively from a and b, a neighbor similarity score (NSS_{pg}) is determined.

$E(X_g)$ is searched for $E(x_p)$, $E(Y_g)$ is searched for $E(y_p)$, and $E(B_g)$ is searched for $E(\beta_p)$. If any of these searches fail, NSS_{pg} is immediately 0. Otherwise, Δx and Δy respectively inherit the values at the locations in $|\Delta x|$ and $|\Delta y|$ corresponding to where $E(x_p)$ and $E(y_p)$ were respectively found in $E(X_g)$ and $E(Y_g)$.

Unless previously assigned a 0,

NSS_{pg} is determined using (2) and (4).

From this point onward, the DTNT and TNT matching processes are the same.

Once the NSSs of all neighbor combinations for ab have been determined, a greedy algorithm is used to select neighbor pairs.

The greedy algorithm eliminates all neighbor combinations with NSSs of 0 and then follows two steps. First, it selects the neighbor combination with the highest NSS. Second, it eliminates all other neighbor combinations that share a member with the selected combination. These two steps are repeated until there are no remaining neighbor combinations from which to select. The selected neighbor pairs for ab and their NSSs are retained in NP_{ab}.

Unless previously assigned a 0,

$$NHSS_{pg} = \sum \text{NSSs from } NP_{ab} \tag{26}$$

3. Determine neighborhood pairings.

The greedy algorithm described in the previous step is used to select neighborhood pairs based on NHSSs from all neighborhood combinations. The selected neighborhood pairs and their corresponding NPs and NHSSs are retained in NHP.

4. Eliminate neighbor pairs inconsistent with neighborhood pairs.

For each neighborhood pair ab in NHP, a neighbor pair pg in NP_{ab} is eliminated if there does not exist a neighborhood pair pg in NHP. NHSSs are adjusted accordingly based on recalculations of (26). Neighborhood pairs in NHP with NHSSs of 0 are eliminated.

5. *Eliminate non-mutual neighbors.*

   For each neighborhood pair *ab* in NHP, a neighbor pair *pg* in NP_{ab} is eliminated if there does not exist a neighbor pair *ab* in NP_{pg}. for neighborhood pair *pg* in NHP. NHSSs are adjusted accordingly based on recalculations of (26). Neighborhood pairs in NHP with NHSSs of 0 are eliminated.

6. *Identify and filter clusters.*

   Neighboring neighborhood pairs in NHP are associated into clusters by beginning a cluster with an unclustered neighborhood pair in NHP and then repeatedly adding all neighboring neighborhood pairs for all cluster members to the cluster until no new members are added. Neighborhood pairs in NHP that are members of a cluster with fewer members than T_{cs} are eliminated.

7. *Calculate match score.*

   For an enrollment template and an authentication template, respectively containing *m* and *n* neighborhoods,

$$\text{Match Score} = 2\ (\sum \text{NHSSs from NHP}) / (m + n) \tag{27}$$

4 Security Analysis

In order to understand the true value of the DTNT approach as well as its limitations, it is necessary to understand the scenarios of its use and how it may be attacked. The primary benefit of DTNT from the standpoint of compromise is in providing time to people maintaining authentication systems and people victimized by compromise while limiting liability from loss. This time can take several forms, including time to defend and time to recover. It is taken from attackers by prolonging the time needed for their efforts, increasing the risks and challenges they face. Attacks need not be directly targeted at an authentication system to cause damage. An effective authentication system must be secure from both internal and external events.

To compromise an authentication system implementing DTNT an attacker would need to acquire either a correct fingerprint and the encryption keys or the enrollment template from the server. Because the server never transmits the enrollment template for any reason during the authentication process, the enrollment template would need to be stolen either during the enrollment process or directly from the server. This is an instance of breaking through the back door to steal the keys for the front door. Once such a compromise was noticed, which would likely be immediately, every compromised enrollment template would be instantly revoked and reissued later using different encryption keys. The compromised enrollment template in its current form would be worthless towards compromise of this or any other authentication system implementing DTNT. To be of any value to an attacker, the stolen enrollment template would need to be heavily analyzed.

Due to the construction of an enrollment DTNT, E(X_g, k_x), E(Y_g, k_y), and E(B_g, k_β) are respectively ordered according to $|\Delta x|$, $|\Delta y|$, and $|\Delta\beta|$. An attacker could attempt three separate brute-force attacks to obtain (X_g, Y_g, B_g) and (k_x, k_y, k_β). Because none of the plaintext values in X_g, Y_g, or B_g are known, for each range the attacker would need to repeatedly make a guess for both the plaintext value and the encryption key until a match is found for the encrypted value in the first column of the range. This alone does not confirm a compromise because many plaintext-key combinations can result in the same ciphertext value. While the ranges of possible correct values for X_g, Y_g, and B_g are limited, there are sufficiently-many possible correct values to greatly hinder an attacker, especially when compounded by the exceedingly-large number of possible encryption keys with which they can be paired. In order to confirm a compromise of a given range, the attacker would need to add and subtract one from the guessed plaintext value, encrypt the results, and compare them with the ciphertexts found in the second and third columns of the range. Even if they match, the attacker has no way of knowing if the order of the two underlying plaintext values matches the order of the two guessed plaintext values because each pair of columns past the first column in every neighbor row was randomly permuted during the generation process. The attacker could further proceed to confirm each additional pair of columns in similar fashion with similar uncertainty. A lengthy set of matches would provide some confirmation that the guessed plaintext values and encryption key are in fact correct. This process would require a significant amount of time and computing power to compromise the underlying information of even a single enrollment template. Even if such an attack were successful, the compromised ranges and encryption keys would be worthless against a different authentication system implementing DTNT because the attacker would not possess the encryption keys for that system. An attempted brute-force attack of those encryption keys through repeated authentication attempts on that system using the compromised ranges would be halted by the system after a limited number of failed attempts, making such an attack statistically useless. The same situation would result if the attacker obtained the correct fingerprint somehow. An attacker could try to compromise the device containing the encryption keys.

Compromising the device containing the encryption keys would require identifying the particular device or its owner, which could be very difficult, such as in the case of a smart phone. A local compromise accomplished by tracking down the device's owner and gaining access to the device would be particularly risky and challenging. Even if the device were found or stolen, an attacker would need to obtain the correct fingerprint, which would prove difficult in a reasonable span of time unless the attacker knew exactly where and how to acquire it. This gives time to allow remote locking or wiping of the device, rendering the stolen encryption keys practically inaccessible. Additionally, the entities maintaining the authentication systems associated with the stolen encryption keys could be informed to revoke the associated enrollment templates, rendering the stolen encryption keys useless. Applying a simple, password-based encryption scheme to the

encryption keys on the device would thwart an attacker even further as it would take time to attack the encryption on the encryption keys, which would likely already be revoked by the time such an attack were successful.

In the event that an attacker should acquire the correct fingerprint and the encryption keys, all associated enrollment templates could be revoked and reissued using different encryption keys. If an attacker should somehow acquire an authentication template during its brief existence, as it is never stored, the attacker would be unable to analyze the stolen authentication template in any of the previously-described manners and therefore be unable to compromise the underlying ranges and encryption keys. Handling of this situation would require detection of the compromise along with revocation and reissue of the associated enrollment template using different encryption keys. In all of these scenarios, DTNT allows the perpetual use of fingerprints for authentication, even in the face of compromise, which proves that the DTNT approach is durable.

5 Experiment

A standard benchmark experiment was conducted using the FVC2006 [26] and FVC2002 [27] protocols and fingerprint databases. The first database of FVC2006 was omitted from this benchmark due to the low resolution of its fingerprint images. Each FVC2006 database contains 12 samples per finger of 140 fingers. Each FVC2002 database contains 8 samples per finger of 100 fingers. The VeriFinger SDK [2] was used to extract minutia locations and directions. The benchmarked parameters listed in Table 1 were used for all fingerprint databases. The experiment consisted of two segments.

For the first segment of the experiment, per the FVC2006 and FVC2002 protocols, false rejection rates (FRRs) were determined by calculating match scores for all non-repeated pairs of all fingerprints from the same finger, applying a decision threshold below which false rejections were counted, and dividing by the total number of trials (9240 for FVC2006 and 2800 for FVC2002). Similarly, false acceptance rates (FARs) were determined by calculating match scores for all non-repeated pairs of the first fingerprint samples from all fingers, applying a decision threshold at or above which false acceptances were counted, and dividing by the total number of trials (9730 for FVC2006 and 4950 for FVC2002).

For the second segment of the experiment, FRRs were determined by calculating match scores for all non-repeated pairs of the first two fingerprints from the same finger, applying a decision threshold below which false rejections were counted, and dividing by the total number of trials (140 for FVC2006 and 100 for FVC2002). Similarly, FARs were determined by calculating match scores for all non-repeated pairs of the first fingerprint samples from all fingers, applying a decision threshold at or above which false acceptances were counted, and dividing by the total number of trials (9730 for FVC2006 and 4950 for FVC2002).

Both segments of the experiment were designed to evaluate the authentication performance of DTNT against TNT as well as several well-known approaches, including: MCC [22][28], PMCC [23], and 2PMCC [25]. The software for these approaches was obtained directly from the respective authors. For more details

regarding these approaches and their parameters, please see the respective papers. For DTNT and 2PMCC, the same keys were used for all trials to allow for fair comparison with TNT, MCC, and PMCC. For $PMCC_k$ and $2PMCC_{k,c}$, only $PMCC_{128}$ and $2PMCC_{64,64}$ are referenced in the discussion herein because they demonstrated superior authentication performance amongst all parameter configurations for PMCC and 2PMCC.

For the first segment of the experiment, Table 2 shows the FRRs for a FAR of 0. Lower FRR indicates superior authentication performance. DTNT demonstrated comparable authentication performance across all of the fingerprint databases with respect to TNT with the exception of FVC2006 4A, which is composed of artificial fingerprints. DTNT demonstrated superior authentication performance across all of the fingerprint databases with respect to $PMCC_{128}$ and $2PMCC_{64,64}$.

Table 1 Benchmarked Parameters

Method	Generation Parameters	Match Parameters
DTNT	F = 1.49	$T_\beta = 19°$, $\sigma = 20 / 3$, $T_{cs} = 5$
TNT		$T_\beta = 19°$, $\sigma = 20 / 3$, $T_{cs} = 4$, $T_\theta = 90°$
MCC	Optimal Parameters per MCC SDK v2.0	
PMCC	PMCC Paper Parameters per MCC SDK v2.0	
2PMCC		

Table 2 FRRs for 0_{FMR} for First Experiment Segment

Method	FVC2006			FVC2002			
	2A	3A	4A	1A	2A	3A	4A
MCC_{16b}	1.76	12.12	6.90	1.14	1.21	9.64	2.79
$PMCC_{16}$	30.31	73.55	87.52	53.89	52.75	79.11	77.96
$PMCC_{32}$	14.83	39.47	60.12	19.11	14.71	60.29	44.36
$PMCC_{64}$	8.35	27.59	31.24	6.14	5.36	38.82	15.64
$PMCC_{128}$	4.70	17.06	24.46	3.00	3.14	29.36	9.43
$2PMCC_{32,24}$	16.85	54.37	83.46	25.25	19.93	58.71	60.82
$2PMCC_{32,32}$	14.83	39.42	60.38	19.14	14.54	60.21	44.32
$2PMCC_{64,48}$	11.04	29.07	41.56	7.29	7.36	47.68	21.64
$2PMCC_{64,64}$	8.35	27.59	31.18	6.14	5.29	38.79	15.57
TNT	1.33	11.26	4.65	0.96	1.07	6.89	2.79
DTNT	1.28	9.82	8.02	1.14	1.07	7.46	3.32

Table 3 FRRs for 0_{FMR} for Second Experiment Segment

Method	FVC2006			FVC2002			
	2A	3A	4A	1A	2A	3A	4A
MCC_{16b}	3.57	19.29	5.00	0.00	0.00	6.00	3.00
$PMCC_{16}$	37.86	82.86	82.86	23.00	26.00	57.00	71.00
$PMCC_{32}$	20.00	52.14	57.14	3.00	2.00	36.00	49.00
$PMCC_{64}$	12.86	38.57	25.71	0.00	1.00	23.00	14.00
$PMCC_{128}$	7.86	27.14	19.29	0.00	0.00	20.00	12.00
$2PMCC_{32,24}$	19.29	61.43	79.29	6.00	5.00	37.00	56.00
$2PMCC_{32,32}$	20.00	52.14	57.14	3.00	2.00	36.00	49.00
$2PMCC_{64,48}$	16.43	39.29	40.00	0.00	1.00	30.00	23.00
$2PMCC_{64,64}$	12.86	38.57	25.71	0.00	1.00	23.00	14.00
TNT	2.14	19.29	5.00	0.00	0.00	4.00	2.00
DTNT	2.86	17.14	5.71	0.00	0.00	4.00	3.00

For the second segment of the experiment, Table 3 shows the FRRs for a FAR of 0. DTNT demonstrated comparable authentication performance across all of the fingerprint databases with respect to TNT. DTNT demonstrated superior authentication performance for all of the fingerprint databases with respect to $PMCC_{128}$ and $2PMCC_{64,64}$ with the exceptions of FVC2002 1A where the FRRs for all three were 0 and FVC2002 2A where the FRRs for DTNT and $PMCC_{128}$ were 0.

6 Conclusion

The singular nature of biometric features and events that have brought into question the assumption of their confidentiality have challenged the growing interest in enhancing the security of authentication systems using credentials based on biometric features. This paper proposed the Durable True-Neighbor Template (DTNT), an enhancement to our existing True-Neighbor-Template (TNT) approach to fingerprint-based authentication that diversifies fingerprint biometrics through encryption to address these challenges in order to produce biometric credentials that not only provide strong security, but are also durable by being diversifiable and reissuable, enabling use in multiple authentication systems and recovery from compromise. The DTNT approach demonstrated comparable authentication performance with respect to TNT and generally-superior authentication performance with respect to several well-known approaches, including: PMCC and 2PMCC. These findings were based on the results of a standard benchmark experiment utilizing the FVC2006 and FVC2002 protocols and fingerprint databases.

References

1. Hackers Publish German Minister's Fingerprint | WIRED. http://www.wired.com/2008/03/hackers-publish
2. VeriFinger SDK – Neurotechnology. http://www.neurotechnology.com/verifinger.html
3. Hacker fakes German minister's fingerprints using photos of her hands | Technology | The Guardian. http://www.theguardian.com/technology/2014/dec/30/hacker-fakes-german-ministers-fingerprints-using-photos-of-her-hands
4. Alsaadi, F.E., Boult, T.E.: Furthering Fingerprint-Based Authentication: Introducing the True-Neighbor Template. Unpublished
5. Juels, A., Wattenberg, M.: A fuzzy commitment scheme. In: Proceedings of the 6th ACM Conference on Computer and Communications Security, pp. 28–36. ACM (1999)
6. Juels, A., Sudan, M.: A fuzzy vault scheme. Designs, Codes and Cryptography **38**, 237–257 (2006)
7. Bringer, J., Chabanne, H., Kindarji, B.: The best of both worlds: applying secure sketches to cancelable biometrics. Science of Computer Programming **74**, 43–51 (2008)
8. Kelkboom, E.J.C., Breebaart, J., Kevenaar, T.A.M., Buhan, I., Veldhuis, R.N.J.: Preventing the decodability attack based cross-matching in a fuzzy commitment scheme. IEEE Transactions on Information Forensics and Security **6**, 107–121
9. Simoens, K., Tuyls, P., Preneel, B.: Privacy weaknesses in biometric sketches. In: 30th IEEE Symposium on Security and Privacy, 2009, pp. 188–203. IEEE (2009)

10. Tams, B.: Decodability attack against the fuzzy commitment scheme with public feature transforms. arXiv preprint arXiv:1406.1154
11. Kaizhi, C., Aiqun, H.: An enhancing fingerprint template protection method. In: 5th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 275–279. IEEE (2013)
12. Jain, A.K., Prabhakar, S., Hong, L., Pankanti, S.: FingerCode: a filterbank for fingerprint representation and matching. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, vol. 2. IEEE (1999)
13. Jin, A.T.B., Ling, D.N.C., Goh, A.: Biohashing: two factor authentication featuring fingerprint data and tokenised random number. Pattern Recognition **37**, 2245–2255 (2004)
14. Kong, A., Cheung, K.-H., Zhang, D., Kamel, M., You, J.: An analysis of BioHashing and its variants. Pattern Recognition **39**, 1359–1368 (2006)
15. Nanni, L., Lumini, A.: Empirical tests on biohashing. Neurocomputing **69**, 2390–2395 (2006)
16. Scheirer, W.J., Boult, T.E.: Cracking fuzzy vaults and biometric encryption. In: Biometrics Symposium, 2007, pp. 1–6. IEEE (2007)
17. Nagar, A., Nandakumar, K., Jain, A.K.: Securing fingerprint template: fuzzy vault with minutiae descriptors. In: 19th International Conference on Pattern Recognition, ICPR 2008, pp. 1–4. IEEE (2008)
18. Nagar, A., Nandakumar, K., Jain, A.K.: A hybrid biometric cryptosystem for securing fingerprint minutiae templates. Pattern Recognition Letters **31**, 733–741
19. Nandakumar, K., Jain, A.K., Pankanti, S.: Fingerprint-based fuzzy vault: implementation and performance. IEEE Transactions on Information Forensics and Security **2**, 744–757 (2007)
20. Hartloff, J., Dobler, J., Tulyakov, S., Rudra, A., Govindaraju, V.: Towards fingerprints as strings: secure indexing for fingerprint matching. In: 2013 International Conference on Biometrics (ICB), pp. 1–6. IEEE (2013)
21. Nagar, A.: Biometric template security. Michigan State University
22. Cappelli, R., Ferrara, M., Maltoni, D.: Minutia cylinder-code: a new representation and matching technique for fingerprint recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**, 2128–2141
23. Ferrara, M., Maltoni, D., Cappelli, R.: Noninvertible minutia cylinder-code representation. IEEE Transactions on Information Forensics and Security **7**, 1727–1737
24. Fukunaga, K.: Statistical pattern recognition. Academic Press (1990)
25. Ferrara, M., Maltoni, D., Cappelli, R.: A two-factor protection scheme for MCC fingerprint templates. In: 2014 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–8. IEEE (2014)
26. Cappelli, R., Ferrara, M., Franco, A., Maltoni, D.: Fingerprint verification competition 2006. Biometric Technology Today **15**, 7–9 (2007)
27. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2002: second fingerprint verification competition. In: Proceedings of 16th International Conference on Pattern Recognition, 2002, vol. 3, pp. 811–814. IEEE (2002)
28. Cappelli, R., Ferrara, M., Maltoni, D.: Fingerprint indexing based on minutia cylinder-code. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**, 1051–1057

Certificate-Based IP Multimedia Subsystem Authentication and Key Agreement

Wei-Kuo Chiang and Ping-Chun Lin

Abstract While more and more mobile devices making requests to get access to the IP multimedia subsystem (IMS), authenticating these devices leads to heavy load on the IMS entities. So, how to shorten the IMS authentication procedure is an important issue. This paper proposes a certificate-based IMS authentication and key agreement scheme, abbreviated as C-IMS AKA. We modify the E-UTRAN attach procedure to obtain the certificate of a UE, and enable the home subscriber server (HSS) and serving call session control function (S-CSCF) to authenticate the UE by verifying the certificate as certification authority (CA) in the registration procedure. Therefore, the UE can be authenticated without the challenge-response procedure; moreover, the number of message exchanges will be reduced. In addition, we analyze and compare the C-IMS AKA with other existing schemes in terms of message propagation delay and queuing delay. Analytical results show that the proposed scheme could make the IMS registration more efficient than current alternatives.

Keywords IP multimedia subsystem · Authentication · Certificate · Delivery cost · Queuing delay

1 Introduction

While more and more users want to access various services in the mobile network, 3GPP defines the IMS [1] that can provide various multimedia services like voice session and video delivery. The IMS is responsible for helping each user equipment (UE) establish multimedia session by session initiation protocol (SIP) [2].

W.-K. Chiang(✉) · P.-C. Lin
Department of Computer Science and Information Engineering,
National Chung Cheng University, Min-Hsiung, Chia-Yi 621, Taiwan, ROC
e-mail: wkchiang@cs.ccu.edu.tw, magicshow159@hotmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_17

Before accessing the IMS network, the UE has to be authenticated by the IMS authentication and key agreement (IMS AKA) procedure. However, the procedure is a two-pass and challenge-response based procedure. With more and more UEs requesting access to the IMS network, the procedure causes the IMS entities heavy load. The procedure may also increase the computation time cost and energy consumption of the UE. Therefore, how to shorten the authentication procedure is an important issue.

In order to solve this problem, Lin *et al.* [3] indicated that several steps and parameters of the GPRS authentication can be reused in the IMS AKA because of their similarity. Therefore, the HSS can identify a UE only by checking the IP Multimedia Privacy Identity (IMPI) of the UE in one round-trip message exchanges after the UE have performed its GPRS authentication.

Huang *et al.* proposed Evolutionary-IMS (E-IMS) AKA [4]. In the E-IMS AKA, the UE can obtain a secret key shared with the HSS after the UE attaching the GPRS network. Then, the UE can encrypt its security parameters generated for the registration procedure with the key; so that, the mutual authentication between the UE and the S-CSCF can be complete within the one-pass message exchanges. That is, the E-IMS AKA can enhance the security in the authentication procedure by providing the mutual authentication procedure between the UE and the IMS network.

Long *et al.* proposed another scheme [5] to enhance Lin *et al*. one-pass authentication, abbreviated as "E-OPA" by us. The S-CSCF can validate a UE by using the pair parameters IMPI and International Mobile Subscriber Identity (IMSI) of the UE. The E-OPA can prevent the registration authentication from replay attack.

The certificate of a UE is implemented on a public-key infrastructure, and is issued by an authoritative certificate authority (CA). The UE can use the certificate as its identification; thus, the UE can be authenticated without other authentication procedures. Therefore, we integrate the IMS AKA with the certificate to improve the registration procedure.

In this paper, we propose a certificate-based IMS AKA scheme, abbreviated as C-IMS AKA. We modify the E-UTRAN attach procedure [6] so that a UE obtains a certificate issued by the HSS after the UE attaching the LTE network. After issuing the certificate for the UE, the HSS plays the role of certificate authority (CA) in the C-IMS AKA. The S-CSCF also is enhanced to verify the UE certificate. So, the UE can be authenticated without the challenge-response procedure; therefore, the number of exchanged messages between the HSS and the S-CSCF will be reduced.

The remaining part of this paper is organized as follows. Section 2 introduces the IMS entities and related works. Section 3 presents the proposed scheme. In section 4, we analyze the performance of the proposed scheme. Finally, we give a conclusion of this paper in section 5.

2 Related Works

2.1 The Evolutionary-IMS AKA Procedure (E-IMS AKA)

The E-IMS AKA [4] replaces two pass authentication in the standard IMS authentication procedure with one pass authentication, and also enhances the security of one-pass authentication. The E-IMS AKA can reduce the delivery cost as compared with the standard IMS AKA. Figure 1 depicts the E-IMS AKA procedure.

2.2 The Enhanced One-Pass IMS Authentication (E-OPA)

The E-OPA [5] can reduce the redundancies in the standard IMS authentication procedure. Moreover, the E-OPA has lower computation time cost of the UE than the E-IMS AKA. The E-OPA is more secure than one-pass authentication proposed by Lin *et al*. Figure 2 depicts the E-OPA procedure.

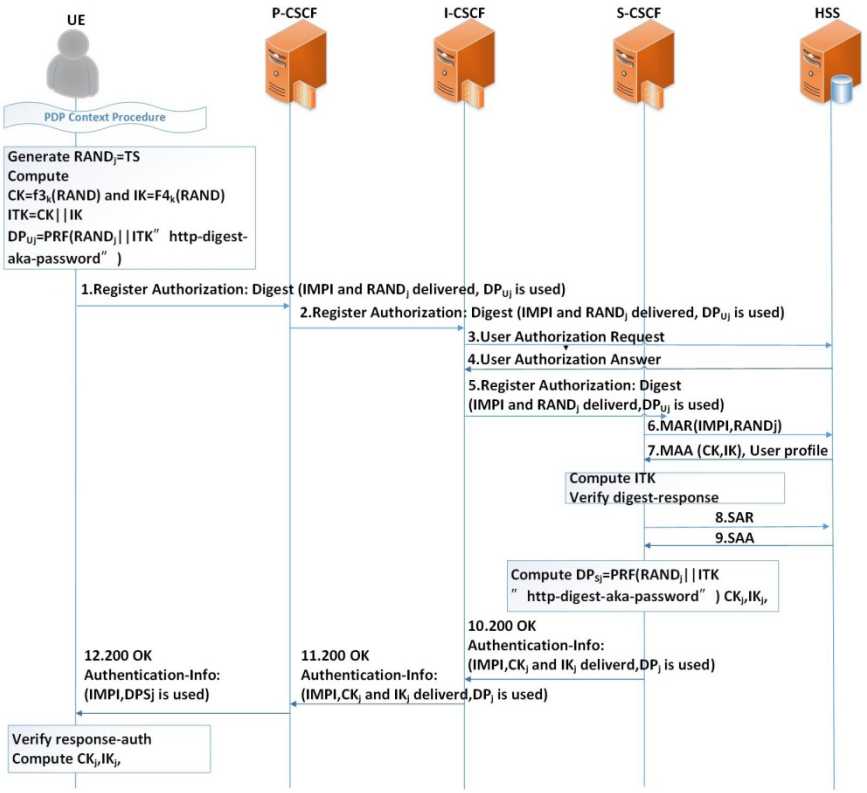


Fig. 1 The E-IMS AKA procedure.

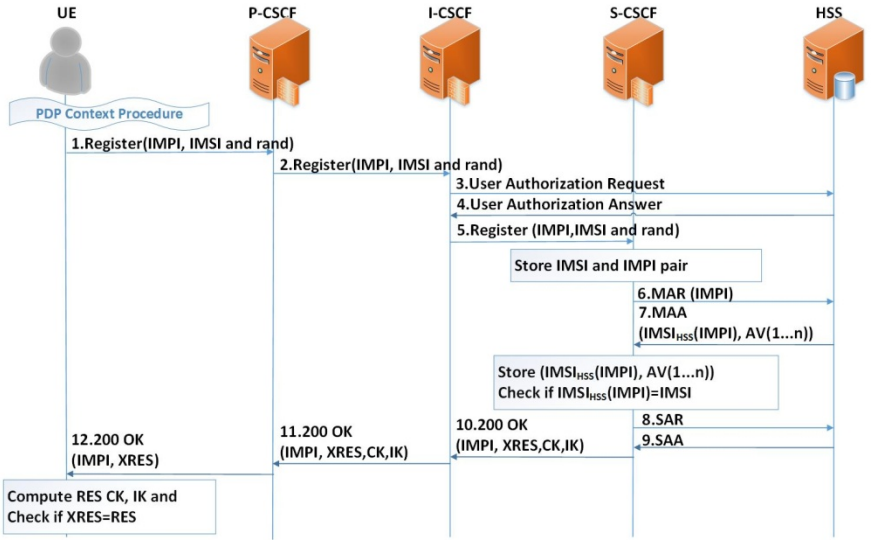


Fig. 2 The E-OPA procedure.

3 The Certificate-Based IMS AKA

This section introduces the certificate-based IMS AKA. In order to improve the authentication procedure, we integrate the certificate into the E-UTRAN and the IMS network. We first present the modification of LTE attach procedure for UE attaching the network. Then, we illustrate the C-IMS AKA for initial registration and re-registration procedures, respectively.

3.1 Modification of LTE Attach Procedure

We modify the E-UTRAN initial attach procedure in [7]. Figure 3 shows the message flow of the modified attach procedure. Steps 1 to 21 and 25 to 29 are the same as those defined in [7]. The modified flow (steps 22~24) is shown in red words, and the corresponding steps are illustrated below.

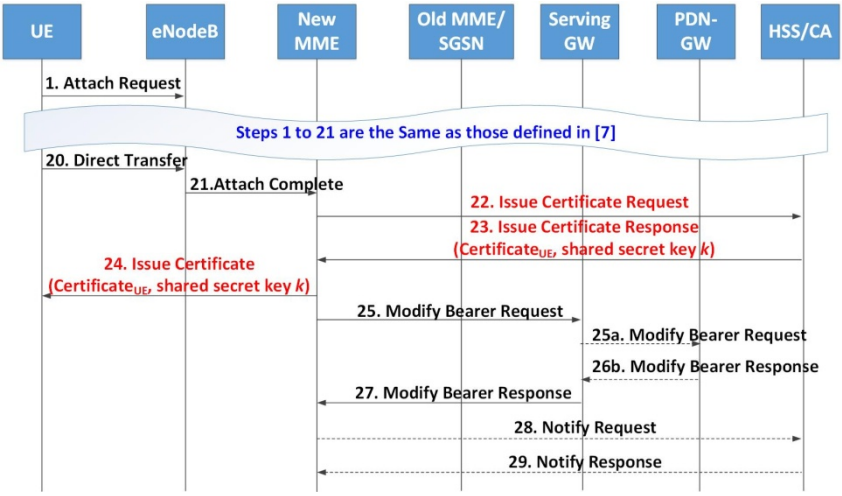


Fig. 3 Modification of the initial attach procedure.

Steps 22~24: The new MME sends the Issue Certificate Request message to the HSS. Then, the HSS will reply to the MME with the certificate for the UE, and the shared secret key *K* carried in the Issue Certificate Response message. The MME will forward the certificate to the UE.

3.2 The C-IMS AKA for the Initial Registration Procedure

In the initial registration procedure, a UE performs the IMS authentication by integrating the certificate obtained from the UE attaching procedure. The UE encrypts the Register message with the secret key K shared with the HSS to protect the certificate confidentiality. We enable the HSS not only to store the UE subscription information in the E-UTRAN environment but also to verify the UE certificate in the IMS registration procedure. That is, the HSS is responsible for issuing and verifying the certificate for the UE registration. After authenticating the UE certificate, the HSS will directly store the name of the S-CSCF without Sever Assignment Request and Sever Assignment Response messages. Thus, the number of exchanged messages can be reduced. Figure 4 depicts the C-IMS AKA initial registration procedure, and the corresponding steps are illustrated below.

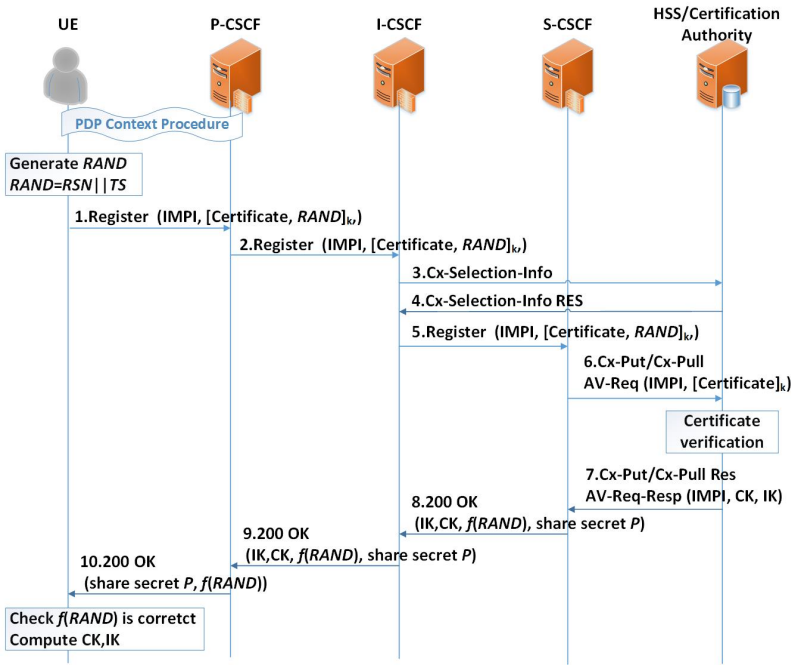


Fig. 4 The C-IMS AKA initial registration procedure.

Steps 1~2: A UE sends a SIP Register message with IMPI, $\text{Certificate}_{\text{UE}}$, and random number *RAND*, encrypted with shared secret key *k*, to the P-CSCF. We adopt the Long's proposed method [5] parameters that the *RAND* is derived from random sequence number *RSN* and timestamp *TS*. Then, the P-CSCF forwards Register message to the I-CSCF. The certificate is obtained from the E-UTRAN initial attach procedure, and carried in the SIP message Authorization header field [2].

Steps 3~5: The I-CSCF sends a User-Authorization Request message to the HSS. The HSS selects the S-CSCF for serving the UE, and sends a User-Authorization-Answer with the S-CSCF address to the I-CSCF. Then, the I-CSCF forwards the Register message to the S-CSCF.

Steps 6~7: Upon receiving the Register message, the S-CSCF will check whether *TS* is in the acceptable time window. If so, the S-CSCF sends the Register message to the HSS. The HSS will verify the certificate. If the certificate is valid, the HSS will store the name of the S-CSCF, and send AV response message including *IMPI, CK,* and *IK* to the S-CSCF. Then the S-CSCF computes *f*(*RAND*) where *f* is a function that performs some transformation on *RAND* (e.g., adding one).

Steps 8~10: The S-CSCF sends 200 OK message with *CK, IK, f*(*RAND*), and secret key *P* to the UE. Upon receiving the 200 OK message, the UE will derive and verify the parameter *f*(*RAND*); so that, the UE can determine whether this message is fresh. If so, the UE authenticates the S-CSCF successfully and finishes the initial registration. After that, the subsequently exchanged messages can be protected by *CK* and *IK*.

3.3 The C-IMS AKA for the Re-registration Procedure

After a UE completing the initial registration, the selected S-CSCF for serving the UE is unchangeable even when the UE performs the re-registration in case of handover. Thus, we enable the S-CSCF to verify the UE certificate in order to decrease the number of messages exchanged between the S-CSCF and the HSS. If any UE revokes its certification, the HSS/Certification Authority will update the certificate revocation list (CRL) to the corresponding S-CSCF to ensure the accuracy of certificate validation. Figure 6 shows the message flow of the C-IMS AKA re-registration procedure, and the corresponding steps are illustrated below.

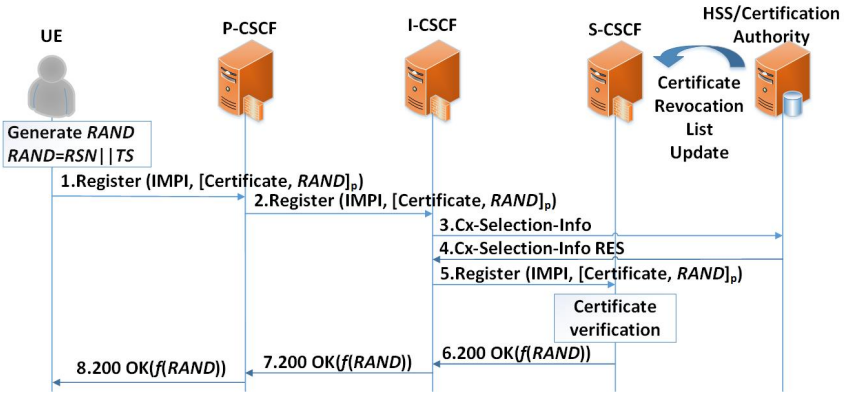


Fig. 5 The C-IMS AKA re-registration procedure.

Steps 1~5: When the UE needs to re-register in case of handover or registration expire. The UE sends a Register message with IMPI, certificate, and *RAND*, where the latter two parameters are encrypted with shared secret key *P*. Then the message will pass through the P-CSCF and the I-CSCF arrive at the S-CSCF.

Steps 6~8: Upon receiving the Register message, the S-CSCF will check the parameter *TS* to determine whether the message is acceptable. Then the S-CSCF will verify the certificate carried in the Register message. If the result is positive, the S-CSCF will send a 200 OK message with the parameter *f*(*RAND*) to the UE via the I-CSCF and the P-CSCF. The UE will check whether *f*(*RAND*) is correct. Then, the UE authenticate the S-CSCF successfully.

3.4 Security Analysis

The following offers security analysis of the C-IMS AKA to demonstrate the feasibility.

Mutual Authentication between the UE and the IMS: With the C-IMS AKA, the IMS network can authenticate the UE by verifying the certificate carried in the Register message, and the UE can also authenticate the S-CSCF by checking

whether the parameter *f*(*RAND*) carried in the 200 OK message is correct. Thus, the C-IMS AKA provides mutual authentication between the UE and the IMS.

Message Confidentiality: To protect the message with the certificate from being intercepted and misappropriated, the UE encrypts the Register message with the shared secret key obtained from the HSS in the attach network procedure. In addition, the UE also encrypts the Register message when it performs the re-registration. That is, the C-IMS AKA can protect message confidentiality.

Replay Attacks: Replay attack is very common, and it may cause the IMS entities overloading. To defeat replay attacks, each message in the C-IMS AKA contains the timestamp *TS*. When the S-CSCF receives the Register message, it checks whether *TS* is in the acceptable time window. The S-CSCF can determine whether the message is fresh; thus, the C-IMS AKA can prevent replay attack.

4 Performance Analysis

We compare the C-IMS AKA with the existing schemes in terms of message propagation delay [8] and queuing delay [9].

4.1 Message Propagation Delay

According to Figures 1, 2, and 4, the message propagation delay of the initial registration procedure in the standard IMS AKA, the E-IMS AKA, the E-OPA, and the C-IMS AKA, can be derived as follows:

$$T_{IMS\text{-}reg} = 4T_{UE-P-CSCF} + 4T_{P-CSCF-I-CSCF} + 2T_{I-CSCF-HSS} + 4T_{I-CSCF-S-CSCF} + 4T_{S-CSCF-HSS} \quad (1)$$

$$T_{E\text{-}IMS\text{-}reg} = 2T_{UE-P-CSCF} + 2T_{P-CSCF-I-CSCF} + 2T_{I-CSCF-HSS} + 4T_{I-CSCF-S-CSCF} + 4T_{S-CSCF-HSS} \quad (2)$$

$$T_{E\text{-}OPA\text{-}reg} = 2T_{UE-P-CSCF} + 2T_{P-CSCF-I-CSCF} + 2T_{I-CSCF-HSS} + 4T_{I-CSCF-S-CSCF} + 4T_{S-CSCF-HSS} \quad (3)$$

$$T_{C\text{-}IMS\text{-}reg} = 2T_{UE-P-CSCF} + 2T_{P-CSCF-I-CSCF} + 2T_{I-CSCF-HSS} + 2T_{I-CSCF-S-CSCF} + 2T_{S-CSCF-HSS} \quad (4)$$

We adopt the parameters of the message propagation used in [8] that assume the propagation delay between two nodes in distinct network is α unit, and that between two nodes in the same network is β unit. We define that parameters α and β are ratio relation for this analysis, but not specific value. Global System for Mobile Communications (GSM) world illustrates that almost 20% roaming subscriber and 80% non-roaming subscriber [10]. With the above set of data,

the improvement ratio of message propagation delay for the E-IMS AKA, the E-OPA, and the C-IMS AKA over the standard IMS AKA can derived respectively:

$$\frac{T_{IMS}-T_{E\text{-}IMS}}{T_{IMS}}=\frac{5.6\beta+0.4a}{17.2\beta+0.8a}=\frac{1.4\beta+0.1a}{4.3\beta+0.2a} \tag{5}$$

$$\frac{T_{IMS}-T_{E\text{-}OPA}}{T_{IMS}}=\frac{1.4\beta+0.1a}{4.3\beta+0.2a} \tag{6}$$

$$\frac{T_{IMS}-T_{C\text{-}IMS}}{T_{IMS}}=\frac{7.6\beta+0.4a}{17.2\beta+0.8a}=\frac{1.9\beta+0.1a}{4.3\beta+0.2a} \tag{7}$$

Equations (5), (6) and (7) are plotted as the function of the value α/β in Figure 6. The message propagation delay for the E-IMS AKA and the E-OPA are the same. The improvement ratio is increased with the value of α/β growing. Figure 6 shows that the C-IMS AKA improvement ratio is the highest.

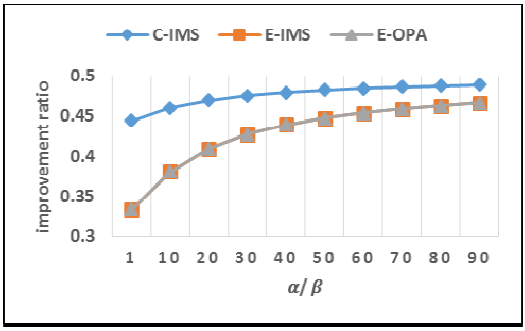


Fig. 6 The improvement ratio of message propagation delay.

4.2 Queuing Delay

We analyze the queuing delay of handling message exchanges at each node, i.e. UE, P/I/S-CSCF, and HSS. We model these nodes with M/M/1 queuing model [9]. Table 1 shows the notations and description used in the queuing delay model.

Table 1 Notations for the Queuing Delay Model.

Notation	*Description*
C_N	The number of UE (assume 10000)
μ	Service rate
λ	Arrival rate of requests for the registration
λ^{other}	Arrival rate of other requests
Q_N	Queuing delay time of node N
n	The number of messages handled at each node

We assume the arrival rate of other requests λ^{other} for the HSS is 0.7μ, and that for the UE/CSCF is 0.6μ, since the load of the HSS is usually higher than that of other nodes. The queuing delay of messages handled at the UE is:

$$\frac{n_{UE}}{\mu_{UE} - C_N \times n_{UE} \times \lambda - \lambda^{other}} = \frac{n_{UE}}{\mu_{UE} - 1 \times n_{UE} \times \lambda - 0.6 \times \mu} \tag{8}$$

The queuing delay of messages handled at the P/I/S-CSCF is:

$$\frac{n_{CSCF}}{\mu_{CSCF} - C_N \times n_{CSCF} \times \lambda - \lambda^{other}} = \frac{n_{CSCF}}{\mu_{CSCF} - C_N \times n_{CSCF} \times \lambda - 0.6 \times \mu} \tag{9}$$

The queuing delay of messages handled at the HSS is:

$$\frac{n_{HSS}}{\mu_{HSS} - C_N \times n_{HSS} \times \lambda - \lambda^{other}} = \frac{n_{HSS}}{\mu_{HSS} - C_N \times n_{HSS} \times \lambda - 0.7 \times \mu} \tag{10}$$

We only evaluate the queuing delay for the initial registration procedure because the number of messages handled at each node in the re-registration procedure for the above-mentioned schemes are the same. The total queuing delay for the initial registration is given by summing the queuing delay of total messages handled at each node:

$$Q_{E-IMS} = \frac{1}{\mu_u - 1 \times 1 \times \lambda - 0.6 \times \mu} + \frac{2}{\mu_C - C_N \times 2 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_C - C_N \times 3 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_C - C_N \times 3 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_H - C_N \times 3 \times \lambda - 0.7 \times \mu} \tag{11}$$

$$Q_{E-OPA} = \frac{1}{\mu_u - 1 \times 1 \times \lambda - 0.6 \times \mu} + \frac{2}{\mu_C - C_N \times 2 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_C - C_N \times 3 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_C - C_N \times 3 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_H - C_N \times 3 \times \lambda - 0.7 \times \mu} \tag{12}$$

$$Q_{C-IMS} = \frac{1}{\mu_u - 1 \times 1 \times \lambda - 0.6 \times \mu} + \frac{2}{\mu_C - C_N \times 2 \times \lambda - 0.6 \times \mu} + \frac{3}{\mu_C - C_N \times 3 \times \lambda - 0.6 \times \mu} + \frac{2}{\mu_C - C_N \times 2 \times \lambda - 0.6 \times \mu} + \frac{2}{\mu_H - C_N \times 2 \times \lambda - 0.7 \times \mu} \tag{13}$$

We assume the service rate μ of all nodes is 250 messages/sec [11]. Equations (11), (12) and (13) are plotted as the function of the value λ in Figure 7. The queuing delay of the C-IMS AKA is the lowest among the three schemes.

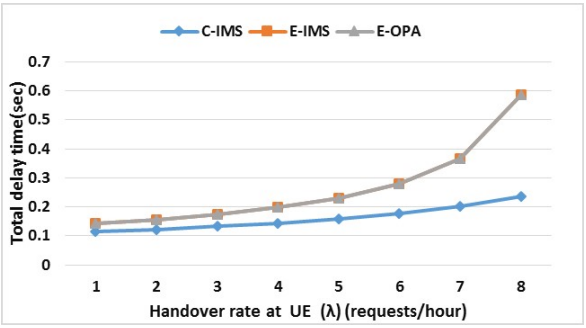


Fig. 7 The queuing delay of the initial registration procedure.

5 Conclusions

This paper proposes a certificate-based IMS authentication and key agreement (C-IMS AKA). It can reduce the number of messages handled at the S-CSCF and the HSS, as compared with current alternatives. We analyze the C-IMS AKA in terms of message propagation delay and queuing delay. Analytical results show that the C-IMS AKA is more efficient than the E-IMS AKA and the E-OPA. Moreover, the certificate will be not only used in the registration procedure but applied to access other IMS services; thus, the C-IMS AKA could reduce the number of messages exchanged at the IMS entities when the UE requests a certificate-based IMS service.

Acknowledgment This work was sponsored in part by *Ministry of Science and Technology of Taiwan* under grant *MOST 104-2221-E-194-009-*.

References

1. GPP TS 23.228: Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS). v13.4.0 (2015)
2. Rosenberg, J., Schulzrinne, H., Camarillo, G., Peterson, J., Sparks, R., Handley, M., Schooler, E.: SIP: session initiation protocol. IETF RFC 3261, June 2002. Available at http://www.ietf.org/rfc/rfc3261.txt
3. Lin, Y.-B., Chang, M.-F., Hsu, M.-T., Lin-Yi, W.: One-pass GPRS and IMS authentication procedure for UMTS. IEEE Journal on Selected Areas in Communications **23**(6), 1233–1239 (2005)
4. Huang, C.-M., Li, J.-W.: Efficient and provably secure IP multimedia subsystem authentication for UMTS. The Computer Journal **50**(6), 739–757 (2007)
5. Long, X., Joshi, J.: Enhanced one-pass IP multimedia subsystem authentication protocol for UMTS. In: 2010 IEEE International Conference on Communications (ICC). IEEE (2010)

6. GPP TS 23.401: General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access. V12.5.0 (2014)
7. Niemi, A., Arkko, J., Torvinen, V.: Hypertext transfer protocol (HTTP) digest authentication using authentication and key agreement (AKA). IETF RFC 3310 (2002)
8. Chiang, W.-K., Kuo, P.-C.: IMS-based centralized service continuity. Wireless Personal Communications **68**(3), 1177–1195 (2013)
9. Kleinrock, L.: Queuing systems, Theory, vol. 1. Wiley (1975)
10. Bessis, T.: Improving performance and reliability of an IMS network by co-locating IMS servers. Bell Labs Technical Journal **10**(4), 167–178 (2006)
11. Munir, A., Gordon-Ross, A.: SIP-based IMS signaling analysis for wimax-3g interworking architectures. IEEE Transactions on Mobile Computing **9**(5), 733–750 (2010)

Software Optimizations of NTRUEncrypt for Modern Processor Architectures

Shay Gueron and Fabian Schlieker

Abstract This paper describes software optimizations for the post-quantum public-key encryption scheme NTRUEncrypt. We build upon the, to the best of our knowledge, fastest open-source NTRUEncrypt library and optimize it by taking advantage of AVX2 and AVX512 SIMD instructions as well as the AES-NI built-in encryption functions.

We show that, on modern processors, using AVX2 yields performance gains of 23% for encryption and 37% for the decryption operation. For the future AVX512 we use a publicly available emulator, since no supporting processor is on the market yet, and report that for the decryption only about half of the instructions compared to the current code are needed to be executed.

Furthermore, we propose replacing the SHA hash functions by pipelined AES-NI for faster randomness generation. With both optimizations enabled, we achieve performance improvements of 1.82x for encryption and 1.74x for decryption with a parameter set that provides 256 bits of security.

Keywords NTRU · NTRUEncrypt · Post-quantum public-key encryption · Software optimization · AVX2 · AVX512 · AES-NI

1 Introduction

NTRUEncrypt is a public-key encryption scheme (part of the NTRU cryptosystem), whose security is based on hard problems on lattices [9]. This makes it different

S. Gueron(✉)
Department of Mathematics, University of Haifa, Haifa, Israel
e-mail: shay@math.haifa.ac.il

S. Gueron
Intel Corporation, Israel Development Center, Haifa, Israel

F. Schlieker
Horst Görtz Institute for IT-Security, Ruhr University Bochum, Bochum, Germany

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_18

from today's most popular public-key cryptosystems such as RSA (based on the hardness of factoring large integers) and (Elliptic-Curve-) Diffie-Hellman Key Exchange (based on the hardness of calculating discrete logarithms). In addition to the high performance of NTRUEncrypt, its reliance on a different hard problem makes it an interesting research topic for the following reason.

Factorization and discrete logarithm problems can be solved in polynomial time using Shor's algorithm on a quantum computer [18], effectively rendering most currently deployed public-key cryptography broken [2]. To date, a sufficiently large quantum computer has not been built, but putting it into practice is an active area of academic research, and also being explored by large companies as well as government agencies [17, 19]. Estimations for the remaining time-frame until such a quantum computer appears vary from ten to forty years.

This motivates the search for alternative cryptosystems (so called post-quantum cryptography) already today. New cryptographic standards only gain confidence over time. For example, Elliptic Curve Cryptography (ECC) along with its advantages over RSA was introduced in 1985. However, it was not until last year that robust ECC standards and implementations were rolled out in significant parts of the internet. Likewise for confidential material like health-care data or state secrets, long-term protection is needed. Even if it is not feasible to decrypt it today, all this data could potentially be stored by an adversary until a quantum computer exists and easily be decrypted then.

There is a variety of post-quantum schemes which are based on different mathematical problems: from lattice theory, coding theory, hash functions to multivariate-quadratic-equations. NTRU was one of the earliest proposals, and has therefore already received some scrutiny by cryptographers. This, in turn, has already led to more secure and more efficient parameter choices for NTRUEncrypt and some general confidence in its soundness. Updated parameter choices were proposed in [7, 8], after a meet-in-the-middle (MitM) attack and recent quantum cryptanalysis rendered some of the initial recommendations insecure [3, 10]. In 2008, NTRUEncrypt was standardized in IEEE Standard 1363.1 [12].

The designers of NTRU contributed an open source reference implementation in 2013 [16]. Earlier (2011), another NTRU software library implementation was publicly released, and is still under active development. It offers much better performance than the reference implementation [1]. To the best of our knowledge, there are currently no other complete and self-contained open-source C implementations available, and we therefore use the faster code as a basis for our optimizations and comparison.

1.1 AVX2 and AVX512 Architectures

The Advanced Vector Extensions 2 (AVX2) architecture [14] was introduced in 2013 with Intel's processor generation codenamed "Haswell". It has several characteristics: Single-Instruction-Multiple-Data (SIMD) registers of 256 bits; many new

(integer) instructions that operate on these registers, including enhanced versions of Streaming SIMD Extensions (SSE) to operate on wider registers; non-destructive destination (i.e., the result does not necessarily overwrite one of the sources). For example, one useful new AVX2 instruction that we use here, is VPERM2I128: permute a selection of the four 128-bit-lanes from two source registers or zeros into the two lanes of the destination register.

The subsequent evolution of AVX2 is AVX512. It doubles the registers' widths (to 512 bits) and doubles the number of registers (from 16 to 32) [15]. Furthermore, a handy new feature is the introduction of masks as an additional operand for many instructions. Zero-masks can be used to zero-out chosen bits during operations while write-masks allow to blend in selected bits from another source operand. For example, it is possible to add registers a and b, then write to the destination register either the bits of the result or bits from a third register c, according to a mask k. Processors with AVX512 support are not yet available, but are expected to be released in the near future.

1.2 AES-NI

Increased demand for high bandwidth encryption led Intel to introduce built-in hardware support to its processors, in the form of new processor instructions. The Advanced Encryption Standard New Instructions (AES-NI) are a set of instructions that can be used by software to carry out AES encryption, decryption and key expansion, in various modes of operation, and achieve high performance. They also provide inherent side-channel protection against (cache-) timing attacks, with an easy-to-use API that minimizes the risk of implementation errors (compared to a complete software implementation) [4, 5]. By now, processors from different vendors such as AMD, ARM, IBM and Sun, followed Intel's architecture, and added AES-NI (or equivalent) instructions to their architectures.

2 Preliminaries

We give a short overview of the encryption scheme and the mathematical operations that NTRUEncrypt relies upon, as these are our optimization targets. A comprehensive description of the algorithms can be found in [12].

2.1 NTRUEncrypt

The NTRUEncrypt lattice-based public-key encryption scheme consists of three basic algorithms: key generation, encryption and decryption. We focus here on the

encryption and decryption algorithms. The associated calculations are carried out in a ring of polynomials, which completely hides the underlying lattice structure and allows for efficient computation.

The cryptosystem is parameterized by a number of variables. First are the ring parameters (N, q) where N is a prime and q is usually a power of 2, for efficient modular reduction. They define the ring $\mathcal{R} = \mathbb{Z}_q[X]/(X^N - 1)$, also denoted as ring of truncated polynomials or convolution polynomials. The message space modulus is denoted p, and is chosen to be much smaller than, and co-prime to the modulus q (the standard fixes $p = 3$).

In the encryption scheme, hash functions are used for generating randomness and two functions, SHA-1 and SHA-256, are standardized, based on the security level. During the encryption/decryption of one message, they are invoked in two different contexts and in each a specified number of times.

The standard defines several flavors of parameter choices, with emphasis on speed, key and message sizes, or a trade-off between both. Some exemplary parameter sets from the balanced category are listed in Table 1. The set we chose for measurements of our optimizations is highlighted.

Table 1 Four standardized parameter sets for different levels of security. Hash function runs from the two different contexts in the scheme are summed up here.

Security level	Name	N	q	p	Hash function	Hash function invocations
112 bit	EES541EP1	541	2048	3	SHA-1	27
128 bit	EES613EP1	613	2048	3	SHA-1	30
192 bit	EES887EP1	887	2048	3	SHA-256	26
256 bit	**EES1171EP1**	**1171**	**2048**	**3**	**SHA-256**	**36**

The algorithms are carried out as follows.

Key Generation

- Generate two ternary polynomials in $\mathcal{R}$, denoted f and g, with coefficients chosen uniformly at random from $\{-1, 0, 1\}$. Ensure that they are invertible modulo q and p (otherwise, re-generate).
- Compute the inverses $f_q = f^{-1} \mod q$ and $f_p = f^{-1} \mod p$.
- Compute $h = f_q * g \mod q$.
- The pair (f, g) is the private key, h is the public key.

Encryption

- Encode the plaintext message into a polynomial m.
- Generate a polynomial r with N small coefficients chosen uniformly at random (not necessarily from $\{-1, 0, 1\}$).
- Compute the ciphertext as $e = pr * h + m \mod q$ using the public key h.

Decryption

– Given e and f, the private key, compute $a = f * e \mod q$.
– Represent the coefficients of a in the interval $[-\frac{q}{2}, \frac{q}{2}]$.
– Obtain the plaintext by computing $c = f_p * a \mod p$.

For simplicity, we omit here the description of the whole SVES-3 encryption scheme. SVES-3 instantiates NTRUEncrypt in order to allow provable protection against adaptive chosen ciphertext attacks in the random oracle model (similar to OAEP+ for RSA). Further information is given in [11] and in the standard.

2.2 Polynomial Multiplication and Reduction

Due to the reduction modulo $(X^N - 1)$, polynomial multiplication in the ring $\mathcal{R}$ is equivalent to cyclic convolution. The product of two polynomials f and g is computed as $f * g = \sum_{i+j=k \mod N} f_i \cdot g_j \mod q$, which can be easily parallelized by convolving multiple coefficients at once. The vectorized implementation does this with 8/16/32 coefficients processed in parallel by each SSE/AVX2/AVX512 instruction execution.

In case one of the two operands is a ternary polynomial, the convolution can be computed even more efficiently with no multiplications needed at all. This is done by just iterating over the nonzero coefficients of the ternary polynomial and, depending on whether it is 1 or -1, either add or subtract the corresponding coefficient of the other polynomial.

Each multiplication operation is finalized by modular reduction of all coefficients by the parameter q. This is also very easy to do in parallel and, due to q being a power of two, performed with a "cheap" logical `and` operation.

2.3 Polynomial Generation

As shown above, each algorithm needs to generate random numbers (from different ranges). One way to achieve this is using a hash function with a seed and varying inputs, and use the hash digest outputs to generate pseudorandom numbers. The pseudorandom generation standards allow for using SHA-1 and SHA-256 as the underlying hash functions. This is the method used by [1].

Profiling multiple NTRUEncrypt operations shows that the computation of these hashes is a significant workload. Yet, it is not the one-way function property that is needed here, but rather that the outputs are indistinguishable from random (i.e., the function acts as a pseudorandom function (PRF)).

This can also be achieved with a keyed pseudorandom permutation (PRP). We propose to replace the use of the standardized hash functions with AES-256, which is (by its design) believed to be a good approximation for a PRP (when the number of

times it is invoked is $<< 2^{64}$ times). As a matter of fact, the NTRUEncrypt standard already limits this to 2^{32}, which is more than enough in practice. This modification allows us to leverage the high performance of AES instructions which are part of the modern processor architectures.

3 Implementation

The source code from the GitHub repository [1] implements the NTRUEncrypt encryption scheme according to the standard, as explained in Section 2. We focus our optimizations on the functions for polynomial generation and arithmetic, because these computations consume the largest portion of the encryption and decryption time. In the baseline implementation, polynomials are generated by a pseudorandom bit generator based on the SHA-1 or SHA-256 hash functions and highly optimized code from the OpenSSL library. We replace these hash functions with the AES block cipher which can be computed substantially faster using native instructions.

The arithmetic (i.e., polynomial convolution and coefficient-wise modular reduction) is already vectorized using SSE instructions. Moreover, a function for encoding the polynomial coefficients into an array of bytes is also vectorized for q fixed to 2048 (this choice is common to all the standardized parameter sets). In total, there are six different functions that are vectorized with SSE instructions. These are the optimization targets. However, in this work, we did not further improve the encoding function because performance profiling revealed that it consumes only around 1-2% of the total computation time. For the remaining functions we implemented both, AVX2 and AVX512 optimized versions.

3.1 AVX2 and AVX512 Vectorization

The polynomial convolution and modular reduction functions are promoted to AVX2 and AVX512 by: a) converting the code to use double (quadruple) the register size; b) using the AVX2 (AVX512) instructions; c) halving (quartering) the number of loop iterations. Hence, more data is processed in parallel, leading to better performance.

Most of the instructions translate naturally from SSE to their analogs over wider registers. An exception is the full-register shifting. While the SSE architecture provides the instruction PSLLDQ for (left-) shifting a 128-bit register by a chosen number of bytes, the AVX2/AVX512 VPSLLDQ analogs do not operate on the full 256-bit register. Rather, their 128-bit "lanes" are shifted independently. The resulting performance is likely to be better if the code is adapted to this behavior and written to use the 128-bit lanes separately, instead of straightforwardly imitating the AVX2/AVX512 analogous flow of the SSE implementation. In general, when one wants to port existing SSE code to use wider registers, it might be helpful to keep this in mind.

Nonetheless sometimes there is no way around this, so in our source code we explore ways to replicate the shift on a full 256/512-bit register with the lowest possible latency (for AVX2 using VPERM2I128 followed by a VPALIGNR; in AVX512 with only one VPERMW instruction, thanks to the new mask operands).

3.2 AES-NI for Pseudorandom Number Generation

We use highly optimized and pipelined AES code (based on [4, 5]) to replace the SHA-1 / SHA-256 hash functions. The number of times the hash function is called is specified in the standard for each parameter set. As the AES output is 16 bytes, half of the SHA-256 hash digest, we call the function twice as many times. However, because AES runs at ~0.65 cycles/byte compared to around ~2.7 cycles/byte using an optimized SHA-256 (parallelized to use AVX2 [6]), this is more than amortized and we save cycles.

To have a deterministic initial state, we first hash the (variable length) seed at the beginning like in the baseline implementation to have a fixed length starting point. In the subsequent iterations, the baseline implementation concatenates the hashed seed with an incrementing counter value. We adjust this to use the hash digest as the AES-256 key and use the counter value as plaintext to be encrypted under this key.

By operating on multiple plaintext blocks in parallel, we leverage pipelining capabilities of AES-NI and achieve high throughput (see Listing 1 for a descriptive source code snippet).

```
uint8_t aes_key[32];
memset(aes_key, 0, 32);

uint8_t aes_input[32*min_calls_mask];
memset(aes_input, 0, 32*min_calls_mask);

uint8_t aes_output[32*min_calls_mask];
memset(aes_input, 0, 32*min_calls_mask);

AES_KEY enc_key;

memcpy(aes_key, Z, hlen);

AES_set_encrypt_key(aes_key, 256, &enc_key);

while (counter < 2*min_calls_mask) {
    uint16_t counter_endian = htons(counter);
    memcpy(aes_input+counter*16, counter_endian, sizeof(counter_endian));
    counter++;
}

AES_ECB_encrypt(aes_input, aes_output, 32*min_calls_mask, enc_key.KEY, enc_key.nr);
```

Listing 1 C code to generate the needed stream of pseudorandom bytes using AES-NI instead of SHA-1/SHA-256. The seed is passed in Z, hashed, and used as the AES-256 key; the incrementing counter value is the plaintext to be encrypted; `min_calls_mask` is the number of hash function runs as specified in the standard

4 Results

We report here two types of results. For the AVX2 vectorization and AES-NI pseudorandom generator, we compare the encryption and decryption performance measured on an Intel processor of the latest 6th Intel® Core™ Generation (architecture codename "Skylake"). For the experiments, we disabled the Intel® Turbo Boost Technology, Intel® Hyper-Threading Technology, and Enhanced Intel Speedstep® Technology. We use the test bench included in the baseline implementation, compiled with `gcc` version 5.2.0 and full optimizations enabled ("`-O3`"). For consistency, we depict and compare only measurements obtained with the EES1171EP1 parameter set that offers 256 bits of security.

For the AVX512 vectorization, we follow a different approach, since supporting processors are not yet available. Therefore, we use the publicly available Intel® Software Development Emulator [13] tool to run the different implementations. The emulator is able to measure the number of instructions required for carrying out the encryption/decryption.

4.1 AVX2 and AVX512 Simulation Results

We measure the number of instructions executed to perform one encryption and one decryption using the Intel SDE tool. This way, we can also give estimations on further improvements through AVX512.

For encryption, our AVX2 implementation needs only 79% of the instructions executed by the baseline implementation, while AVX512 further reduces this to 68%. Decryption yields an even greater improvement with 69% of the number of SSE instructions (AVX2) down to 53% using AVX512 (Figure 1, left graph). Although the instruction count does not necessarily correspond to the future observed performance, it is a good indicator for what can be expected (see the correspondence between the two graphs for the AVX2 results).

4.2 AES-NI and Overall Results on Silicon

The test bench included in the library implementation reports the number of encryption/decryption operations that can be executed per second. By leveraging wider AVX2 instructions, we are able to perform 23% more encryptions and 37% more decryptions per second than with SSE.

Interestingly, using the AES instructions for randomness generation has much more impact in the encryption than in the decryption (as opposed to the vectorization where it is the other way round). A speedup of 34% more encryption and 19% more decryption operations per second can be observed.

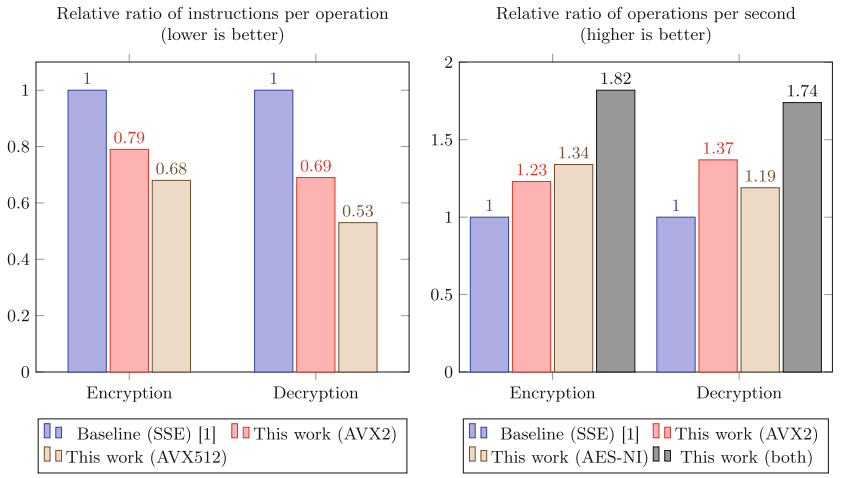


Fig. 1 The relative performance improvement achieved for the EES1171EP1 parameter set, where the values are ratios relative to the SSE baseline implementation [1] that is scaled to 1. In the graph on the left, the metric is the number of instructions needed for a single encryption/decryption operation (i.e., lower is better) for AVX2 and AVX512 vectorization (Section 3.1). Since there is not yet a processor with AVX512 support, the values were measured with the Intel SDE tool. The graph on the right displays the ratio of possible encryption/decryption operations per second (i.e., higher is better) using AVX2 as described in Section 3.1, AES-NI explained in Section 3.2, and both optimizations at the same time. The numbers were obtained using the test bench included in the baseline implementation.

In the graph on the right of Figure 1, we illustrate the speedup using AES-NI as well as the cumulative performance improvement achieved by using both our techniques together. Overall, our optimizations achieve 82% more NTRUEncrypt encryption operations and 74% more decryption operations per second than the baseline implementation.

5 Discussion

This paper explored the performance advantage of leveraging advanced architectures with wide registers and hardware encryption in order to speed up NTRUEncrypt. We showed how an open-source implementation with SSE optimizations can be easily ported to AVX2 and to the future AVX512 architectures. This reduces the number of executed instructions, and therefore improves the performance of the associated cryptographic operations.

We note that porting an existing SSE code is not always straightforward, because some SSE instructions do not map naturally to an AVX2 (AVX512) analog. This can be overcome by using permutation instructions, though at some additional (noticeable) latency overhead. We argue that a re-structured algorithm can potentially turn

out more efficient than naively elevating it to larger vector sizes and their according instructions.

Furthermore, the current NTRUEncrypt standard uses the cryptographic hash functions SHA-1 and SHA-256 for simplicity and efficiency in generating pseudorandom bytes indistinguishable from a stream of truly random bytes. Hardware instructions for cryptographic functions were not available at the time the standard was established. With the introduction of hardware instructions for the AES block cipher, whose output is also indistinguishable from a stream of random bytes, much higher performance is possible. Consequently, replacing a hash function by AES, as shown in this paper, can be a worthwhile consideration.

The threat of quantum computers to currently deployed cryptosystems is still only theoretical today, but cannot be dismissed as huge effort is invested into making it practical. Therefore, the post-quantum secure NTRUEncrypt could possibly get rolled out to many production systems (e.g., web servers), and our performance improvements would directly benefit such usages.

Finally, we remark that our code is contributed to the open source community. It is available online at https://github.com/fschlieker/libntru.

Acknowledgements This research was supported by the PQCRYPTO project, which was partially funded by the European Commission Horizon 2020 research Programme, grant #645622.

References

1. C Implementation of NTRUEncrypt (2015). https://github.com/tbuktu/libntru
2. Buchmann, J.A., May, A., Vollmer, U.: Perspectives for cryptographic long-term security. Commun. ACM **49**(9), 50–55 (2006)
3. Fluhrer, S.: Quantum cryptanalysis of NTRU. IACR Cryptology ePrint Archive 2015, 676 (2015). http://eprint.iacr.org/2015/676
4. Gueron, S.: Intel's new AES instructions for enhanced performance and security. In: 16th International Workshop on Fast Software Encryption, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers, pp. 51–66 (2009)
5. Gueron, S.: Intel® Advanced Encryption Standard (AES) New Instructions Set, September 2012. https://software.intel.com/sites/default/files/article/165683/aes-wp-2012-09-22-v01.pdf
6. Gueron, S., Krasnov, V.: Simultaneous hashing of multiple messages. J. Information Security **3**(4), 319–325 (2012)
7. Hirschhorn, P.S., Hoffstein, J., Howgrave-Graham, N., Whyte, W.: Choosing ntruencrypt parameters in light of combined lattice reduction and MITM approaches. In: Proceedings of the 7th International Conference on Applied Cryptography and Network Security, ACNS 2009, Paris-Rocquencourt, France, June 2-5, 2009, pp. 437–455 (2009)
8. Hoffstein, J., Pipher, J., Schanck, J.M., Silverman, J.H., Whyte, W., Zhang, Z.: Choosing parameters for NTRUEncrypt. IACR Cryptology ePrint Archive 2015, 708 (2015). http://eprint.iacr.org/2015/708
9. Hoffstein, J., Pipher, J., Silverman, J.H.: NTRU: a ring-based public key cryptosystem. In: Proceedings of the Third International Symposium on Algorithmic Number Theory, ANTS-III, Portland, Oregon, USA, June 21-25, 1998, pp. 267–288 (1998)

10. Howgrave-Graham, N.: A hybrid lattice-reduction and meet-in-the-middle attack against NTRU. In: Proceedings of the 27th Annual International Cryptology Conference on Advances in Cryptology - CRYPTO 2007, Santa Barbara, CA, USA, August 19-23, 2007, pp. 150–169 (2007)
11. Howgrave-Graham, N., Silverman, J.H., Whyte, W.: Choosing parameter sets for NTRUEncrypt with NAEP and SVES-3. IACR Cryptology ePrint Archive 2005, 45 (2005). http://eprint.iacr.org/2005/045
12. IEEE Std. 1363.1-2008: IEEE Standard Specification for Public Key Cryptographic Techniques Based on Hard Problems over Lattices (2009)
13. Intel Corporation: Intel® Software Development Evaluator (SDE). http://software.intel.com/en-us/articles/intel-software-development-emulator
14. Intel Corporation: Intel® 64 and IA-32 Architectures Software Developer's Manual, September 2015
15. Intel Corporation: Intel® Architecture Instruction Set Extensions Programming Reference, August 2015
16. NTRUOpenSourceProject: Open Source NTRU Public Key Cryptography Algorithm and Reference Code (2015). https://github.com/NTRUOpenSourceProject/ntru-crypto
17. Rich, S., Gellman, B.: NSA seeks to build quantum computer that could crack most types of encryption, January 2, 2014. https://www.washingtonpost.com/2014/01/02/8fff297e-7195-11e3-8def-a33011492df2_story.html
18. Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: 35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, USA, November, 20-22, 1994, pp. 124–134 (1994)
19. Simonite, T.: IBM shows off a quantum computing chip, April 29, 2015. http://www.technologyreview.com/news/537041/ibm-shows-off-a-quantum-computing-chip/

Vulnerabilities and Mitigation Methods in the NextGen Air Traffic Control System

Sachiko Sueki and Yoohwan Kim

Abstract The air traffic control (ATC) systems have been modernizing and standardizing the automation platforms in recent years in order to control increased number of flights. In 2004, FAA started transforming the nation's ground-based ATC system to a system which uses satellite-based navigation and other advanced technology, called NextGen. The NextGen system deploys Internet Protocol based network to communicate and heavily relies on computerized information system and digital data, which may introduce new vulnerabilities for exploitations. Many vulnerabilities of NextGen stem from the increased interconnection of systems through wireless networks. For instance, a critical part of the NextGen, Automatic Dependent Surveillance – Broadcast, which transfers essential information via wireless network without encryption, is an easy target for attackers. There have been some deployments of security measures but still lack in critical system. In this study, we present the potential vulnerabilities of the NextGen ATC systems and their possible solutions.

Keywords ATC · Automatic dependent surveillance – broadcast · Data communications · System wide information management · En route automation modernization and replacement · Terminal automation modernization and replacement

1 Introduction

In 2013, Presidential Policy Directive 21 identified 16 critical infrastructure sectors which provide essential services that are vital to the nation's safety, prosperity, and well-being. One of the sectors is transportation systems, which includes aviation such as aircrafts, air traffic control (ATC) systems, airports, and landing strips. Cyber systems including ATC, tracking, and communication systems provide a fundamental capability in keeping the nation's transportation system safe and operational [1].

S. Sueki(✉) · Y. Kim
Department of Computer Science, University of Nevada, Las Vegas, NV, USA
e-mail: suekis@unlv.nevada.edu, Yoohwan.Kim@unlv.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_19

The ATC systems have been modernizing and standardizing the automation platforms in recent years. Comprehensive ATC under the direction of the Federal Government started in 1936 in the United States [2]. The ATC system evolved as the number of flights increased. At present the so-called legacy system is managed based on a combination of radar and computer technology. Some of the technologies used in the legacy system were developed as far back as the 1940s [3]. The system is not capable of navigating in oceanic airspace and remote land regions because of its ground-based operations. In general, aircrafts operating in these regions follow inefficient procedural separation methods. These inefficient control systems are causing flight delays.

According to FAA long-range forecasts, aircraft operations are going to increase to approximately 81 million and 96 million in 2020 and 2030, respectively [4]. In order to control increased number of flights, in 2004, FAA started transforming the nation's ground-based ATC system to a system which uses satellite-based navigation and other advanced technology, called NextGen [5]. The improvements from the legacy system to the NextGen [6] are listed in Table 1. The NextGen system deploys Internet Protocol (IP)-based network to communicate and heavily relies on computerized information system and digital data, which may not be adequately secure and thus vulnerable to exploitations. The facilities, aircrafts and pilots communicate using point-to-point communication lines in the legacy system while in the NextGen they happen through system-wide interconnectivity. Furthermore, modern aircrafts increasingly rely on Internet for many purposes. Such interconnectivity in their information systems presents elevated cyber-attack opportunities.

There have been reported cyber-attack incidents in the ATC systems. For instance, in 2006, Federal Aviation Administration (FAA) ATC system was infected by a virus forcing it to shut down a portion of the ATC systems in Alaska. In 2008, an attacker took over the critical FAA network servers and gained an access to shut down the servers [7]. Earlier in 2015, FAA network was attacked with malicious software [8]. According to the report [7], more than 800 cyber incident alerts were issued to the Air Traffic Organization responsible for ATC operations during the Fiscal Year 2008. As the NextGen ATC systems replace the legacy systems, opportunities for cyber attackers can further increase. Even though FAA has taken steps to protect the systems from cyber-based threats, significant security control weaknesses still exist [5]. Therefore, it is crucial to understand and mitigate the vulnerabilities that exist in the ATC systems and its counter measures taken such as use of encryption and authentication technologies.

In this study, a literature review, surveys and analyses are being conducted to identify vulnerabilities that exist in the NextGen ATC systems and we suggest possible mitigation measures. First, a brief explanation of the NextGen air traffic control systems is given in section 2. Then, vulnerabilities and their possible solutions are discussed in sections 3 and 4, respectively. Finally, some promising mitigation methods are discussed in section 5.

Table 1 Improvements from the legacy system to the NextGen

Legacy System		NextGen
Voice Communication	→	Digital Communication
Ground-Based Navigation	→	Performance-Based Navigation
Radar Surveillance	→	Satellite-Based Surveillance
Constrained Automation	→	Flexible Automation, Decision-Support Tools
Disparate Point-to-Point System	→	Integrated System and Information Distribution

2 NextGen Air Traffic Control Systems

The NextGen ATC systems consist of six major programs, which are primarily FAA internal system upgrades that are necessary to deploy additional capabilities. The six programs are Automatic Dependent Surveillance – Broadcast (ADS-B), Data Communications (Data Comm), En Route Automation Modernization (ERAM), Terminal Automation Modernization and Replacement (TAMR) National Airspace System (NAS) Voice Switch (NVS), and System Wide Information Management (SWIM) [9].

ADS-B uses a Global Navigation Satellite System to determine aircraft's own position and broadcasts its position, speed and altitude to ground stations or other aircrafts in the vicinity over a radio frequency. On board GPS receiver gives aircraft's own position and velocity. Then, the transmitting subsystem, ADS-B Out, periodically broadcasts its information via a message. The ATC stations on the ground and nearby aircrafts equipped with the receiving subsystem, ADS-B In, can receive these messages. The ADS-B functions most likely are integrated into currently used 1090ES data link, which predominantly uses the 1090 MHz frequency for communications and data is transmitted by blocks utilizing pulse position modulation (PPM). ADS-B is the central component in the NextGen and ADS-B Out must be equipped in aircrafts by January 1, 2020 [9].

Data Comm communicates with digitally-delivered messages between ATC and pilots replacing radio voice communications. Routine instructions such as departure clearances and weather-avoiding reroutes are directly sent to the flight deck, reducing potential miscommunications. The initial en route services are expected in 2019 and full operational capability at air route traffic control centers in 2021 [9].

ERAM is a scalable system combining flight plan information with information from surveillance sources such as ADS-B data to automate many air traffic control functions and support controller decisions. ERAM serves as the platform of data sharing, digital communications and trajectory-based operation. The system will be used in air traffic controllers at the air route traffic control centers. Full deployment of ERAM is planned to be completed by 2015 [9]. Other air traffic facilities and government agencies such as airport towers, FAA command center, automated flight service stations, Department of Homeland Security, Department of Defense, and U.S. Customs and Border Protection, are now connected to the centers via ERAM.

The TAMR program converts the automation platforms for near airports and high altitude to a single automation platform called Standard Terminal Automation Replacement System (STARS). The system meets operational requirements for ADS-B and improves flight plan processing with a 4-D trajectory (lateral, vertical, horizontal and time). Full deployment of TAMR is planned for 2020 [9].

NVS uses router-based communications linked through the FAA Telecommunications Infrastructure network. NVS provides a capability of sharing communication resources unlike the current voice switches operated independently at individual facilities. The capability of NVS is still in development and NVS is currently on schedule to start operational test in Seattle in the fiscal year 2019 [9].

SWIM is the base for data-sharing and currently distributes weather and flight planning information to the NAS users through a single point of access. The SWIM program is to implement a set of information technology principles in the NAS and provides users with relevant and commonly understandable information. Raw surface data from airport towers are converted to accessible information via SWIM Terminal Data Distribution System (STDDS). The information is, then, available from Terminal Radar Approach Control (TRACON) to airlines and airports through SWIM messaging services. The SWIM Flight Data Publication Service (SFDPS) will improve flight data sharing using standard Flight Information Exchange Model with a Globally Unique Flight Identifier. SFDPS is currently available only in the SWIM research and development domain [9].

As NVS is still in the development phase, not much information is available, therefore NVS is excluded from the study. On the other hand, ADS-B is a critical part of the NextGen, which transfers essential information via wireless network. Therefore, there have been several studies in its vulnerabilities and mitigations as reported in [10]-[19]. The majority of vulnerabilities and solutions reported here is related to ADS-B.

3 Vulnerabilities

ADS-B has been developed without security in mind and its signals are public over a known frequency. Furthermore, transmissions are not encrypted or authenticated. Therefore, ADS-B is subject to various types of attacks. The attacks include eavesdropping, jamming, message injection, message deletion and message modification [11]. Eavesdropping is highly possible since the complexity of an attack is low and ADS-B sends unencrypted messages over a broadcast medium. Furthermore, as pointed out by [12], there are services to aid eavesdropping such as commercially available ADS-B receivers [20], digitized live ADS-B data available to public via the Internet [21], and open-source GNU radio module available for sophisticated traffic analysis [22]. Jamming is another simple attack that can cause denial-of-service (DOS) problems [13]. Jamming is a common problem to all wireless communication. However, the impact is severe because of importance and criticality of the transmitted data. Since ADS-B does not have any authentication measures, an attack with cheap and simple technological means can be used

to inject non-legitimate message into the communication system. Message injection can display ghost aircrafts on a cockpit display forcing the pilots to change their course and/or velocity. Injecting multiple messages can cause ghost aircrafts flooding which can lead to DOS of the controller's surveillance system [12]. Message deletion can be achieved by transmitting the inverse of the signal broadcast by a legitimate sender [11] or by causing large enough number of bit errors for the receiver to recognize a message as corrupted. Message deletion can cause aircraft disappearance. Message modification can be typically done by overshadowing or bit-flipping during transmission. Overshadowing is to replace part of the message by sending a high-powered signal. Bit-flipping is the signal converting by flipping bits from 1 to 0 or the other way around by superimposing the signal. By combining aforementioned attacks, attackers can achieve trajectory modification, indication of false alarms such as hijacking, and aircraft spoofing [14].

Data Comm allows controllers to electronically send instructions to the cockpit display with a push of a button. Instructions are sent via data link without any authentication, which can be susceptible to possible cyber-attacks [23]. The information is supposed to be seen by controllers and applicable pilots. However, hobbyists who have appropriate radio equipment can monitor and decode transferred information, making this vulnerable to cyber-attacks similar to the ADS-B messages.

The major vulnerability of SWIM is attributed from net-centric exchanges, which potentially increases the chance of the system to be compromised and damaged. Damage can potentially spread to other systems on the network when one of the systems connected to an IP network is compromised. Furthermore, SWIM is vulnerable to a man-in-the-middle attack since it does not provide any end-to-end confirmation that messages are sent or received on the network [24]. The other concern is the use of various software. Since SWIM does not include control information and safety-critical information such as surveillance data, any relatively inexpensive commercial software or internally developed software, and open source software can be used. Such software may have a list of widely known vulnerabilities that can increase opportunities for unauthorized access and malicious-code execution.

ERAM and TAMR are used for analyzing data within the centers. However, they also communicate with other facilities in order to obtain data or transfer information. Vulnerabilities of the programs are associated with the interconnectivity of systems, which can lead to an unauthorized access or a malicious code attack to the ATC system.

Finally, NextGen's potential vulnerability arises from the IP network connectivity. Approximately 36 percent of the ATC system in the NAS is connected to IP network and the connections are projected to grow to 50 to 60 percent by 2020 [5]. The legacy system's point-to-point connections, which co-exist with the NextGen system, can be also compromised because of increased connectivity with IP network.

4 Mitigation Methods

Most of the security mitigation is focused on ADS-B because of its importance in the NextGen System and an existence of several potential vulnerabilities. In order to secure ADS-B, there are two distinct approaches, secure broadcast authentication and secure location verification [15]. Secure broadcast authentication is to secure the communications and can be used to prevent and/or detect attacks in a unidirectional broadcast network. Node-based authentication (the authenticity of the hardware) includes non-cryptographic schemes on the physical layer and cryptography. Secure location verification authenticates the claimed location using data from the senders and other ADS-B participants. The techniques include multilateration, group verification, distance bounding, Kalman filtering, data fusion, and traffic modeling.

Non-cryptographic schemes such as fingerprinting are to identify suspicious activity in a network. Fingerprinting is to identify what they are based on the unique characteristics of devices such as operating system, drivers, clocks, and radio circuit [25]. ADS-B currently does not utilize the schemes to secure the system. However, there are three possible techniques that may be employed in ADS-B. The techniques are software-based fingerprinting, hardware-based fingerprinting, and channel/location-based fingerprinting [26]. Software-based fingerprinting uses distinctly different patterns or behavior of software operating on wireless equipment. Hardware-based fingerprinting is to identify devices based on unique hardware differences such as differences in turn-on/off transient, modulation of a radio signal, and clock skew. Another recent technique is physically unclonable functions (PUF) [27], which uses specifically implemented circuits to create unique and secure signatures. Channel/location-based fingerprinting is based on received signal strength, channel impulse response, or the carrier phase. Randomize/uncoordinated frequency hopping/spreading is one type of non-cryptographic schemes, which is different from fingerprinting. Such schemes use Uncoordinated Frequency Hopping, Uncoordinated Direct-Sequence Spread Spectrum or Randomized Differential Direct-Sequence Spread Spectrum [28]. By regularly changing communication frequencies of a sender and a receiver, they wait for a chance to be at the same communication channel.

Cryptography is one of the common methods to secure communication in wireless networks, which requires distribution of encryption keys to vast participants of ADS-B systems [15]. One of the proposed methods in [16] is the use of public key cryptography with challenge/response format. Retroactive key publication is a variation of public key cryptography, which sends a partial public key with every message [29]. The receivers who buffer all the messages can decrypt them using the collected public key. On the other hand, a recent study by [17] suggested the use of Staged Identity-Based Encryption, which uses receiving parties' identities as public keys for encryption.

Multilateration technique can geometrically calculate an unknown location from a precise distance between four or more known locations [18]. Currently a preferred method for location verification is multilateration by utilizing the time difference of arrival. Time difference of arrival can be obtained from several antennas in different locations that receive the same signal at different times. The other way of utilizing multilateration is group verification. Group verification is to verify the location claimed by a non-group aircraft in-flight using multilateration done by a group of aircrafts [15].

Distance bounding is to find the upper bound of locations by sending a challenge to the receiver and getting a response [15]. The upper bound is calculated based on the fact that electromagnetic waves do not travel faster than the speed of light. The actual location can be found via a difference in distance between the measurements from the various ground stations.

Kalman filtering is already used to filter and smoothen GPS position data in messages in ADS-B [19]. In every time step, the measured variables and the error covariance are projected. It then updates the estimations and error covariance with the actual measurements. The filtering is an important tool for sorting out noisy signals and smoothing over missing data for multilateration approach. The filter was also suggested to use in one of the distance bounding protocols.

Data fusion is to verify the data obtained within the system by comparing it with the data coming from other independent sources, e.g., the fusion of ADS-B and radar [19].

Traffic modeling can be created from historical data and machine learning methods to detect deviations from normal ADS-B behavior [15]. The technique can be also applied to establish red flags for intrusion detection system so that the technically and physically impossible data are reasonably dropped to reduce the strain on the ADS-B system and prevent spoofing and DOS attacks.

Similar to the ADS-B message protections, messages sent using Data Comm can be secured using cryptography such as Elliptical Curve Cryptography Asymmetric Public-Key Infrastructure. One such application suggested by [23] was the Protected Aircraft Communications Addressing and Reporting System (PACARS), which provides end-to-end message protection and/or authentication. PACARS uses Elliptical Curve Cryptography Asymmetric Public-Key Infrastructure.

SWIM provides a comprehensive set of technical security controls via its infrastructure and the Common Data Transport (CDT) security services implemented in the FAA telecommunications infrastructure [30]. The CDT provides firewalls that can defend unauthorized access, IP address spoofing, traffic rerouting, session hijacking, and some forms of DOS attacks. The CDT security also includes link and network layer cryptographic security providing authentication, integrity and confidentiality. The SWIM security provides identity and access management, which includes authentication, authorization and auditing. Authentication may be provided by Kerberos or Public Key Infrastructure. Kerberos is a network authentication protocol based on the shared-secret cryptography [31]. From a functional perspective, SWIM and CDT share an intrusion detection system, security information management and a public key infrastructure certificate authority to monitor events, keep event logs, and provide digital certificates.

In order to guard the overall IP-networked systems, an enterprise approach is being developed, which views IP-networked systems as subsystems within the large enterprise-wide system [5]. The subsystems can interoperate while enterprise-wide set of shared cybersecurity controls such as continuous monitoring, incident detection and response, internal policy enforcement, and identity and key management are available. However, in order to fully utilize an enterprise approach, a holistic threat model can be a valuable tool. A holistic threat model might provide a likely compromise and dangers associated with potential consequences [5].

The NextGen ATC systems' vulnerabilities and their solutions are listed in Table 2.

5 Advantages and Disadvantages of Methods

Mitigation methods must be adaptable to large-scale deployment in order to be implemented in the ATC systems. Furthermore, for practical purpose, the cost and the complexity of deployment need to be considered. Therefore, adding new hardware or modifying the existing systems is hard to implement. For instance, ADS-B is a unidirectional broadcast while many of the proposed methods such as PUF, channel/location-based fingerprinting, and distance bounding require bidirectional communication. Hardware-based fingerprinting with clock skew and data fusion require additional data while PUF and Randomize/uncoordinated frequency hopping/spreading require additional hardware.

Table 2 The NextGen ATC System's Vulnerabilities and Their Solutions

NextGen Program	Vulnerabilities	Mitigation Methods
ADS-B	Eavesdropping	Randomize/uncoordinated frequency hopping/spreading, Public key infrastructure
	Jamming	Randomize/uncoordinated frequency hopping/spreading
	Message Injection	Fingerprinting, Public key infrastructure, Multilateration, Distance bounding, Kalman filtering, Traffic modeling Group verification, Data fusion
	Message deletion	Randomize/uncoordinated frequency hopping/spreading
	Message modification	Fingerprinting, Randomize/uncoordinated frequency hopping/spreading, Public key infrastructure, Multilateration, Distance bounding, Kalman filtering, Traffic modeling Group verification, Data fusion
Data Comm	Cyberattack, Eavesdropping	PACARS (Elliptical Curve Cryptography Asymmetric Public key infrastructure)
SWIM	Network compromise (unauthorized access, spoofing, DOS, etc)	Common data transport security service, Public key infrastructure, Kerberos
	Man-in-middle attack	Public key infrastructure, Kerberos
	Commercial, internal, open source software	Timely installation of patches
Others	IP-networked system	Enterprise wide system, Holistic threat model

Multilateration and Kalman filtering appear to be promising methods to incorporate in ADS-B. Multilateration is currently in use in comparatively short distances while this technique still cannot fully utilize ADS-B because of decreased accuracy over long distances. Kalman filtering is already being used in ADS-B related systems but it is slightly difficult in positional claim verification of aircraft-to-aircraft systems. However it is feasible and highly scalable. For secure communications, one of the promising methods is retroactive key publication even though this technique can be susceptible to a memory-based DOS attack and requires a slight modification of data. A successful implementation of these three methods can provide protections from message injection and modification attacks. In order to address all possible vulnerabilities in ADS-B, new protocols or methods may be required.

Some security measures are already in place in SWIM but not in rest of the systems. Although those systems are not as crucial as ADS-B, it is important to mitigate their vulnerabilities. Therefore, further studies are needed.

6 Summary and Conclusions

The NextGen ATC systems' vulnerabilities come mainly from the increased interconnection of systems through wireless network. There have been some deployments of security measures such as in SWIM but not in other critical systems such as ADS-B. Even though there are many security measures proposed, their practical use is still questionable because of its broadcast nature and wide operational range. One solution cannot protect variety of attacks. Therefore, the system must be protected with a defense-in-depth and an enterprise approach.

References

1. Homeland Security: Transportation systems sector-specific plan, an annex to the national infrastructre protection plan (2010). http://www.dhs.gov/transportation-systems-sector (accessed September 15, 2015)
2. Federal Aviation Administration: FAA historical chronology, 1926–1996 (2011). https://www.faa.gov/about/history/chronolog_history/ (accessed October 3, 2015)
3. Federal Aviation Administration: Navigation programs – history (2015). https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/history/ (accessed October 3, 2015)
4. Federal Aviation Administration: FAA long-range aerospace forecasts, fiscal years 2020, 2025 and 2030 (2007). http://www.faa.gov/data_research/aviation/long-range_forecasts/media/long07.pdf (accessed October 3, 2015)
5. United States Government Accountability Office: Air Trafic Contorl, FAA needs a more comprehensive approach to address cybersecurity as agency transition to NextGen, Report to Congressional Requesters, GAO-15-370 (2015)
6. Bradford, S.: NextGen progress and ICAO. In: Integrated Communications, Navigation and Surveillance Conference (ICNS 2014), pp. 1–22, April 8–10, 2014

7. Office of Inspector General: Review of web applications security and intrusion detection in air traffic control systems, 2009. Audit Report, Report ID: FI-2009-049, p. 23 (2009)
8. Sternstein, A.: Exclusive: FAA computer systems hit by cyberattack earlier this year. In: Nextgov (2015). http://www.nextgov.com/cybersecurity/2015/04/faa-computer-systems-hit-cyberattack-earlier-year/109384/ (accessed October 3, 2015)
9. Federal Aviation Administration: NextGen implementation plane (2015). https://www.faa.gov/nextgen/media/NextGen_Implementation_Plan-2015.pdf (accessed October 3, 2015)
10. Danev, B., Zenetti, D., Capkun, S.: On physical-layer identification of wireless devices. ACM Computer Surveys **45**(1), 1–29 (2012)
11. Strohmeier, M., Lenders V., Martinovic, I.: Security of ADS-B: state of the art and beyond. arXiv preprint arXiv:1307.3664 (2013)
12. Schäfer, M., Lenders, V., Martinovic, I. (eds.): Experimental analysis of attacks on next generation air traffic communication. In: 11th International Conference on Applied Cryptography and Network Security. Lecture Note in Computer Sciences, pp. 253–271 (2013)
13. McCallie, D., Butts, J., Mill, R.: Security analysis of the ADS-B implementation in the next generation air transportation system. International Journal of Critical Infrastructure Protection **4**(2), 78–87 (2011)
14. Amin, S., Clark, T., Offutt, R., Serenko, K.: Design of a cyber security framework for ADS-B based surveillance systems. In: Systems and Information Engineering Design Symposium (SIEDS 2014), pp. 304–309, April 25, 2014
15. Strohmeier, M., Lenders, V., Martinovic, I.: On the security of the automatic dependent surveillance-broadcast protocol. IEEE Communications Surveys & Tutorials **17**(2), 1066–1087 (2015). Secondquarter 2015
16. Viggiano, M., Valovage, E., Samuelson, K., Hall, D.: Secure ADS-B authentication system and method, U.S. Patent 7730307 B2, June 1, 2010
17. Hableel, E., Baek, J., Byon, Y., Wong, D.S.: How to protect ADS-B: confidentiality framework for future air traffic communication. In: IEEE Conference of on Computer Communications Workshops (INFOCOM WKSHPS), April 26–May 1, 2015, pp. 155–160 (2015)
18. Nijsure, Y., Kaddoum, G., Gagnon, G., Gagnon, F., Yuen C., Mahapatra, R.: Adaptive air-to-ground secure communication system based on ADS-B and wide area multilateration. IEEE Transactions on Vehicular Technology **99**, 1. doi:10.1109/TVT.2015.2438171
19. da Silva, J.L.R., Brancalion, J.F.B., Fernandes, D.: Data fusion techniques applied to scenarios including ADS-B and radar sensors for air traffic control. In: 12th International Conference on Information Fusion, Fusion 2009, pp. 1481–1488, July 6–9, 2009
20. iPad Pilot News: Which ADS-B reciever should I buy? (2015). http://ipadpilotnews.com/2015/10/ads-b-receiver-buy-2/ (accessed November 10, 2015)
21. Flightrader24: Live Air Traffic. Available from http://www.flightradar24.com/
22. Sharan, R., West, N.: The comprehensive GNU radio archieve network. http://www.cgran.org/ (accessed November 10, 2015)
23. Storck, P.E.: Benefits of commercial data link security. In: Integrated Communications, Navigation and Surveillance Conference (ICNS 2013), pp. 1–6, April 22–25, 2013

24. Jaatun, M.G., Faegri, T.E.: Sink or SWIM: information security requirements in the sky. In: Eighth International Conference on Availability, Reliability and Security (ARES 2013), pp. 794–801, September 2–6, 2013
25. National Air Traffic Controllers Association: NextGen now. Quaterly E-Publication **1**(4) (2015). http://www.natca.org/safety.aspx?zone=Safety%20and%20Technology&pID=4586 (accessed October 25, 2015)
26. Zensg, K., Govindan, K., Mohapatra, P.: Non-cryptographic authentication and identification in wireless networks [security and privacy in emerging wireless networks]. IEEE Wireless Communications **17**(5), 56–62 (2010)
27. Devadas, S., Suh, E., Paral, S., Sowell, R., Ziola, T., Khandelwal, V.: Design and Implementation of PUF-Based "Unclonable" RFID ICs for anti-counterfeiting and security applications. In: 2008 IEEE International Conference on RFID, pp. 58–64, April 16–17, 2008
28. Chengzhi, L., Huaiyu, D., Liang, X., Peng, N.: Analysis and optimization on jamming-resistant collaborative broadcast in large-scale networks. In: 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), pp. 1859–1863, November 7–10, 2010
29. Kwon, T., Hong, J.: Secure and efficient broadcast authentication in wireless sensor networks. IEEE Transactions Computer **59**(8), 1120–1133 (2010)
30. Stephens, B.: Security architecture for system wide information management. In: The 24th DASC 2005 Digital Avionics Systems Conference, vol. 2, p. 10, October 30–November 3, 2005
31. Neuman, C., Yu, T., Hartman, S., Raeburn, K.: The Kerberos network authentication service (V5), RFC4120 (July 2005)

Pushing the Limits of Cyber Threat Intelligence: Extending STIX to Support Complex Patterns

Martin Ussath, David Jaeger, Feng Cheng and Christoph Meinel

Abstract Nowadays, attacks against single computer systems or whole infrastructures pose a significant risk. Although deployed security systems are often able to prevent and detect standard attacks in a reliable way, it is not uncommon that more sophisticated attackers are capable to bypass these systems and stay undetected. To support the prevention and detection of attacks, the sharing of cyber threat intelligence information becomes increasingly important. Unfortunately, the currently available threat intelligence formats, such as YARA or STIX (Structured Threat Information eXpression), cannot be used to describe complex patterns that are needed to share relevant attack details about more sophisticated attacks.

In this paper, we propose an extension for the standardized STIX format that allows the description of complex patterns. With this extension it is possible to tag attributes of an object and use these attributes to describe precise relations between different objects. To evaluate the proposed STIX extension we analyzed the API calls of the credential dumping tool Mimikatz and created a pattern based on these calls. This pattern precisely describes the performed API calls of Mimikatz to access the LSASS (Local Security Authority Subsystem Service) process, which is responsible for authentication procedures in Windows. Due to the specified relations, it is possible to detect the execution of Mimikatz in a reliable way.

Keywords Cyber threat intelligence · STIX · CybOX · Complex pattern · Attribute relation

1 Introduction

Nearly every day, attackers try to compromise computer systems and network infrastructures to steal, for example, bank account information or to exfiltrate intellectual

M. Ussath(✉) · D. Jaeger · F. Cheng · C. Meinel
Hasso Plattner Institute (HPI), University of Potsdam, 14482 Potsdam, Germany
e-mail: {martin.ussath,david.jaeger,feng.cheng,christoph.meinel}@hpi.de

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_20

property. Different signature-based [12] and anomaly-based approaches were developed to prevent and detect such attacks. These approaches are implemented in different security systems, such as firewalls, Intrusion Detection Systems (IDSs) and anti-virus products, which are widely used in current environments. In general, the deployed security systems provide a decent protection level against standard attacks, which are often highly automated and do not have specific targets. Beside standard attacks, there are also more sophisticated attacks that target specific companies, public authorities and organizations. The main characteristics of such sophisticated attacks, which are sometimes also called Advanced Persistent Threats (APTs), are that they are able to bypass existing security systems and they are also capable to largely infiltrate the target network without triggering alerts. To achieve this, the adversaries behind these attacks use different approaches, such as encryption, polymorphic malware or the imitation of legitimate activities to hinder the prevention and detection. Due to the fact that current security systems are not able to reliably prevent and detect sophisticated attacks, the sharing of cyber threat intelligence information becomes increasingly important.

The sharing of attack details, such as used command-and-control servers, malware or attack tools, is not a new approach. Different organizations and companies provide such information for quite some time, for example through investigation reports, blacklists or web services. Due to the fact that the different information feeds most often use a custom format, the automated processing and utilization of this information is complex. To simplify the usage of cyber threat intelligence information different formats, such as YARA, Open Indicators of Compromise (OpenIOC) or Structured Threat Information eXpression (STIX), were proposed. With the help of these standardized formats not only the efficiency of threat intelligence sharing should be improved, but also the integration of this information into prevention and detection systems should be much simpler [15]. Through the sharing and the automated processing of threat information, potential targets are capable to detect attacks in a faster manner. Also the sharing of classified threat intelligence information within a group of trusted partners is possible. By directly sharing such information, it is no longer necessary to wait until the vendor of a utilized security system releases for example a signature or pattern update.

Most publicly available cyber threat intelligence feeds parse other information feeds that use a custom format and provide the same information in a standardized format. Through the conversion of the information, various threat information details can be easily used to support prevention and detection activities. Although the new formats allow a very detailed description of attacks, most feeds only share basic information, such as IP addresses, hash sums, static patterns and domain names. This type of information might be helpful for standard attacks, but for complex and sophisticated attacks more details are needed for a reliable and precise detection. Especially for complex multi-step attack activities it is relevant to identify all related steps, otherwise it might not be possible to detect them, as single steps might look benign. The existing cyber threat intelligence formats do not support the description of complex patterns that are needed to detect more sophisticated attack activities.

In this paper, we propose an extension of the STIX format that allows the description of complex patterns. These pattern are based on object attributes and their relations. Through the sharing of such patterns it is possible to detect relevant attack activities that are performed during sophisticated attacks. Furthermore, we will evaluate the proposed extension by describing a pattern of the well-known credential dumping tool Mimikatz.

The remainder of this paper is structured as follows: Section 2 briefly describes three commonly used formats to share cyber threat intelligence information. Furthermore, in this section we also present three public services that provide threat intelligence information. Afterwards, in Section 3 we propose a STIX extension that allows the description of complex patterns. In the following section we evaluate the extension by describing the characteristic API calls of the credential dumping tool Mimikatz within a complex pattern. Finally, we conclude our paper and propose future work in Section 5.

2 Related Work

Through the sharing of threat intelligence information it is possible to create and provide new patterns and signatures, which should allow an effective prevention and detection of attacks in a timely manner. The usage of standardized formats [9] is one relevant requirement for the sharing of threat intelligence among different parties [14], [7]. Only if all involved parties have a common understanding about the meaning of the shared information [10], [4], an efficient and automated processing of this information can be implemented. In the following we will describe three common formats for the sharing of threat intelligence information.

Different public services, such as Open Threat Exchange from AlienVault [1] and X-Force Exchange from IBM [8], already use some of these formats to provide attack details. To show what kind of threat intelligence is currently shared, we will briefly describe the provided information of three different public services.

2.1 YARA

YARA [2] is an application that uses rules to identify potentially malicious files or suspicious processes in memory. Each YARA rule contains different characteristics that can be used to detect malicious files or processes. A rule usually has a *strings* and a *condition* section to describe relevant characteristics. The most common method is to use text strings, hexadecimal strings and regular expressions to describe relevant parts of a file or process. In the condition section a Boolean expression describes the relations of the strings and other properties that need to be matched to correctly identify the described file or process. With various properties it is possible to specify for example the file size, number of sections or exported functions. However, if various operators are used, the specified condition of a YARA rule can be relatively complex.

Listing 1.1 shows a YARA rule that describes characteristics of the TinyZBot malware. This rule was published as part of an investigation report [5] to allow potential targets to search for this type of malware. The rule defines nine different strings that are characteristic for this malware. The condition shows that the rule will return true if a file or process contains one of the four possible combinations of the specified strings.

```
rule TinyZBot
{
 strings:
    $s1 = "NetScp" wide
    $s2 = "TinyZBot.Properties.Resources.resources"
    $s3 = "Aoao WaterMark"
    $s4 = "Run_a_exe"
    $s5 = "netscp.exe"
    $s6 = "get_MainModule_WebReference_DefaultWS"
    $s7 = "remove_CheckFileMD5Completed"
    $s8 = "http://tempuri.org/"
    $s9 = "Zhoupin_Cleaver"

 condition:
    ($s1 and $s2) or ($s3 and $s4 and $s5) or ($s6 and $s7 and $s8) or $s9
}
```

Listing 1.1 TinyZBot YARA Rule [5]

YARA and the corresponding rules are focused on the detection of potentially malicious files and processes. This is also the reason why the format does not support the description of other characteristics that are for example relevant to detect malicious network traffic. Furthermore, the format does not provide easy to use extension capabilities. Although the detection of suspicious files and processes is crucial for the detection of attacks, there are various other attack details that are relevant and should be shared. Especially in cases where the attackers use legitimate credentials and administrative tools, it is not possible to use YARA to detect such type of malicious activities.

2.2 *Open Indicators of Compromise (OpenIOC)*

OpenIOC [11] is an XML-based format that can be used to describe forensic artifacts of attacks. The single indicators, which contain details of an intrusion, are called IOCs (Indicators of Compromise). The format was proposed by the company Mandiant. Due to the standardization and public availability of the OpenIOC format, also other vendors started supporting the usage of IOCs as threat intelligence format.

OpenIOC supports, by default, more than 500 different IOC terms[1], which can be used to describe different artifacts, such as file and process attributes or network indicators. By sharing and using such IOCs, the identification of forensic artifacts, which belong to an attack, should be much simpler. The design of the format also allows to add extensions for additional artifacts. A single IOC can describe various artifacts of an attack and these characteristics can be linked through operators, such as *OR* as well as *AND*.

[1] http://openioc.org/terms/Current.iocterms

Listing 1.2 shows an IOC that describes the artifacts of the CHOPSTICK backdoor that was used during the APT28 campaign [6]. To ensure clarity of the example, some parts with meta-information were removed. This IOC contains two single artifacts and one combined artifact that are linked with *OR*. The two single artifacts describe a file name (*edg6EF885E2.tmp*) and a name of a handle (*\Device\Mailslot\check_mes_v5555*). The combined artifact is linked with AND and consists of a registry path (*Microsoft\MediaPlayer\E6696105-E63E-4EF1-939E-15DDD83B669A*) and a name of a regstiry value (*chnnl*). If one of the three artifacts, one of the two single artifacts or the combined artifact can be found on a system, then this strongly indicates that this system is or was infected with the CHOPSTICK malware.

```
<ioc id="[...]" last-modified="2014-10-20T18:50:53Z">
 <definition>
  <Indicator id="[...]" operator="OR">
   <IndicatorItem id="[...]" condition="is">
    <Context document="FileItem" search="FileItem/FileName" type="mir"/>
    <Content type="string">edg6EF885E2.tmp</Content>
   </IndicatorItem>
   <IndicatorItem id="[...]" condition="is">
    <Context document="ProcessItem" search="ProcessItem/HandleList/Handle/Name" type="mir"/>
    <Content type="string">\Device\Mailslot\check_mes_v5555</Content>
   </IndicatorItem>
   <Indicator id="[...]" operator="AND">
    <IndicatorItem id="[...]" condition="contains">
     <Context document="RegistryItem" search="RegistryItem/Path" type="mir"/>
     <Content type="string">Microsoft\MediaPlayer\{E6696105-E63E-4EF1-939E-15DDD83B669A}</Content>
    </IndicatorItem>
    <IndicatorItem id="[...]" condition="contains">
     <Context document="RegistryItem" search="RegistryItem/ValueName" type="mir"/>
     <Content type="string">chnnl</Content>
    </IndicatorItem>
   </Indicator>
  </Indicator>
 </definition>
</ioc>
```

Listing 1.2 CHOPSTICK IOC[2]

The OpenIOC standard is capable to describe various forensic artifacts of an attack, which can be used to identify potentially compromised systems. By default, the format only supports the combination of artifacts with an OR or AND operator. In most cases these two operators are enough to describe the relations between artifacts, but for sophisticated attacks it is often necessary to specify the relation in a more detailed way. Especially for multi-step attacks, which sometimes use legitimate credentials and built-in administrative programs, it is essential to identify artifacts that are related for example through a session identifier to ensure that these steps belong to a malicious activity. Furthermore, the standardized OpenIOC format does not support the enrichment of artifacts with relevant context information.

2.3 STIX

STIX [3] is another XML-based format that is still under active development. The corresponding project is led by OASIS organization. Nevertheless, the development

[2] https://raw.githubusercontent.com/fireeye/iocs/master/APT28/bdf7929c-3f0b-4fdd-bcc5-b4a82554ad92.ioc

of the format is open and community-driven. This might be also one reason why numerous vendors and products[3] started to support the STIX format.

The objective of STIX is to describe attacks and relevant context information in a comprehensive way to enable an efficient and effective threat information sharing. The format can make use of eight different core constructs, which are *Campaign*, *CourseOfAction*, *ExploitTarget*, *Incident*, *Indicator*, *Observable*, *ThreatActor* and *TTP* (tactics, techniques and procedures). The CybOX format is used within STIX to describe Observables, but it is an additional project that is developed in conjunction with STIX. To allow a comprehensive description of attacks and context information, the objects of the different core constructs can be linked with other objects. Due to the increasing complexity of sophisticated attacks, STIX was designed to be flexible and extensible to support most cases. Therefore, it is also possible to embed other formats and patterns, such as YARA rules or IOCs, into STIX. Another standard that is developed in conjunction with the STIX project, is the so-called Trusted Automated eXchange of Indicator Information (TAXII) standard that enables a secure and automated sharing of cyber threat intelligence information.

Listing 1.3 shows a simple STIX document[4] that describes an IP address of a command-and-control infrastructure of an attacker. To ensure clarity, the sample was shortened to the relevant parts of the STIX document. First of all, an IPv4 address object (*198.51.100.2*) is defined as an observable. In the second part of the sample the defined observable is referenced to describe the used infrastructure of the attacker as part of a TTP object.

```
<stix:STIX_Package id="[...]" version="1.2">
 <stix:Observables cybox_major_version="1" cybox_minor_version="1">
  <cybox:Observable id="example:observable-c8c32b6e-2ea8-51c4-6446-7f5218072f27">
   <cybox:Object id="[...]">
    <cybox:Properties xsi:type="AddressObject:AddressObjectType" category="ipv4-addr">
     <AddressObject:Address_Value>198.51.100.2</AddressObject:Address_Value>
    </cybox:Properties>
   </cybox:Object>
  </cybox:Observable>
 </stix:Observables>
 <stix:TTPs>
  <stix:TTP xsi:type="ttp:TTPType" id="[...]">
   <ttp:Title>Malware C2 Channel</ttp:Title>
   <ttp:Resources>
    <ttp:Infrastructure>
     <ttp:Type>Malware C2</ttp:Type>
     <ttp:Observable_Characterization cybox_major_version="2" cybox_minor_version="1">
      <cybox:Observable idref="example:observable-c8c32b6e-2ea8-51c4-6446-7f5218072f27"/>
     </ttp:Observable_Characterization>
    </ttp:Infrastructure>
   </ttp:Resources>
  </stix:TTP>
 </stix:TTPs>
</stix:STIX_Package>
```

Listing 1.3 STIX Sample[4]

Due to the design of the STIX format, various issues of previous threat intelligence formats could be addressed. The approach of STIX is to allow a holistic description of attack details and relevant context information, such as confidence levels, which are crucial for sharing threat intelligence information. Although the format allows to describe various relations between different objects, it is not possible to describe more complex attack steps that are connected through certain characteristics.

[3] http://stixproject.github.io/supporters/

[4] https://stixproject.github.io/documentation/idioms/c2-ip-list/

2.4 Standardized Cyber Threat Intelligence Usage

According to a survey of the SANS Institute [15], only 38 % of the organizations are using cyber threat intelligence information in standardized formats. About 46 % of these companies use the STIX format for cyber threat intelligence information. The related standards CybOX (26 %) and TAXII (33 %) are also relatively often utilized. The OpenIOC format is used in 33 % of the organizations that employ standardized formats for their threat intelligence information. This survey shows that the overall adoption of standardized threat intelligence formats is not very high and this probably also has a negative impact on the efficient sharing of such information. One reason for this might be that currently most public services only provide already available information in a standardized format. Therefore, no additional threat intelligence is offered, but further overhead is introduced through the usage of relatively complex formats.

2.5 Existing Services

Hail a Taxii[5] and X-Force Exchange[6] from IBM are two public services that provide threat intelligence information in the STIX format. For transferring the information, both services use the TAXII protocols and data formats. Hail a Taxii converts already existing threat intelligence information, for example from the different malware trackers of abuse.ch or the CyberCrime tracker, into a STIX compliant format. In general this service offers mainly so-called STIX watchlists for domains, URLs, file hashes and IPs. These lists consist of the corresponding indicators without providing any further details or insights. The service of IBM provides mainly investigation reports in the STIX format. Unfortunately, the indicators of the different reports are most often not correctly converted into the corresponding STIX or CybOX objects. In these cases the indicators are only mentioned in the description field of the STIX document, which makes the automated processing of the threat intelligence information very complex. Nevertheless, the X-Force Exchange service also offers context information about IPs, URLs and vulnerabilities.

Open Threat Exchange (OTX)[7] from AlienVault is another service that provides threat intelligence information. This service should provide a platform for the community to share threat intelligence information. Users can create so-called pulses to share information with other community members. To create a pulse it is possible to upload PDF reports, plain text files or STIX documents. The parser of OTX tries to extract relevant indicators like IP addresses, domains, file hashes, file paths and email addresses. The shared information can be downloaded in different formats, such as CSV, OpenIOC 1.0 or 1.1 and STIX. A brief evaluation of the OTX service

[5] http://hailataxii.com/

[6] https://exchange.xforce.ibmcloud.com/

[7] https://otx.alienvault.com/

showed that different indicators were only extracted, because of the fact that the parser identified a product version number as an IP address. Furthermore, most often the shared information was taken from investigation reports or articles and the threat intelligence information was not enriched with relevant context information.

3 STIX Extension

The mentioned public services only provide basic threat intelligence information, such as IP addresses and domain names. These services do not offer any complex patterns that are needed to detect the increasing number of complex malicious activities. Furthermore, also the standardized cyber threat intelligence formats do not support the description of complex patterns and therefore it is not possible to use these formats to describe sophisticated attack activities.

Therefore we extended the existing STIX format in a way that it supports the description of complex patterns with object attribute relations.

3.1 Current Capabilities

The current version of STIX (1.2), including CybOX (2.1), supports mainly two different ways to describe relations or compositions of elements. For very simplistic relations it is possible to use *Composite_Indicator_Expression* (*indicator* schema) or *Oberservable_Composition* (*cybox* schema) elements together with the corresponding *operator* attribute. The value of this attribute can be *OR* or *AND*. Due to the limited values of the *operator* attribute and the fact that it is not possible to describe relations between the different object attributes, the defined relations are very imprecise. Another possible method to describe relations in STIX is to use a *Related_Object* (*cybox* schema) element and one term of the predefined relationship vocabulary [13] (*ObjectRelationshipVocab-1.1*), such as *Created*, *Deleted_by* or *Uploaded_By*. By using a specific relationship vocabulary term, it is also implicitly defined which source and related objects can be used. For example, to describe that a specific process locks a certain file, the *Locked* relationship term can be used in conjunction with a *Process* and *File* object. Although, the second method allows to describe relations in a more fine-granular way, it is also not possible to describe connections between different object attributes.

3.2 Complex Patterns

For the description of complex patterns, it is necessary to specify how single elements are related with each other. These relations are often based on characteristics

or attributes that are identical or similar for different connected elements. Without describing such relations, the identification of complex patterns might be imprecise, due to the fact that elements from different activities are wrongly connected. For example, if different events recorded the malicious activity of an attacker, these events are often connected through a session identifier that ensures that all these events belong to the session of the attacker and not to an unrelated benign activity. Therefore, complex patterns are not only relevant for describing more sophisticated multi-step attacks, but also to define very precise patterns to reduce the false positive rate. Furthermore, complex attack activities often consist of multiple steps and it is usually not possible to identify one characteristic step that can be used to detect such malicious actions in a reliable way. Thus, all relevant steps of an activity and their relations have to be identified. These characteristics can then be used to describe precise patterns.

3.3 Extension

Our proposed STIX extension allows the description of complex patterns by using object attribute relations. Furthermore, this extension uses the flexible design of STIX for a seamless integration into the format. To achieve the latter it is necessary to have a thorough understanding of the format to correctly extend the format schemas and minimize the required changes.

The main idea is to tag certain attributes of an object that are later used to describe object relations. The different relations between the object attributes are specified in a separate part of the STIX document to ensure flexibility.

To allow a seamless integration and the usage of existing CybOX objects, the extension adds an attribute group to the *BaseObjectPropertyType* of the *cybox_common* schema. Due to the fact that all CybOX elements are indirectly based on the *BaseObjectPropertyType*, the added attribute group *ReferenceFieldGroup* will be available within all object properties. Therefore, it is easily possible to extend already described patterns with detailed information about the relations of the different objects. The *ReferenceFieldGroup* only contains the attribute *reference_name* that is used to address the object property and allow the description of related properties of different objects.

For describing complex object property relations, the extension makes use of the *CompositeIndicatorExpressionType* of the *indicator* schema to provide the necessary details. As already explained the *Composite_Indicator_Expression* element, which is based on the *CompositeIndicatorExpressionType*, allows by default the description of relatively abstract object relations. To reuse this existing feature of STIX, the extension adds *COMPLEX* to the list of valid *operator* values. Furthermore, an additional *Relation* element is specified within the *CompositeIndicatorExpressionType*. Each *Relation* element contains a mandatory attribute that describes the type of the object attribute relation, such as *Equals*. Within the *Releation* element, the related objects with their corresponding *id* are referenced, including the *ReferenceName* of

the element property that is used to describe the relation. Due to the fact that every relation consists of at least two elements, the minimal number of related elements within the *Relation* element is two.

4 Evaluation

For the evaluation of the proposed STIX extension, we will describe a pattern for Mimikatz. Mimikatz is a tool that is often used within more sophisticated attacks, because it is capable to dump credentials. The attackers can misuse the gathered credentials for example to compromise further systems or to move laterally in the environment. It is very difficult to distinguish between benign activities and malicious activities that make use of legitimate credentials. Therefore, it is very crucial to detect the execution of Mimikatz. Due to the fact that this tool is open source, potential attackers can easily modify the source code and re-compile it. Thus, it is not possible to use hash sums to detect the Mimikatz executables in a reliable way. However, it is very unlikely that an attacker will change the general approach of the tool.

Therefore, it is necessary to identify the main characteristics and afterwards use these characteristics to create a meaningful pattern. In the case of Mimikatz we analyzed the performed API calls and identified six different calls that are characteristic for this tool and also relevant to get access to the memory of the LSASS (Local Security Authority Subsystem Service) process and dump credentials. The mentioned LSASS process is responsible for authentication procedures in Windows. Based on the identified API calls it is possible to reliably identify Mimikatz processes.

Due to the fact that the API parameters are dynamic and related, it is not possible to define a static pattern with the different functions and the corresponding parameters to precisely detect the execution of Mimikatz. We rather need to describe a pattern that consists of the different API calls and their relations. These relations are based on API parameters as well as values and ensure that the calls are related. This also improves the precision of the described pattern, because it will not match with processes that call the identified APIs in an unrelated way.

For the identification of the characteristic API calls of Mimikatz (version 2.0 alpha), we executed the *privilege::debug* and the *sekurlsa::logonpasswords* command from the Mimikatz command-line interface. Through these commands the process requests additional privileges and afterwards the logon passwords are dumped, including hashes and clear text passwords. Figure 1 visualizes the different API calls and the relations between these calls. All parameters that are statically define within the pattern are printed italic and the colored parameters are used within object relations. First of all the *RtlAdjustPrivilege* API is called to request further privileges that are needed to access the process memory of the LSASS process. Afterwards, Mimikatz needs to identify the process id of the LSASS process. Therefore, the Unicode string "*lsass.exe*" is loaded into the memory with the help of the *RtlInitUnicodeString* API and then this memory address is used within the *RtlEqualUnicodeString* API call to identify the correct process information of the LSASS process

and get access to the process id. To get all information about the running processes, Mimikatz uses the *NtQuerySystemInformation* API call. Thereafter, the identified id is used for the *OpenProcess* API call, which returns a handle to the LSASS process. This handle is then used to read the memory of the process with the help of the *ReadProcessMemory* API call.

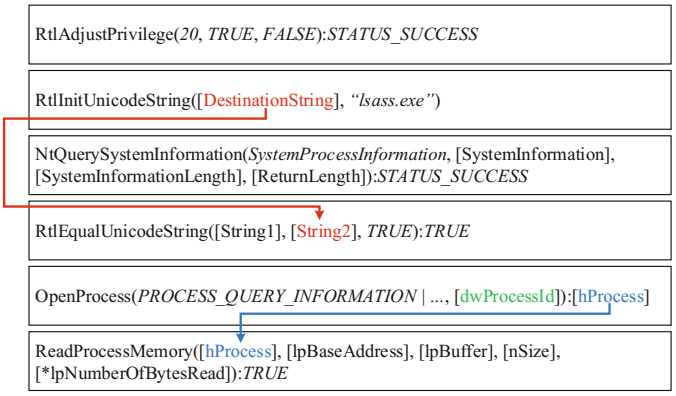


Fig. 1 Characteristic API Calls of Mimikatz

To describe a complex STIX pattern that consists of the identified API calls and the corresponding relations between them, it was necessary to create a new object type that allows the precise description of Windows API calls. This new object type is called *WindowsAPIObjectType* and is based on the default *APIObjectType*. It allows the description of the parameters that are used for the API call and also the return value.

Listing 1.4 shows the complex STIX pattern that describes the different API calls of Mimikatz. To ensure readability, the shown STIX document only contains the description of one relation and one API call. In the first part of the document the different element relations are described. For example the handle that is returned from the *OpenProcess* API call has to be identical to the handle that is used for the *ReadProcessMemory* call. Therefore, within the *Relation* element the corresponding API calls are referenced via the *idref* attribute and the *ReferenceName* of the related element attributes. The single API calls, with their attributes, are defined as *Indicators*. The element attributes that are used for describing relations, have a *reference_name* attribute so that they can be easily referenced within the relations. The values of the different parameters of the API calls are only defined statically, if they are necessary for the successful execution of Mimikatz. Next to the already mentioned API calls, the complex pattern also defines a process object for the LSASS process to obtain the process id of it. The reason for this is that the *NtQuerySystemInformation* returns a complex structure that contains various information about all processes that are

running on the system. This structure cannot be easily used within a pattern and therefore it is necessary to obtain the process id of LSASS from somewhere else. Nevertheless, the Mimikatz process calls *NtQuerySystemInformation* and this is also a relevant characteristic for the execution and cannot be omitted for a precise pattern.

```
<stix:STIX_Package>
 <stix:Indicators>
  <stix:Indicator id="[...]" xsi:type='indicator:IndicatorType'>
   <indicator:Composite_Indicator_Expression operator="COMPLEX">
    [...]
    <indicator:Relation relation_type="Equals">
     <indicator:RelationElement idref="1">
      <indicator:ReferenceName>hProcess</indicator:ReferenceName>
     </indicator:RelationElement>
     <indicator:RelationElement idref="2">
      <indicator:ReferenceName>hProcess</indicator:ReferenceName>
     </indicator:RelationElement>
    </indicator:Relation>
    [...]
    <indicator:Indicator id="[...]" xsi:type='indicator:IndicatorType'>
     <indicator:Observable id="[...]">
      <cybox:Object id="2">
       <cybox:Properties xsi:type="WinAPIObj:WindowsAPIObjectType">
        <Function_Name xmlns="[...]#APIObject-2">ReadProcessMemory</Function_Name>
        <WinAPIObj:Parameter reference_name="hProcess" position="1"></WinAPIObj:Parameter>
        <WinAPIObj:Return_Value>TRUE<WinAPIObj:Return_Value>
       </cybox:Properties>
      </cybox:Object>
     </indicator:Observable>
    </indicator:Indicator>
   </indicator:Composite_Indicator_Expression>
  </stix:Indicator>
 </stix:Indicators>
</stix:STIX_Package>
```

Listing 1.4 Complex STIX Pattern of Mimikatz API Calls (shortened)

5 Conclusion and Future Work

In this paper, we proposed a generic STIX extension that allows the description of complex patterns. These patterns are based on attribute relations that can be described through the tagging of relevant attributes and the specification of object attribute relations. Due to the fact that the extension changes the *cybox_common* schema and the *indicator* schema, all default objects can be used to define complex patterns.

For the evaluation of our STIX extension we analyzed the API calls of the credential dumping tool Mimikatz and described the characteristic calls within a STIX document that uses complex relations. To support the description of Windows API calls, we also defined a new *WindowsAPIObjectType*. With the specified API calls and their relations, it is possible to detect the execution of Mimikatz in a reliable way.

The evaluation showed that it is possible to use the STIX extension to describe all relevant characteristics of more sophisticated attack activities. Therefore, it is possible to share relevant attack details about complex attacks and use this information to support prevent and detection systems.

For future work it might be interesting to improve existing patterns with additional relations and evaluate the change of the precision. This evaluation could also reveal, if it is always reasonable to invest additional efforts and describe complex patterns with relations. Another task could be to propose and implement a system that is capable to automatically enhance existing patterns, for example with the help of a knowledge base.

References

1. AlienVault: AlienVault Open Threat Exchange (OTX)™ User Guide, October 2015. https://www.alienvault.com/doc-repo/OTX/user-guides/AlienVault-OTX-User-Guide.pdf
2. Alvarez, V.M.: Yara User's Manual (2011). https://yara-project.googlecode.com/files/YARA%20User's%20Manual%201.6.pdf
3. Barnum, S.: Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX™). MITRE Corporation, February 2014. https://stixproject.github.io/getting-started/whitepaper/
4. Costa, D.L., Collins, M.L., Perl, S.J., Albrethsen, M.J., Silowash, G.J., Spooner, D.L.: An ontology for insider threat indicators: development and application. In: Proceedings of the 9th Conference on Semantic Technology for Intelligence, Defense, and Security (2014)
5. Cylance: Operation Cleaver, December 2014. http://www0.cylance.com/assets/Cleaver/Cylance_Operation_Cleaver_Report.pdf
6. FireEye Labs: APT28: A Window Into Russia's Cyber Espionage Operations? October 2014. https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-apt28.pdf
7. Haass, J.C., Ahn, G.J., Grimmelmann, F.: Actra: a case study for threat information sharing. In: Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security, pp. 23–26. ACM (2015)
8. IBM: IBM X-Force Exchange Data Sheet, April 2015. http://public.dhe.ibm.com/common/ssi/ecm/wg/en/wgd03055usen/WGD03055USEN.PDF
9. Kampanakis, P.: Security Automation and Threat Information-Sharing Options. Security Privacy, 42–51. IEEE, September 2014
10. Kul, G., Upadhyaya, S.: A preliminary cyber ontology for insider threats in the financial sector. In: Proceedings of the 7th ACM CCS International Workshop on Managing Insider Security Threats, pp. 75–78. ACM (2015)
11. Mandiant: An Introduction to OpenIOC (2011). http://openioc.org/resources/An_Introduction_to_OpenIOC.pdf
12. Meier, M.: A model for the semantics of attack signatures in misuse detection systems. In: Information Security. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg (2004)
13. MITRE Corporation: Object Relationships. http://cyboxproject.github.io/documentation/object-relationships/
14. Serrano, O., Dandurand, L., Brown, S.: On the design of a cyber security data sharing system. In: Proceedings of the 2014 ACM Workshop on Information Sharing and Collaborative Security, pp. 61–69. ACM (2014)
15. Shackleford, D.: Who's Using Cyberthreat Intelligence and How? SANS Institute, February 2015. http://www.sans.org/reading-room/whitepapers/analyst/cyberthreat-intelligence-how-35767

Analyzing Packet Forwarding Schemes for Selfish Behavior in MANETs

Asad Raza, Jamal Nazzal Al-Karaki and Haider Abbas

Abstract The selfish behavior of nodes in Mobile Ad hoc Networks (MANETs) is a serious issue, which negatively impacts the packet forwarding service, causes the issue of availability and degrades the quality of service in networks. Selfish nodes try to conserve their energy and bandwidth by avoiding participation in routing and packet forwarding. To address this problem; the research community has proposed many solutions. We have studied these solutions from two perspectives i.e. context based solutions and context free solutions. In this paper, we are going to present the analysis of some of the well-known protocols which offer solutions to mitigate selfish behavior of nodes.

Keywords Mobile adhoc networks · Selfish behavior · Mobile nodes · Packet forwarding

1 Introduction

MANETs are collection of mobiles nodes which form multi-hop wireless network that is self-organizing and it doesn't rely on conventional infrastructure. Almost all the routing protocols for MANETS work with the hypothesis that participating mobile nodes will forward the data packets to neighboring nodes voluntarily [19]. Due to the fact that there is no infrastructure in MANETs, the participating nodes

A. Raza(✉) · J.N. Al-Karaki
Abu Dhabi Polytechnic (Institute of Applied Technology), Abu Dhabi, UAE
e-mail: {asad.raza,jamal.alkaraki}@adpoly.ac.ae

H. Abbas
Center of Excellence in Information Assurance, King Saud University, Riyadh, Saudi Arabia
e-mail: hsiddiqui@ksu.edu.sa

H. Abbas
National University of Sciences and Technology, Karachi, Pakistan
e-mail: haiderabbas-mcs@nust.edu.pk

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_21

are responsible for the execution of networking functions. Mobile nodes rely on one another for packet forwarding to far-off destinations in such type of networks. In this scenario, intermediate nodes operate as routers in order to provide the services of packet forwarding between remote nodes. Forwarding of packets from source to destination takes places through the selected intermediate routes which are described by routing protocols and algorithms. Packet routing and forwarding are at the core functionalities of the network layer [20] and mutual cooperation among the nodes is vital for the ad-hoc networks to remain operational.

Packet forwarding service depends on the goal of the application. In case when the nodes have a common goal their mutual cooperation is essential to complete the task for example, in battlefield and rescue operations. In these situations the nodes will be cooperative by nature but they may not cooperate if the packet forwarding is only utilizing their resources and not giving them any benefit. Nodes try to save their energy resources in the later case. The energy resource limitation of mobile nodes causes a new predicament in MANETs which degrades the network performance and in severe cases leads to deny of service. In order to save its battery a node may start behaving selfishly in the network by not forwarding packet received from other nodes. A Selfish node is a node in a packet forwarding process that utilizes the resources of other nodes to forward its packets but it does not participate in forwarding their packets. In mobile ad-hoc networks saving energy resources for a node is desirable, but behaving selfishly is considered to be illegitimate. This deviation corresponds to a potential threat to one of the critical network requirement i.e. availability.

This paper discusses the selfish behavior of mobile nodes during packet forwarding in MANETs and critically analyzes some of the well-known techniques for solving this problem. This paper is organized in four sections. In section 2 we have discussed the related research work. In Section 3 we have analyzed some of the popular techniques along with advantages and disadvantages and finally in Section 4 we present our conclusion.

2 Related Work

An extensive amount of research work has been carried out in MANET on selfish behavior of nodes. Different researchers studied selfish nodes from different dimensions. Selfish node keeps its own resources while consuming the resources of the other nodes. The selfish behavior of node is due to the lack of infrastructure in MANET and different solutions have been suggested for determining selfish nodes in mobile ad-hoc networks [6]. Yang yang et al proposed service negotiation model for mobile ad-hoc networks [1]. Chengqi Song and Qian Zhang (2010) classified the solutions to tackle with selfish behavior of nodes into two categories which are context-based solutions and context free solutions. Context-based solutions are based on the technique of recording selfish behavior of nodes and later on punishing them in future whereas in context-free solutions a packet should be transmitted through a route path no matter how many nodes behave selfishly.

Context-based solutions have common features like they observe the nodes' behavior, identify the nodes having selfish behavior and punish these selfish nodes. Context-based techniques are categorized into two types i.e reputation based and credit exchange based. Reputation-based techniques observe the behavior of mobile nodes and based on this behavior the reputation score is increased/decreased [2]. In typical reputation based design nodes observe other nodes behavior and based on these observation calculates the reputation scores. Reputation scores of those nodes increase which forward packets successfully and decreases for those which drop the packets. This reputation scores is then shared with other nodes. With the help of reputation information, nodes which want to transmit packets avoid selfish nodes. Punishment is an important feature with which selfish nodes are punished. This may be in a form that other nodes not to forward their packets and also not to share route information with these selfish nodes. In credit-exchange based techniques the candidate nodes are rewarded with credit point each time they participate in packet forwarding service for other nodes. Whenever a mobile node wants to send datagrams through its neighboring mobile nodes, it has to pay credit to those nodes. In this scenario if the selfish node wants to send its packets, it also has to pay credit points to other nodes. If a selfish node doesn't contribute in packet forwarding, it will automatically run out of credits and will be punished by other nodes by not forwarding its packets. Therefore the selfish nodes are forced to participate in routing data packets if they want to forward their own data packets. The methodology of credit exchange based design is somewhat similar to reputation-based. It also requires monitoring of others nodes behavior. On this observation it is decided to reward credit or not. This credit based system one might question about the security of this protocol. For example, a malicious/selfish node may try to modify its credit point value in order to fool the other nodes. According to the protocol a virtual bank maintains information on nodes' credits and will prevent the modification of credits by malicious nodes [2]. Contrary to this approach, in context-free solutions the mobile nodes transmit their data packets without discovering or identifying selfish nodes. Since there is no identification of selfish nodes in context-free solutions, therefore a different methodology is used to ensure that all the nodes including selfish nodes participate in packet forwarding. In this methodology the destination address of data packets is kept hidden from all the nodes and it is exposed only when all the nodes participate in packet forwarding. It ensures that a selfish node will not drop the packet because it is unaware of the destination address which could be its own address. The COFFEE protocol is one of the most popular context-free solutions for forwarding data packets in MASNETs [2]. It is based on the idea of hiding destination address until packet forwarding is done by all the nodes. In COFFEE path the destination will be an intermediate node and it will have such a path that the destination node will receive the packet twice. It uses layered encryption to ensure that the identity of destination is not revealed by intermediate nodes [4]. In this protocol a secret key is used to encrypt each data packet and the key itself is also encrypted using a layered encryption approach. The COFFEE packets are

broadcasted to all the nodes since there is no information about route. Hyun Jin Kim and Jon M. Peha have presented a novel technique for identifying all the selfish nodes by one or more neighboring nodes [5]. The availability of service and quality are affected immensely by the selfish behavior of nodes, i.e. each node do not want to forward packets to the next node arrived from another node in order to save its own power. To tackle with this challenge fact some solution has been proposed in literature. Different approaches have been discussed in previous works which are based on monitoring and detecting selfish behavior. End-to-end ACKs is methodology for monitoring and ensuring the reliability of routes and routing protocols by acknowledging end-to-end transmission of packets [7]. A problem with end-to-end approach was misbehavior of nodes, which is tackled by another approach called watchdog [8] which monitors the nodes in promiscuous mode. Later on another approach called Activity-Based Overhearing was proposed for the same purpose which is basically the generalization of the watchdog approach. In Activity-Based Overhearing (ABO) a node regularly monitors activities of all of its neighbors in promiscuous mode and keeps an eye on the packets being forwarded by a node its neighbor [9]. In ABO and watchdog approach there are issues related to power control of nodes, which is tackled in two-hop ACK [10]. Probing is another variant of this approach which is the amalgamation of route monitoring and node monitoring. Details on probing technique can be found in [9] [11]. Some reputation based techniques/ approaches have also been discussed in previous work. Based on this general principle of reputation some solutions are presented in studies which are Signed token [13], CORE [14], CONFIDANT [15] and Friends and foes [16].

3 Packet Forwarding Solutions

Two approaches are used to mitigate the packet forwarding problem by selfish nodes: context based and context free. Context based solutions have common features like they observe the nodes' behavior, identify the nodes behaving selfishly and punish these selfish nodes. Context free solutions broadcast data packets without identifying selfish nodes; therefore a different methodology is used as compared to context-based solutions. Context-free solutions hide the identity of the destination of the transmission of the packets to be sent. Some of the renowned protocols based on above two approaches are discussed in this section.

3.1 Watchdog

Watchdog protocol states that the selfish behavior of nodes should be identified with high probability despite adverse channel condition and interface [8]. Watchdog is a scheme proposed to distinguish selfish behavior of nodes by the process of overhearing. Many further solutions were then based on this technique. Watchdog tries to detect nodes behaving selfishly by over hearing. Watchdog monitors

the behavior of mobile nodes in promiscuous mode. If node seems to be acting selfish, a new route would selected by the source with the help of "path rater" to avoid selfish nodes. Once a selfish node is detected on a route by "path rater" mechanism, this information is relayed back to the source. The source node upon receiving information about selfish node chooses another route. Watchdog mechanism has the following drawbacks: First, It does not punish the detected selfish node. Secondly, watchdog might not be able to detect the selfish nodes in the presence of collision, collusion or partial dropping [8]. Figure1 shows the protocol path of Watchdog.

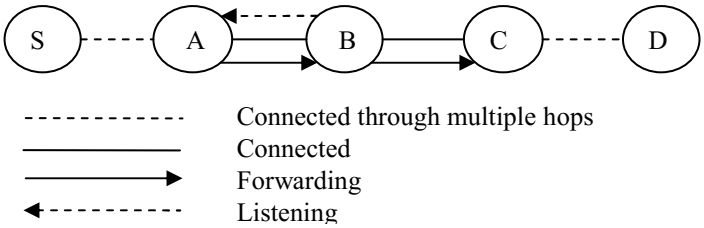


Fig. 1 Watchdog protocols path

3.2 End-to-end ACK's

End-to-end ACK's ensures the reliability of route by acknowledging forwarded packets in an end-to-end manner. When destination node receives the packet successfully, it acknowledges this by sending feedback to originator node. In end-to-end ACK's, successful acknowledgment means that the route does not contain any selfish node. In case if feedback is received after a specified timestamp, it will be considered that the route on which packet is transmitted may have a selfish node on it. A rating is maintained by routing protocols for each route that reflects reliability of route. The rating depends on the successful reception or the failure of ACK's. It is updated when a piece of data is transmitted across the route. This rating, as a metric is used by the routing protocol to choose the most reliable routes. To ensure non-repudiation acknowledge must be signed, otherwise a selfish node misbehaves by not forwarding packets and sending a falsified acknowledgment to the source without being detected. The drawback of this approach is the lack of detection of nodes behaving selfishly. This indicates only the failure or broken route instead of identifying selfish nodes in a network. As this technique helps to avoid transmitting data through unreliable routes, therefore it can be combined with other techniques to improve performance [17].

3.3 CONFIDANT

The CONFIDANT (Cooperation of nodes and Fairness in Dynamic Ad hoc Networks) scheme uses a reputation based mechanism to identify and isolate non cooperative nodes. The main idea of CONFIDANT [15] protocol is to punish non-cooperative selfish nodes by not forwarding their traffic. Detection of selfish nodes is made through neighborhood watch mechanism. The CONFIDANT protocol consists of four components which should be present in every node.

Monitor. The first component is Monitor: which registers the abnormal behavior of nodes and report it urgently to Reputation System.

Trust Manager. Trust Manager is the second component which manages both the inbound and outbound ALARM messages. ALARM messages inform legitimate nodes about the selfish behavior of non-cooperative nodes.

Reputation System. The third component is Reputation System which manages the node view on the reputation of the others.

Path Manager. The Path Manager which is the fourth component of CONFIDANT not only prevents packet forwarding to misbehaving nodes but it also deletes path to such nodes and eventually degrades their ranking.

The CONFIDANT protocol builds a distributed system of Reputation. It enables re-socialization through timeouts in the black lists. The drawback of CONFIDANT protocol is that, the views of nodes reputation are often exchange with each other which causes an over head. A false accusation present by path manager is also a potential threat.

3.4 Nuglets

Nuglets is an economic base approach which is a per-hop payment scheme [18In this methodology the payment units known as Nuglets are stored in a protected and temper-proof hardware module in every node. It is basically based on credit exchange based solution. The principle idea of Nuglet is to utilize virtual currency for buying the coordinated effort. Nuglets are appended to the message and each handing-off node takes Nuglets from the message which can be later utilized to purchase the routing or forwarding of its own message. The implementation of Nuglet module should be in a temper-proof hardware called security module. This module is responsible for maintaining the Nuglets value. Nuglets in a packet are protected from illegitimate modification by cryptographic mechanisms. In order to maintain and establish security association with neighboring nodes, the security module utilizes hello protocol. There are three models for packet forwarding charging [18]. Packet pressure model (PPM), pressure trade model (PTM) and the hybrid one. In PPM model the source node is required to do the Nuglets payment

in advance for its data packet to be forwarded by the other nodes. In PTM the intermediate nodes do the trading of packets i.e. buying packets from the previous node and selling it to the next node.

PPM model reduces network congestion but preventing nodes from forwarding useless data but the downside is that it's difficult to predetermine the number of required Nuglets. However, in PTM model source does not have to predetermine know the number of required Nuglets. The drawback is that it cannot prevent nodes from overloading the network with useless packets.

3.5 COFFEE

The COFFEE Protocol is a context-free solution for resolving the issue of packet forwarding in mobile ad-hoc networks. The main idea of this protocol is to hide the destination address until all packet forwarding is done by all the nodes. It ensures that the identity of destination is not leaked by any node before the completion of packet forwarding process. The COFFEE protocol route path is selected in such a way that the datagram will be delivered at the destination node two times. In COFFEE protocol the packet body and keys are encrypted using layered encryption mechanism. This encryption ensures the secrecy of destination identity. Therefor despite of receiving the packets, intermediate nodes are not able to identify the actual destination. In order to fool the selfish nodes this protocol works in such a way that until the transmission is not completed, any mobile node could be the possible destination of packet. Since the selfish nodes have no information about the destination they don't drop the packets thinking that the packet might be for them. In COFFEE protocol the packets are broadcasted by every node because there is no information about the route. Figure 2 shows the COFEE path established.

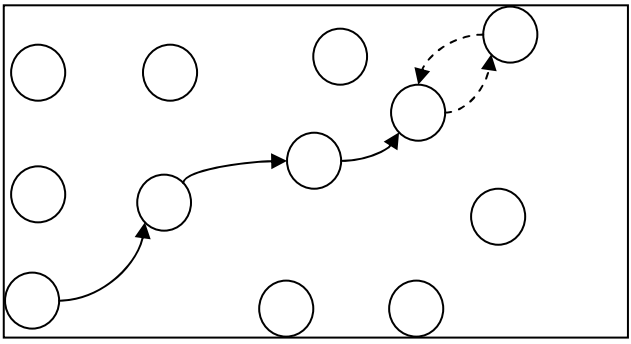


Fig. 2 The COFFEE Path

There are advantages and disadvantages for all these protocols, which are summarized, in the following table.

Table 1 Advantages and disadvantages of protocols

Protocol	Advantages	Disadvantages
Watch-dog	Monitors selfish behavior in promiscuous mode.	No punishment for selfish nodes. Ineffective during collision.
End-to-end ACK's	End-to-end acknowledgment process. Rating mechanism for reliable routes. Avoids transmission through unreliable routes.	Requires a non-repudiation mechanism for ACK's. Cannot detect selfish nodes.
CONFIDANT	Can identify selfish nodes. Punishment for selfish nodes Reputation system for every node	Creates and overhead on network traffic due to reputation system Chances of false accusation
Nuglets	Credit exchange based solutions Tamper proof hardware security module Cryptographic protection of Nuglets.	Computationally expensive due to encryption decryption Causes network congestion if PTM mode is used
COFFEE	Destination address is hidden Packet body and keys are encrypted	Packet is received twice at destination Computationally expensive

4 Analysis and Conclusion

We have discussed some of the well-known solutions for the problem of selfish behavior of nodes in MANETs. We classified these protocols in context based and context free solutions. We have analyzed these protocols with respect to their advantages and disadvantages. The watchdog protocol listens to the nodes promiscuously/passively. The primary downside of this protocol is that it does not punish the detected selfish node. Secondly, watchdog method does not perform well in certain situations like adverse channel, interference, partial dropping hence allowing the selfish nodes to corrupt a single packet and still remain unidentified with a very high likelihood. In End-to-End ACK the route reliability is ensured through ACK and rating and the major issue of this protocol is maintaining ranking tables which creates an overhead. It is also unable to detect selfish nodes like watchdog protocol. In CONFIDANT the major advantage is that it detects the selfish node

by reputation mechanism and punishes them. In this protocol the views of nodes reputation are often exchanged with each other which cause an overhead on the network and a false accusation presented by path manager is another drawback of CONFIDANT protocol. Next we discussed Nuglets protocol, which is credit based solution, credit is used to reward or punish node based on their participation packet forwarding service. Nuglets requirement estimation is difficult task for each node. Based on the above discussion it is concluded that first two approaches .i.e. watchdog and End-to-end ACK only avoid routes having selfish nodes, they neither detect nor punish them. The later approaches i.e. CONFIDANT and Nuglets solve the problems of detection and punishment of nodes but they create overhead problem in networks. The above discussed protocols are context based, next we discussed the context free protocol i.e. COFFEE. In COFFEE packet body and keys are encrypted. This encryption makes sure that nothing can be used to imply the identity of the destination. However Route path distance and processing of packets by each node are issues in this protocol.

Table 2 Comparative Analysis of Protocols

Protocol	Detection of Selfish Nodes	Punishment for Selfish nodes	Network Congestion	Cryptographic protection
Watchdog	Yes	No	No	No
End-to-end ACK's	No	No	No	No
CONFIDANT	Yes	Yes	Yes	No
Nuglets	Yes	Yes	Yes	Yes
COFFEE	No	No	No	Yes

A comparative analysis of these protocols is presented in Table 2. This comparison is done on the basis of four parameters that are capability of detecting selfish nodes, punishment mechanism for nodes which do not participate in packet forwarding, network congestion due to the traffic generated by these protocols and cryptographic protection for providing security services like confidentiality of destination address, origin authentication and non-repudiation. This comparison doesn't suggest which protocol is best but it can be used as a reference to decide which protocol should be implemented based on the requirements.

References

1. Yang, Y., Guo, S.-Y., Qiu, X.-S., Meng, L.-M.: Service negotiation model for selfish nodes in the mobile ad hoc networks. In: IEEE ICC Proceedings, pp. 1–5, 5-9 June 2011
2. Song, C., Zhang, Q.: Protocols for stimulating packet forwarding in wireless ad hoc networks, 50–55 (2010)

3. Buchegger, S., Le Boudec, J.-Y.: Self-policing mobile ad hoc networks by reputation systems. IEEE Commun. Mag., 101–107, July 2011
4. Han, Z., Poor, H.V.: Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks. IEEE Trans. Commun. (1), 203–213, January 2009
5. Kim, H.J., Peha, J.M.: Detecting Selfish Behavior in a Cooperative Commons (2008)
6. Varnagar, N., Lathigara, A.: Review paper of selfish node detection in MANET. International Journal of Advance Research in Computer Sciences and management, February 2015
7. Conti, M., Gregori, E., Maselli, G.: Improving the perform-ability of data transfer in mobile ad hoc networks. In: The 2nd IEEE International Conference (2005)
8. Marti, S., Giuli, T., Lai, K., Baker, M.: Mitigating routing misbehavior in mobile ad hoc networks. In: ACM Conference on Mobile Computing and Networking (2000)
9. Kargl, F., Klenk, A., Weber, M., Schlott, S.: Advanced detection of selfish or malicious nodes in ad hoc networks. In: 1st European Workshop on Security in Ad-hoc and Sensor Networks, ESAS 2004 (2004)
10. Djenouri, D., Badache, N.: A novel approach for selfish nodes detection in MANETs: proposal and petri nets based modeling. In: The 8th IEEE International Conference on Telecommunications, ConTel 2005 (2005)
11. Nagar, S.J., Raimagia, D.G., Ghosh, P.A.: Identification and elimination of selfish nodes in ad hoc network. International Journal of Engineering Research and Development, 29–34, April 2014
12. Awerbuch, et al.: An on-demand secure routing protocol resilient to byzantine failures. In: ACM Workshop on Wireless Security (WiSe), Atlanta, Georgia, 28 September 2002
13. Yang, H., Meng, X., Lu, S.: Self-organized network layer security in mobile ad hoc networks (2002)
14. Michiardi, P., Molva, R.: Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. In: Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security: Advanced Communications and Multimedia Security, pp. 107–121 (2002)
15. Buchegger, S., Le-Boudec, J.: Performance analysis of the CONFIDANT, protocol cooperation of nodes fairness in dynamic ad hoc networks. In: Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing, pp 226–236
16. Miranda, H., Rodrigues, L.: Friends and foes: preventing selfishness in open mobile ad hoc networks. In: The 23rd IEEE International Conference on Distributed Computing Systems (ICDCS 2003), pp. 440–445 (2003)
17. Papadimitratos, P., Haas, Z.J.: Secure data transmission in mobile ad hoc networks. In: ACM MOBICOM Wireless Security Workshop (WiSe 2003), San Diego, California, 19 September 2003
18. Buttyan, L., Hubaux, J.: Nuglets: a virtual currency to stimulate cooperation in self-organized mobile ad hoc networks. Swiss Federal Institute of Technology, Lausanne, Switzerland, Tech. Rep. DSC/2001/001, January 2001
19. Liang, G., Agarwal, R., Vaidya, N.: When watchdog meets coding. In: IEEE INFOCOM Proceedings (2010)
20. Fahad, T., Djenouri, D., Askwith, R.: On detecting packets droppers in MANET: a novel low cost approach. In: Information Assurance and Security, IAS 2007 (2007)

Speed Records for Multi-prime RSA Using AVX2 Architectures

Shay Gueron and Vlad Krasnov

Abstract RSA is a popular public key algorithm. Its private key operation is modular exponentiation with a composite *2k*-bit modulus that is the product of two *k*-bit primes. Computing *2k*-bit modular exponentiation can be sped up four fold with the Chinese Remainder Theorem (CRT), requiring two *k*-bit modular exponentiations (plus recombination). Multi-prime RSA is the generalization to the case where the modulus is a product of $r \geq 3$ primes of (roughly) equal bit-length, *2k/r*. Here, CRT trades *2k*-bit modular exponentiation with *r* modular exponentiations, with *2k/r*-bit moduli (plus recombination). This paper discusses multi-prime RSA with key lengths (=*2k*) of 2048/3072/4096 bits, and $r = 3$ or $r = 4$ primes. With these parameters, the security of multi-prime RSA is comparable to that of classical RSA. We show how to optimize multi-prime RSA on modern processors, by parallelizing *r* modular exponentiations and leveraging "vector" instructions, achieving performance gains of up to *5.07x*.

Keywords RSA · Multi-prime RSA · AVX2 · AVX512 · Haswell broadwell skylake

1 Introduction

RSA [1] is a popular public key cryptosystem. It is frequently used in client-server communication protocols such as TLS [2], for encryption and for digital signature.

S. Gueron(✉)
Department of Mathematics, University of Haifa, Haifa, Israel
e-mail: shay@math.haifa.ac.il

S. Gueron
Intel Corporation, Israel Development Center, Haifa, Israel

V. Krasnov
CloudFlare Inc., London, UK
e-mail: vlad@cloudflare.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_22

In the classical TLS handshake, the client uses RSA public key (cheap) operations to encrypt a secret, and the server decrypts it using (expensive) private key operations. The client and the server use this secret to derive a session key. Trust in the server's public key is established through a (trustworthy) certificate that the server submits to the client. The computational overhead from the server's viewpoint is the RSA decryption.

This protocol has the following undesired property: if the server's private key is compromised, then all past recorded session keys are compromised as well, i.e., privacy is lost for the contents of all past sessions. To address this problem, modern TLS handshakes seek "Perfect Forward Secrecy" by using an ephemeral key exchange, based on the ephemeral Diffie-Hellman (EDH) algorithm. The EDH key exchange can be based on $\boldsymbol{F}_p$ arithmetic or on Elliptic Curve computations. Here, the server uses its private key to sign the public Diffie-Hellman (DH) parameters. The signature scheme can be an RSA signature. In such case, the computational overhead from the server's viewpoint is the RSA signature.

We point out that recent industry trends shift towards preferring Elliptic Curves signatures (ECDSA) over RSA signatures. Nevertheless, RSA still enjoys significant popularity, mainly because: a) many existing certificates attest the authenticity of an RSA public key; b) some Certificate Authorities issue only RSA certificates; c) many client applications still request an RSA based handshake, so the server is required to support RSA even if its preferences favor an ECC based handshake.

The security of RSA is typically estimated by the complexity of the best known algorithm for integer factorization. With this approach, *2048*-bit RSA keys are considered to have the equivalent strength of *112*-bit symmetric keys, and *3072*-bit RSA keys have the equivalent strength of *128*-bit symmetric keys. Currently, the vast majority of RSA usages deploy *2048*-bit keys (RSA2048 hereafter). However, the use of longer keys is going to be relevant in the relative near future. NIST's [3] key management guidelines approve RSA2048 until 2030, and recommend migration to *128*-bit security afterwards. This implies migration to at least RSA3072. In addition, NSA guidance is to use at least RSA3072 for Top Secret information [4].

From the viewpoint of servers, that constantly carry out private key operations to support multiple sessions, the computational cost of RSA based handshakes is significant, and will increase if the keys become longer. This motivates software optimization for RSA with the relevant key lengths. In this paper, we show how to optimize multi-prime RSA with $r = 3$ and $r = 4$ primes, over the relevant key lengths of *2048*, *3072* and *4096* bits, to achieve significant performance gains.

2 The Computational Cost of Classical and Multi-prime RSA

The classical RSA scheme uses a $2k$-bit modulus (n) which is a product of two secret k-bit primes (p, q). The modulus $n = p \cdot q$ is made public (called also the "public key"). The algorithm also chooses a public exponent (e) typically

$e = 2^{16}+1$. The private key d is defined by $d = e^{-1} \bmod (p-1)(q-1)$, and its bit-length is also $2k$. The RSA signature on a message M is $s = m^d \bmod n$, where m is a formatted block that consists of a hash digest of M (typically SHA256), appended with proper padding to the full length of $2k$ bits. RSA decryption uses the same type of modular exponentiation (with a differently parsed m).

The most common performance optimization for RSA uses the Chinese Remainder Theorem (CRT). This converts the computation of $s = m^d \bmod n$, to two independent computations $s_1 = m^{d1} \bmod p$, and $s_2 = m^{d2} \bmod q$. Here, $d_1 = d \bmod (p-1)$ and $d_2 = d \bmod (q-1)$. The CRT recombination of s_1 and s_2 yields s. The relative weight of this recombination is small, compared to the computations of s_1 and s_2, so the cost of a $2k$-bit RSA with CRT is closely approximated by twice the cost of computing a k-bit modular exponent. This is approximately 4 times faster than the direct computation of one $2k$-bit modular exponent.

The CRT can be used in a similar way for multi-prime RSA, when the modulus is a product of $r \geq 3$ primes $p_1, p_2, \ldots, p_r$ of (roughly) the same bit length (either $floor(2k/r)$ or $floor(2k/r)+1$). In this case, the r-prime $2k$-bit RSA parameters satisfy $len(p_1) + len(p_2) + \ldots + len(p_r) = 2k$. The CRT computations translate to r modular exponentiations of the type $s_j = m^{dj} \bmod p_j$, where $d_j = d \bmod (p_j-1)$, $j=1, \ldots, r$.

The efficiency of multi-prime RSA stems from the fact that it is faster to compute r modular exponentiations, each with a $(2k/r)$-bit modulus, when r is increased.

On the other hand, the choice of r is limited by security considerations: the complexity of factorizing an integer, known to have small prime factors, decreases as a function of the the smallest factor. Therefore, multi-prime RSA has different security characteristics compared to the classical (2 primes) RSA. Obviously, we are interested in a favorable balance between the security and the performance.

3 Wider Vectorization

Although breaking RSA is not known to be equivalent to factoring the public key, it is common to approximate the security of RSA by the difficulty of the best known integer factorization algorithm. The Number Field Sieve (NFS [5]) is the fastest known factorization technique for a general composite integer of a given size. Its complexity for factoring the integer n scales according to [6]

$$L[n] = e^{1.923 \cdot \sqrt[3]{\log(n) \cdot \left(\log\log(n)\right)^2}} \tag{1}$$

(note that the expression does not depend on the number of factors of n).

If it is known that the prime factors of n are small, then the Elliptic Curve Method (ECM) [7] becomes faster than the NFS. Here, the complexity decreases as a function of the size of the smallest factor (p) of n, and scales according to

$$E[n,p] = \log_2(n)^2 e^{\sqrt{2\log(p)\cdot\log\log(p)}} \quad (2)$$

A recent large integer, known to be factored using the NFS algorithm has *786* bits and this required ~2^{67} operations (see [8]). We use 2^{67} for $L[2^{768}]$ as a baseline, in order to obtain a "practical" approximation of the security of RSA with a given modulus sizes (A. Langley in [9] uses 2^{80} bits for $L[2^{1024}]$, as baseline for extrapolation).

Consider a modulus n whose smallest prime factor is p. We estimate the number of operations required to factor n by the minimum of $2^{67}\times L[n]/L[2^{768}]$ and $2^{67}\times E[n, p]/L[2^{768}]$. We used these estimations as an indicator of the symmetric key equivalent strength of RSA and multi-prime RSA. The approximations are shown in Figure 1, and lead to the following conclusions:

4-prime RSA4096 and *4*-prime RSA3072 have at least the same security as the (classical) RSA4096/RSA3072.

3-prime RSA2048 has the same security as RSA2048.

4-prime RSA2048 gives an only *slightly* reduced (by a single bit) security level compared to RSA2048.

Consequently, multi-prime RSA is obviously advantageous in the first two cases, and is a performance-security tradeoff in the third case.

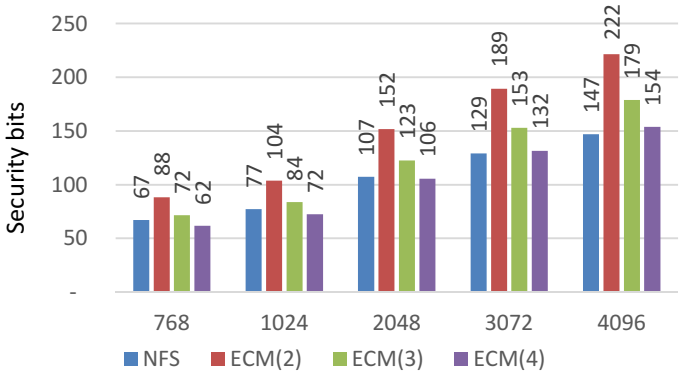


Fig. 1 The "bit-equivalent" complexity of factoring an integer with *768* bits (as the extrapolation base) and *1024/2048/3072/4096* bits, using the NFS and the ECM methods. The NFS method is applied to a general integer. The ECM (r) method is applied under the assumption that is has r=2/3/4 prime factors of the same size. The different bars show the results with a different number of prime factors. See explanations in the text.

4 Simultaneous Modular Exponentiation with Vector Instructions

The Intel© Architecture Codename Haswell is the first processor that introduced AVX2 [10] instructions with integer operations that are relevant to our study.

The architecture has *256*-bit registers (called "ymm"s), that can hold four *64*-bit values (SIMD "elements"), and these instructions can perform four *64*-bit integer operations in parallel, over these registers. For example, the instructions VPADDQ and VPMULUDQ allow for computing four *64*-bit additions or four *64*-bit low half products, with one instruction invocation. For RSA computations (i.e., modular arithmetic with big integers), we can leverage this capability if we move the inputs (outputs) from (to) the standard radix 2^{64} to a "redundant representation", as described in detail in [11]. Optimization of the multi-prime RSA can be achieved by leveraging the redundant representation technique to compute (up to) *4* modular exponentiations simultaneously. To compute the modular operations (Montgomery multiplications in our case) without cumbersome handling of carry bits, we view the operands as multi-digit numbers with *29*-bit "digits". Each *29*-bit digit populates a *64*-bit container in the relevant ymm register. We use the available AVX2 instructions and perform *64*-bit multiply-accumulate operations on these "smaller digits", and can accumulate several products before the accumulator overflows (to see how this was done for accelerating the serial one modular-exponentiations of RSA, see [11]).

We point out that accelerating the classical RSA with AVX2 requires some mix of ALU and SIMD instructions. By contrast, the multi-prime RSA offers an inherent opportunity to parallelize (up to) four independent Montgomery multiplications, for (up to) four independent modular exponentiations, without mixing ALU instructions. This allows for optimizing the software flow in a way that the processor maximizes the number of instructions it executes per cycle. Our optimized AVX2 code computes all the required modular exponentiations in a single function, and defers the CRT recombination to the end of these computations.

Note that the AVX2 parallelization of modular exponentiation can be applied to the classical *2*-prime RSA as well. However, the SIMD unit is under-utilized in this case: only two out of four possible operations are parallelized. Here, we found that the AVX2 implementation is slower than the straightforward approach, and therefore discard it.

5 Results

To evaluate our algorithm, we implemented an optimized parallel modular exponentiation code that uses the above method. The code was integrated into the OpenSSL [12] (git version from August 19, 2015), and contributed to this library as a software patch. It is available from [13], and we assume that it would be integrated to one of the future official OpenSSL releases.

We prepared new parallel modular exponentiation functions for *3*-prime RSA2048, *4*-prime RSA2048 and *4*-prime RSA3072. While we did not prepare a dedicated parallel modular exponentiation for *4*-prime RSA4096 and *3*-prime RSA3072, they implicitly benefit from the serial *1024*-bit modular exponentiation, which was already highly optimized by us [14], and is part of the current official OpenSSL version (as of 1.0.2).

For benchmarking, we used the built-in "openssl speed" utility. Since the three tested processors run at different frequencies, we converted the results that are

reported in signatures-per-second, to the frequency-agnostic metric "Cycles/Signature". The results are summarized in Figures 2, 3 and 4.

The figures demonstrate a significant speed advantage for *4*-prime on all moduli sizes. The greatest gain is observed on Haswell with Hyperthreading enabled, reaching *5.07x* speedup compared to the classical *2*-prime RSA. More importantly the *4*-prime RSA3072 is now on par with *2*-prime RSA2048 and even outperforms it on Skylake. That enables to default to larger moduli, gaining the security equivalent of *256*-bit EC without sacrificing performance.

Even RSA4096 becomes feasible with *4*-prime, being only *2*x slower than *2*-prime RSA2048 (instead *10x* slower on Haswell and almost *7x* slower on Broadwell and Skylake before).

The performance improvement stems from two sources: a) the multi-prime implementation is inheritably faster due to the use of modular exponentiations with smaller operands; b) our parallelized method significantly optimizes the implementation.

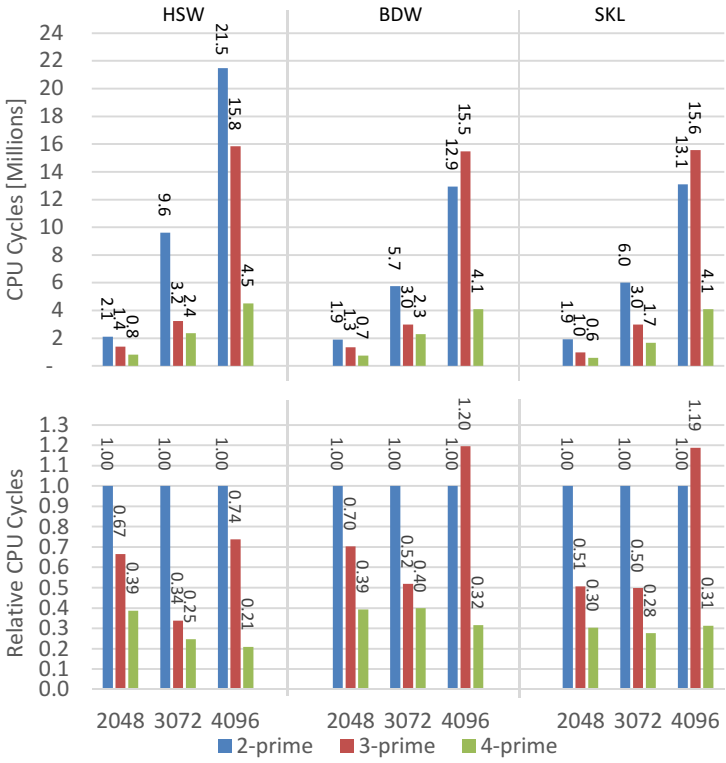


Fig. 2 The performance of our multi-prime RSA implementations with *r=3/4*, compared to the classical *2*-prime RSA. Top panel: CPU cycles per signature (lower is better). Bottom panel: the relative performance of multi-prime RSA versus the *2*-prime version (scaled as *1*). The performance is shown for three Intel processors left to right): Architecture Codename Haswell (HSW), Broadwell (BDW), Skylake (SKL). Hyperthreading and turbo are disabled.

6 Conclusion

Multi-prime RSA is a great way to alleviate the computational overhead that is experienced by servers that constantly open multiple client-server sessions with clients that do not allow (i.e., request) ECDSA signatures during handshake negotiation. With the right parameters choice, the security level of multi-prime RSA is the same as that of the classical RSA. From the client viewpoint, there is no difference between RSA and multi-prime RSA: the (client observable) value of the server's public modulus does not reveal the number of its prime factors. Therefore, a server can choose unilaterally (and silently) to use the multi-prime RSA (with the proper certificate, of course), and enjoy the resulting performance benefits.

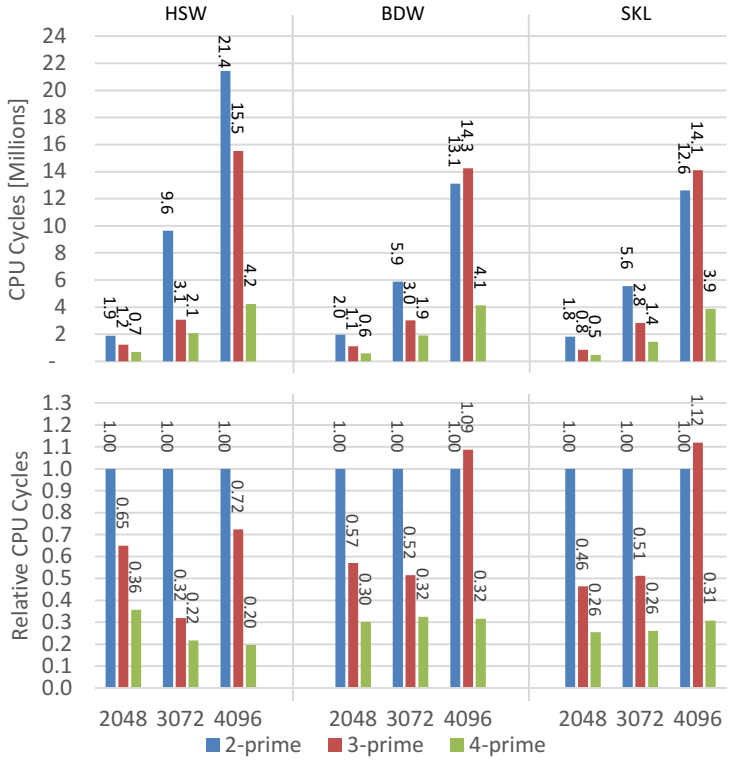


Fig. 3 The performance of our multi-prime RSA implementations with *r=3/4*, compared to the classical (*2*-prime) RSA. Top panel: CPU cycles per signature (lower is better). Bottom panel: the relative performance of multi-prime RSA versus (*2*-prime) RSA, which is scaled as *1*). The performance is shown for three Intel processors left to right): Architecture Codename Haswell (HSW), Broadwell (BDW), Skylake (SKL). Hyperthreading is enabled. Turbo is disabled.

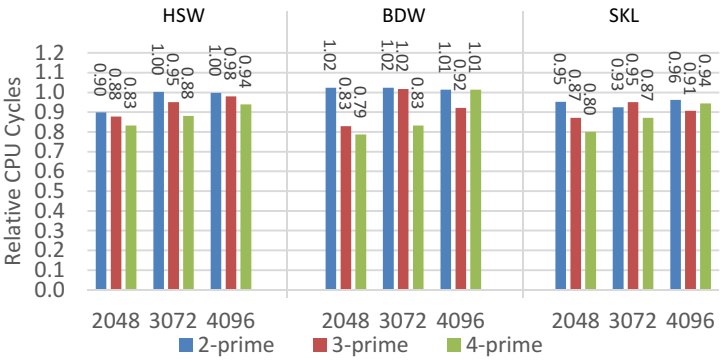


Fig. 4 The *2*-prime and multi-prime RSA performance with Hyperthreading enabled, relative to the single-threaded performance. A value under 1.0 indicates faster relative performance. The performance is shown for three Intel processors left to right): Architecture Codename Haswell (HSW), Broadwell (BDW), Skylake (SKL). Turbo is disabled

The multi-prime RSA method benefits twice on the modern architectures: a) it is inherently faster than the classical RSA; b) it allows to fully utilize the wide SIMD architecture (offered by AVX2), to gain more performance.

To code that implements the optimized parallel modular exponentiation code using AVX2, is contributed to OpenSSL as a software patch. Until it is integrated into a future official OpenSSL releases, it is available from [13], and can be used by user who build an updated OpenSSL version.

6.1 Future Architectures

The upcoming AVX512 [15] instructions set will introduce a larger register file (*32×512* bits). This will allow for implementing the entire multiplication (in redundant form) with no spills to memory, and potentially improving the performance further. In the future, the VPMADD52[L/H] [15] instructions are expected to improve the performance of RSA even further, including the parallel multi-prime implementation here, by providing a much lower instruction count.

Both the AVX512 in VPMADD52 implementations of *1024* and *2048*-bit modular exponentiations are already available as contributed patches for OpenSSL [16], [17], and are expected to further improve the performance of 3-prime RSA3072 and 4-prime RSA4096. They can also be used as a reference for implementing modular exponentiation (single or parallel) with other key lengths. For example, a VPMADD52-based implementation of *1536*-bit modular exponentiation, that speeds up the classical RSA3072, has been recently posted in [18].

Acknowledgements This research was supported by the PQCRYPTO project, which was partially funded by the European Commission Horizon 2020 research Programme, grant #645622.

References

1. Jonsson, J., Kaliski, B.: Public-Key Cryptography Standards (PKCS) #1: RSA Cryptography Specifications Version 2.1. In RFC 3447, Network Working Group, IETF (2003). https://www.ietf.org/rfc/rfc3447.txt
2. Dierks, T., Rescorla, E.: The Transport Layer Security (TLS) Protocol Version 1.2. In RFC5246, Network Working Group, IETF (2008). https://www.ietf.org/rfc/rfc5246.txt
3. Barker, E., Roginsky, A.: Transitions: Recommendation for Transitioning the Use of Cryptographic Algorithms and Key Lengths. In NIST Special Publication 800-131A, p. 5 (2011). http://csrc.nist.gov/publications/nistpubs/800-131A/sp800-131A.pdf
4. NSA: Cryptography Today (Accessed September 2015). https://www.nsa.gov/ia/programs/suiteb_cryptography/index.shtml
5. Lenstra, A.K., Lenstra Jr., H.W. (eds.): The Development of the Number Field Sieve, vol. 1554. Lecture Notes in Mathematics. Springer, Berlin (1993)
6. Lenstra, A.K.: Unbelievable security matching AES security using public key systems. In: Advances in Cryptology, ASIACRYPT 2001, pp. 67–86. Springer, Heidelberg (2001)
7. Lesntra Jr., H.W.: Factoring integers with elliptic curves. The Annals of Mathematics **126**(3), 649–673 (1987)
8. Thorsten, K., et al.: Factorization of a 768-Bit RSA modulus. In: Proceedings of the 30th Annual Cryptology Conference on Advances in Cryptology, CRYPTO 2010, Santa Barbara, CA, USA, 15-19 August 2010, pp. 333–350 (2010)
9. Langley, A.G.: Multi-prime RSA trade offs. In ImperialViolet (blog) (2011). https://www.imperialviolet.org/2011/04/09/multiprime.html
10. Buxton, M.: Haswell New Instruction Descriptions Now Available! Intel Corporation (2011). http://software.intel.com/en-us/blogs/2011/06/13/haswell-new-instruction-descriptions-now-available/
11. Gueron, S., Krasnov, V.: Software implementation of modular exponentiation, using advanced vector instructions architectures. In: Proceedings of the 4th International Conference on Arithmetic of Finite Fields, WAIFI 2012, pp. 119–135 (2012)
12. OpenSSL: The Open Source toolkit for SSL/TLS. http://www.openssl.org/
13. Gueron, S., Krasnov, V.: OpenSSL multi-prime patch. https://github.com/vkrasnov/multiprime
14. Gueron, S., Krasnov, V.: [PATCH] Efficient, and side channel analysis resistant 1024-bit modular exponentiation, for optimizing RSA2048 on AVX2 capable x86_64 platforms. OpenSSL patch, posted July 2012. https://rt.openssl.org/Ticket/Display.html?id=2850
15. Intel Corportaion: Intel® Architecture Instruction Set Extensions Programming Reference. Intel, August 2015. https://software.intel.com/sites/default/files/managed/07/b7/319433-023.pdf
16. Gueron, S., Krasnov, V.: [PATCH] Efficient 1024-bit and 2048-bit modular exponentiation for AVX512 capable x86_64. OpenSSL patch, posted January 2014. https://rt.openssl.org/Ticket/Display.html?id=3240
17. Gueron, S., Krasnov, V.: [PATCH] Fast modular exponentiation with the new VPMADD52 instructions. OpenSSL patch, posted November 2014. https://rt.openssl.org/Ticket/Display.html?id=3590
18. Gueron, S., Drucker, N.: [PATCH] Fast 1536-bit modular exponentiation with the new VPMADD52 instructions. OpenSSL patch, posted September 2015. https://rt.openssl.org/Ticket/Display.html?id=4032

Privacy Preservation of Source Location Using Phantom Nodes

Shruti Gupta, Prabhat Kumar, J.P. Singh and M.P. Singh

Abstract Sensor networks are widely used for subject monitoring and tracking. In this modern era of wireless technology, privacy has become one of the essential concerns of wireless sensor networks. The location information of sensor nodes has to be hidden from an adversary for the sake of privacy. An adversary may trace the traffic and try to figure out the location of the source node. This can expose a significant amount of information about the subject being monitored which may be further misused by the adversary. This paper proposes 2-Phantom Angle-based Routing Scheme (2PARS) designed to improve the Source Location Privacy. The proposed scheme considers a triplet for selecting the phantom nodes. A triplet is a group of three nodes formed on the basis of their distance from the sink, their location information and the inclination angle between them. For every single packet transmission, phantom selection is performed thereby creating alternative paths. As the path changes dynamically, the safety period increases without any significant increase in the packet latency. The analysis shows that this scheme performs better in terms of safety period as compared to single phantom routing and multi-phantom routing schemes.

Keywords WSN privacy · Context privacy · Random walk · Dynamic routing · Safety period

1 Introduction

Wireless Sensor Network (WSN) consists of small, low cost, limited power sensor nodes that can sense, collect and circulate information [1, 2]. One of the major application of sensor-based networks is subject monitoring and tracking [3].

S. Gupta · P. Kumar(✉) · J.P. Singh · M.P. Singh
Department of Computer Science and Engineering, National Institute of Technology Patna, Patna, Bihar, India
e-mail: gupta.shruti78@gmail.com, {prabhat,jps,mps}@nitp.ac.in

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_23

Here, the nodes are randomly deployed to inspect a subject which can be any object of interest. Sensing of the subject by a sensor node is known as an event [4, 5]. Nowadays, privacy has become an important consideration. Privacy means "a state of being anonymous, unobservable and unlinkable". It can also be defined as "the right to be left alone" [6]. Privacy in WSN can be classified into two categories: Content privacy and Context privacy [7]. Content privacy basically focuses on providing non-repudiation, integrity, freshness and confidentiality of the messages. Context privacy deals with hiding the location and identity of nodes and the traffic flow within the network. The contextual information such as the location of the monitored assets can be misused by an adversary.

Hiding the location of the source is a difficult task as there are numerous aspects that influence the effectiveness of the solution [8] such as the mobility of the nodes, resources of the adversary, adversary capability of viewing the network traffic etc.

In this paper, we propose 2PARS which has been designed to improve the SLP. This scheme uses the distance of each node from the BS, its location information and the inclination angle between the nodes to form a triplet which is used to select the phantom nodes. Our analysis shows that the safety period of the proposed algorithm is better than Phantom Single Path Routing Scheme(PSRS) and Multi-Phantom Routing Scheme(MPRS).

The remaining part of the paper is organized as follows. Section 2 discusses Related Work, Section 3 describes the Network model, Section 4 discusses the Adversary Capabilities, Section 5 explains the Proposed protocol, Section 6 provides an analysis of the protocol and Section 7 concludes with a brief discussion of Future Work.

2 Related Work

Kamat et al. [9, 10] introduced the phantom routing scheme(PRS), the first strategy based on a random walk to provide SLP. It involves two phases: the random walk phase and the flooding phase. In random walk phase, when a source senses an event, the message is forwarded in a randomized manner upto h hops. The node at the end of the random walk is treated as a phantom source. After h hops, the packet is flooded towards the BS using baseline flooding. Another protocol named PSRS works similar to PRS. However in PSRS, instead of using baseline flooding, it forwards the message to the BS using shortest path algorithm. Kamat et al. in [9] illustrated that pure random walk is incompetent in keeping phantom source away from the real source as each node has an equal probability of being selected as an intermediate node. This may led to message looping in a cycle near source node. In order to avoid the repetition of paths, directed random walk [11] was introduced which can be either sector-based or hop-based. In sector-based directed random walk, each node divides its neighbors into two sets: north-west and south-east. When the source senses an event, it randomly chooses one of the set and forwards the message to a randomly selected node of that set. Each intermediate node forwards the message to randomly

selected neighbor from its defined set. In hop-based directed random walk, each node divides its neighbors on the basis of their hop-count distance from the BS. It has two sets: neighbors having larger hop-count and neighbors having equal or smaller hop-count. The same procedure for message forwarding is followed as in sector-based directed random walk. L. Zhang et al. [12] introduced Self-Adjusting Directed Random Walk(SADRW) which is an enhancement of sector-based directed random walk. In SADRW, neighbors are divided into four sets: north, south, east, and west. When a node senses an event, it randomly chooses a set and forwards the message to the randomly selected node of that set. Each intermediary node forwards the message in the same direction as selected. If any node cannot forward it in the selected direction, then it randomly selects a fresh direction from remaining sets. This random walk terminates if the packet is at the boundary of the network and has already travelled i hops such that $h \geq i$. Otherwise, it selects another direction from remaining sets. Wei-Ping et al. [13] suggested Phantom Routing with Locational Angle(PRLA) which is an enhancement of PSRS. Here, every node calculates the inclination angle between itself and its neighbors with BS as a vertex. It states that higher the inclination angle, higher would be the forward probability. The forward probability of a node is used to select it as an intermediate node. Each intermediate node uses the same inclination angle to forward the packet. The last node that receives the packet after h hops is considered as phantom source. The shortest path algorithm is then used to forward the packet to the BS. Li et al. [14, 15] introduced the RRIN(Randomly selected intermediate node) which is an improvement of PRS. In this scheme, source node forwards the message to the randomly selected intermediate neighbor node and in the same fashion, the message is sent to the BS. This increases the safety period, but at the cost of higher delay and energy consumption. The safety period is the number of messages that are successfully received by the BS before an adversary reaches the source location. Spachos et al. [16] proposed Angle-based Dynamic Routing Scheme(ADRS) to enhance the SLP. This scheme calculates the inclination angle between the sender and the receiver. It creates a candidate set of nodes based on the inclination angle for forwarding the packets. For every message transmission, the candidate set is reformed which creates multiple paths to the destination. Manjula et al. [17] proposed Energy-efficient Privacy Preserved Routing (EPR). In this scheme, the author uses 2α-angle anonymity concept to generate the location of phantom sources. Kumar et al. [18] introduced a new protocol using two phantom nodes named MPRS. It includes two phases: Configuration phase and Working phase. In the configuration phase, triplets of nodes are formed by the BS. A triplet is a group of three nodes where the inclination angle between every two nodes should be at least 30^o. When the source senses an event in the working phase, it selects a node from its triplet (work as a phantom source) and forwards the message to its neighbor with phantom source as a destination. The phantom source forwards the message to the BS using shortest path algorithm. The selection of the phantom source is based on a randomly generated number.

3 Network Model

In our network model, identical sensor nodes are randomly deployed which can perform three different roles i.e. act as a Source node, Phantom Source node, and the Intermediary nodes. The sensor nodes are static after deployment. Source nodes are those nodes which sense any subject and want to forward the messages to the BS. A subject is a certain kind of entity of interest. The nodes which are neither source node nor phantom node for a message behaves as intermediary nodes and are used to transfer message to next hop. They are not allowed to make any modification in the message. Phantom Source nodes are those nodes which forward the received message from the source node to the destination (BS) using their own specific identity. These nodes behave as a fake source on behalf of the original source for the adversary who eavesdrops the traffic and tries to find out the location of the original source. All the nodes have similar characteristics but perform different works at different times.

4 Adversary Capabilities

In this section, the capabilities of an adversary considered in our proposed scheme has been discussed. We classify the capabilities of an adversary on the basis of its behavior, the vision of the network, the resources it has and the information it has about the network. An adversary tries its best to locate the source node. The characteristics of an adversary have been summarized below:

- An adversary is external and has enough resources.
- An adversary can move from one location to another while tracing the traffic flow.
- An adversary can locate the specific node by eavesdrop-ping and hop-by-hop tracing. Initially, it starts from the BS with its strategy to trace back to the original source location.
- An adversary is passive. It can only eavesdrop the traffic flow and attempt to scrutinize the contents inter-changed among the nodes.
- An adversary is semi-honest. It obeys the protocols followed within the network so that it remains unobserved.
- An adversary has the ability to locally visualize the traffic inside the network. It can only view a part of the network at a glance. It can only overhear the flow of messages within its hearing range.

5 2PARS

We propose a new 2-Phantom Angle-based Routing Scheme for providing source location privacy. Our proposed scheme consists of two phases: Set-up phase and Steady phase. The proposed work assumes that BS has locational information of all nodes deployed.

5.1 Set-up Phase

In the beginning of this phase, BS floods an empty packet to its neighbors with a counter value set to zero. Each node that receives the message stores the sender ID and the counter value. The counter value is used to determine the hop-count of the sender node from the BS. The message is then flooded within the network with the counter value incremented by one. In this way, nodes can determine their hop-count from the BS. Each node reports its hop-count to the BS. The BS maintains a table of the hop-count of each node. This table is utilized for creating the triplets within the network. In a triplet, each node behaves as a phantom for other two nodes. The selection of triplet should be done in such a way that no two nodes are co-linear with the BS and the inclination angle measured between any two nodes with the BS as a vertex should be at least 30^o as shown in Fig. 1. If the nodes from a triplet are co-linear with the BS then the source node may lie in the path between the phantom node and the BS. When this condition arises, the privacy can be easily breached by the adversary.

$$t(slope) = \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{1}$$

$$\phi = \tan^{-1} \frac{(t_2 - t_1)}{(1 + t_1 t_2)} \tag{2}$$

The BS randomly selects three nodes m_1, m_2 and m_3 and calculates the inclination angles ϕ_1 and ϕ_2 between them using eqn. 1 and 2 as shown in Fig. 1. We have fixed the position of the node m_2 in the triplet. m_1 and m_3 may be replaced by same hop-count distant node in case the requirement of minimum 30^o angle between nodes of triplet is not satisfied as discussed in algorithm 2. Calculation of two angles ϕ_1 and ϕ_2 serves our purpose as we are considering only 2D scenarios in this work. In the 2D scenario, if the angle (ϕ_1) between m_1 and m_2 and the angle (ϕ_2) between m_2 and m_3 is greater than 30^o then obviously the angle between m_1 and m_3 would be greater than 30^o.

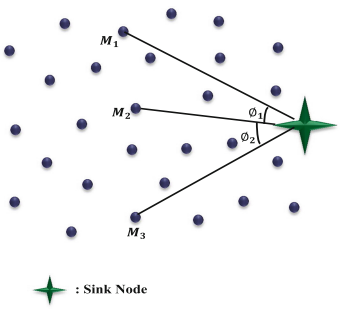


Fig. 1 Triplet Selection

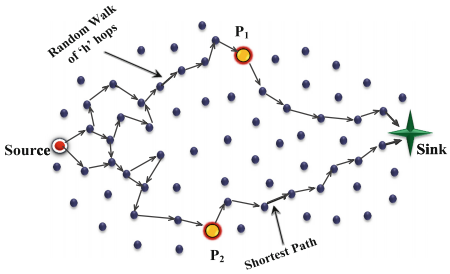


Fig. 2 Distance of source from BS $> n$

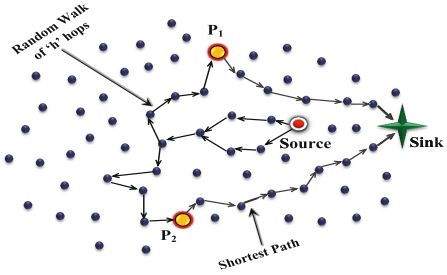


Fig. 3 Distance of source from BS $\leq n$

Normally, the phantom selection and transmission of message is done as shown in Fig. 2. However for scenario when the source node would be closer to the BS needs special care as adversary is assumed to be near the BS initially and it would be relatively easy for it to trace back to the source node. For handling such cases, the position of the phantom node should be in the opposite direction of source node as shown in Fig. 3. Algorithm 1 takes care of such cases and ensures that it is still difficult for the adversary to trace back to the source location. Algorithm 1 uses two user-defined values n and p where n is the minimum hop distance of source node from the BS and p is the minimum hop distance of the phantom node from the BS.

Algorithm 1. Triplet Selection Algorithm (At BS)

1: Arrange the table in ascending order of hop-count value.
2: Initialize n and p, $p > n$.
3: Randomly selects three nodes m_1, m_2, m_3.
4: Select the node with lowest hop-count, called it as m_1.
5: **if**(Hop-count(m_1) $\leq n$)
6: **if**(Hop-count(m_2) $\geq p$)
7: **if**(Hop-count(m_3) $\geq p$)
8: Call Algorithm 2
9: **goto** 3
10: **else**
11: replace(m_3 with other node)

12: **goto** 4
13: **else**
14: replace(m_2 with other node)
15: **goto** 4
16: **else**
17: Call Algorithm 2
18: **goto** 3

The BS selects the node having lowest hop-count among the selected nodes m_1, m_2 and m_3 and treat it as a source node. If the hop-count of the selected node is less than n then it compares the hop-count of the other two nodes with p which need to work as phantom nodes. If the condition of phantom nodes are satisfied, the triplet is selected. Otherwise it is replaced with other nodes until the condition is fulfilled. The entire procedure works as per algorithm 1.

Algorithm 2. Angle Calculation Algorithm (At BS)

1: Arrange the table in ascending order of hop-count value.
2: Calculate angle ϕ_1(m_1 and m_2) and ϕ_2(m_2 and m_3) in degree.
3: **if** ($\phi_1 \geq 30^o$)
4: **if**($\phi_2 \geq 30^o$)
5: Triplet Selected
6: Inform(m_1, m_2, m_3)
7: **Exit**
8: **else**
9: replace(m_3 with other node of nearly same hop-count)
10: **goto** 4
11: **else**
12: replace(m_1 with other node of nearly same hop-count)
13: **goto** 3

After triplet selection, the BS informs all the nodes about its triplet by sending a packet containing the ID of other two nodes. Each node records the ID of other two nodes which will act as a phantom node for it.

5.2 Steady Phase

Steady phase described in algorithm 3 starts after completion of set-up phase. When the source senses a subject within its range, it needs to forward the message to the BS. Let q be the number of neighbors for the source node S. For initial q messages, the source uses different neighbors to route the messages towards the BS. After that, the source generates two random numbers r_1 and r_2 in the range of 1 to 10. r_1 is generated for phantom selection and r_2 works as a threshold value for that selection.

These random numbers are generated for every message transmission and are used to select one among two available phantoms. If r_1 is greater than r_2 then the first node will behave as a phantom otherwise the second one will behave as a phantom. Once the phantom has been selected, S randomly selects one of its neighbors. The message X is forwarded to the selected neighbor m with P as the destination. Selected neighbor forwards the message to the destination phantom node using random walk. The phantom node checks for source ID of all the messages received by it to locate the messages coming from one of its phantoms. The phantom replaces the source ID of only these messages with its own ID and send to the BS using shortest path for the sake of privacy. For other messages it acts only as a intermediary node. The entire procedure is described in algorithm 3.

Algorithm 3. Steady Phase Algorithm

1: Source S directly sends q messages
2: S generates 2 random numbers r_1 and r_2 between [1,10]
3: **if** $(r_1 \leq r_2)$
4: $P = P_1$
5: **else**
6: $P = P_2$
7: S randomly selects node $m \in$ M(Set of neighbors)
8: S forwards message X to m with P as destination
9: m forwards X to P
10: **if** (phantom(P) == S)
11: replace($S_{ID} = P_{ID}$)
12: P forwards X to BS
13: **else**
14: P forwards X to BS

The creation of multiple paths confuses the adversary and make it harder for it to trace down source location. However, BS must be in a position to detect the real source. Hence, the source should also include its ID in the message content as it is replaced by phantom node for the sake of privacy.

6 Analysis

Attaining the source location privacy with increased safety period is a challenging task. In shortest path based algorithms, the local adversary, in worst case, only needs to capture h messages to breach the privacy of a source h hop distance away. The proposed algorithm is successfull in increasing the hit ratio as compared to earlier schemes. This works is also successfull in overcoming the problems of MPRS related to scenarios where the source was closer to BS. In our proposed scheme, we try to overcome these problems and improve the algorithms to provide better SLP. In the

set-up phase, triplet selection is performed and the different scenarios based on the position of the source with respect to the BS are considered. Nodes must comply with certain threshold values of the hop-count to be a part of a triplet. As discussed only two angles are needed to be calculated in the present work. In MPRS, the angle between every pair is calculated which is an extra overhead. Algorithm 2 calculates the triplet such that proper phantoms are selected based on inclination angle. In the steady phase, the source sends first q messages directly through q neighbors as the likelihood of detecting the source by only one message is negligible. With the use of two random numbers, we are providing more randomness to phantom selection. Due to this, the paths changes more dynamically which significantly increases the safety period as compared with MPRS.

We have considered hit ratio to analyze our work. Hit ratio depends on number of messages captured out of total messages sent and can be calculated using eqn. 3.

$$Hit\ Ratio(HR) = \frac{Number\ of\ messages\ captured}{Number\ of\ messages\ sent} \tag{3}$$

Lets assume that the source S having M neighbors senses an event , then the probability of selecting a neighbor would be $1/M$. If it has P phantoms then the probability of selecting a phantom would be $1/P$. Thus, a total of MP different paths would exist. If we assume that an adversary can move one hop towards the source after capturing a single message, then it only needs $(d + p)$ messages to reach the source when every message follows the same path where d and p are the hop-count from the source to the phantom and from the phantom to the BS respectively. The probability that the message is transmitted from a certain path through phantom would be $1/MP$. Hit ratio can be calculated for this work using eqn. 4. However, we have fixed the number of phantoms to two. Putting the value of P as two in eqn. 4 gives HR equal to 1/3M. This shows that hit ratio is inversely propositional to the number of neighbors. Hence, a denser deployment of sensor nodes will further decreases the hit ratio which increases the privacy.

$$HR = \frac{1}{M + MP} \tag{4}$$

The proposed algorithm is successfull in reducing hit ratio by creating alternative paths. If we more denser the network, the hit ratio would be decreased which increases the privacy.

7 Conclusion

The algorithm discussed in the paper obfuscate the adversary by creating alternative paths. This protocol significantly increases the safety period with some overhead of message latency. For energy efficiency and less network congestion, we have not used

flooding technique in the steady phase. Future work may be done by increasing the number of phantom nodes and evaluating their cost benefits. The proposed protocol may be further enhanced by including other privacy preservation techniques.

References

1. Zhou, Y., Fang, Y., Zhang, Y.: Securing wireless sensor networks: a survey. IEEE Communications Surveys & Tutorials **10**(3), 6–28 (2008)
2. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Computer Networks **38**(4), 393–422 (2002)
3. Alomair, B., Clark, A., Cuellar, J., Poovendran, R.: Statistical framework for source anonymity in sensor networks. In: Global Telecommunications Conference (GLOBECOM 2010), pp. 1–6. IEEE (2010)
4. Kamat, P., Xu, W., Trappe, W., Zhang, Y.: Temporal privacy in wireless sensor networks. In: The 27th International Conference on Distributed Computing Systems, ICDCS 2007, p. 23. IEEE (2007)
5. Sheng, B., Li, Q.: Verifiable privacy-preserving range query in two-tiered sensor networks. In: The 27th Conference on Computer Communications, INFOCOM 2008. IEEE (2008)
6. Solove, D.J.: Understanding privacy, pp. 15–17. Harvard University Press, Cambridge (2008)
7. Li, N., Zhang, N., Das, S.K., Thuraisingham, B.: Privacy preservation in wireless sensor networks: A state-of-the-art survey. Ad Hoc Networks **7**(8), 1501–1514 (2009)
8. Conti, M., Willemsen, J., Crispo, B.: Providing source location privacy in wireless sensor networks: a survey. IEEE Communications Surveys & Tutorials **15**(3), 1238–1280 (2013)
9. Kamat, P., Zhang, Y., Trappe, W., Ozturk, C.: Enhancing source-location privacy in sensor network routing. In: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, ICDCS 2005, pp. 599–608. IEEE (2005)
10. Ozturk, C., Zhang, Y., Trappe, W.: Source-location privacy in energy-constrained sensor network routing. In: Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks, pp. 88–93. ACM (2004)
11. Yao, J., Wen, G.: Preserving source-location privacy in energy-constrained wireless sensor networks. In: 28th International Conference on Distributed Computing Systems Workshops, ICDCS 2008, pp. 412–416. IEEE (2008)
12. Zhang, L.: A self-adjusting directed random walk approach for enhancing source-location privacy in sensor network routing. In: Proceedings of the 2006 International Conference on Wireless Communications and Mobile Computing, pp. 33–38. ACM (2006)
13. Wei-Ping, W., Liang, C., Jian-Xin, W.: A source-location privacy protocol in wsn based on locational angle. In: IEEE International Conference on Communications, ICC 2008, pp. 1630–1634. IEEE (2008)
14. Li, Y., Lightfoot, L., Ren, J.: Routing-based source-location privacy protection in wireless sensor networks. In: IEEE International Conference on Electro/Information Technology, eit 2009, pp. 29–34. IEEE (2009)
15. Li, Y., Ren, J.: Preserving source-location privacy in wireless sensor networks. In: 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON 2009, pp. 1–9. IEEE (2009)
16. Spachos, P., Toumpakaris, D., Hatzinakos, D.: Angle-based dynamic routing scheme for source location privacy in wireless sensor networks. In: 2014 IEEE 79th Vehicular Technology Conference (VTC Spring), pp. 1–5. IEEE (2014)
17. Manjula, R., Datta, R.: An energy-efficient routing technique for privacy preservation of assets monitored with wsn. In: 2014 IEEE Students' Technology Symposium (TechSym), pp. 325–330. IEEE (2014)
18. Kumar, P., Singh, J., Vishnoi, P., Singh, M.: Source location privacy using multiple-phantom nodes in wsn. In: TENCON 2015 - 2015 IEEE Region 10 Conference, pp. 1–6, November 2015

Implementing a Bitcoin Exchange with Enhanced Security

Ebru Celikel Cankaya and Luke Daniel Carr

Abstract We design and implement a secure cryptocurrency exchange by combining secure web programming practices with Bitcoin. In this work, the Amazon EC2 cloud is used to store Bitcoin and host the web server. While building the system, we employ a comprehensive coverage of security features including symmetric key encryption, SSL certificates for digital signatures, SHA-512 for hashing, password revalidation against session hijacking attacks, and two factor authentication. The implementation is tested in an experimental setup and is compared with similar applications currently available. The tests demonstrate that our system proves to offer enhanced security outperforming its peers.

Keywords Amazon web services · Cloud storage · Cryptocurrency exchange

1 Introduction

Thanks to the convenience of E-commerce transactions, the way people buy and sell things is slowly shifting to the digital realm. As online transactions become more prevalent, more practical means to support such transactions are needed. A trusted third party (a financial institution such as a bank, as is the most common) well provides this by undertaking the mission of accountability towards the buyer as well as the seller. Though secure, this solution incurs several overheads: Financial institutions are subject to rules and regulations which are country specific. Also, performing transactions via a financial institution involves extra costs such as credit card fees. Moreover a system that relies on a cryptographic proof rather than trust will be more reliable.

Our implementation offers users a secure method to exchange Bitcoin for fiat, therefore enabling the use of Bitcoins for e-commerce transactions to take

E.C. Cankaya(✉) · L.D. Carr
Department of Computer Science, University of Texas at Dallas, Richardson, Texas, USA
e-mail: {exc067000,lxc112230}@utdallas.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_24

advantage of anonymous transactions with no obligation to include a trusted third party, i.e. the financial agents. Bitcoin is comprised of a peer-to-peer network where a sequence of timestamped transaction is hashed to create a blockchain [Nakamoto, 2008, Androulaki, 2013]. This blockchain propagates through [Babaioff, 2012] participants (miners) to generate a secure log.

2 Motivation and Background

[Krieger, 2013] Using financial institutions as a trusted third party in E-commerce transactions is a common method. Yet, it suffers from the inherent weaknesses of a trust based model such as the lack of ability to support non-reversible payments, and the augmented transaction cost rendering small transactions almost impossible due to its minimum transaction size enforcement.

In Bitcoin, the first transaction that is part of a block initiates a set of new coins (block reward). These coins belong to the creator of the block. Progressively, coins are circulated in the network by users as the Bitcoin system lacks a central authority. The participants of the peer-to-peer Bitcoin network generate and then gradually build a public block-chain, therefore collectively forming a consensus to validate transactions and prevent double spending. The blockchain is an appended history of previously agreed upon transactions. It is built by repeatedly calculating the cryptographic hash of the current transaction, other pending transactions, and a random nonce. There is a significant cost incurred due to the cryptographic computation of the blockchain which is reimbursed to the miners through block rewards and transaction fees.

What Bitcoin enables is an E-commerce system that bases itself on a cryptographic proof rather than trust which removes the need for a trusted third party. The security of this proposed system relies on the fact that the collection of honest miners controls more CPU power than any collaborating group of dishonest miners. If a majority of CPU power is controlled by honest miners, the honest chain will grow the fastest and outpace competing dishonest chains, if any. To modify a past block, a dishonest node/group needs to recreate the proof-of-work of the block and propagate it further to surpass that of the honest nodes. Since the probability of a slower attacker catching up diminishes exponentially as subsequent blocks are added, it becomes statistically almost impossible to fool the system.

Bitcoin incentivizes nodes to stay honest so that they can incrementally create more coins and therefore make more profit as otherwise the mechanism will be undermined and terminated. The system furnishes anonymity and privacy by making public keys anonymous and transactions transparent so that the public can observe a transaction yet they cannot associate it to the parties involved. This is analogous to stock exchange transactions that are made available to public without revealing the exchanging parties.

There are several other electronic currency systems on the market. Ripple [Fugger, 2004] is one such currency system. Bitcoin differs from Ripple by allowing its users to accept all issued currencies as proof-of-work helps establish and

propagate trust via digital signatures, whereas in Ripple the trust is established by a chain of trusted intermediaries between users. KARMA [Vishnumurthy, 2003] and PPay [Yang, 2004] are other electronic currency systems with comparatively higher overheads incurred particularly when the transaction volume is high. Mondex relies on smart cards to realize electronic currency, thereby making it a physical scheme [Stalder, 2002]. Yet, the non-physical nature of Bitcoin domain makes it suitable for online trade.

Over Bitcoin's not too distant history, it has seen many exchanges launch and many crash and harm the ecosystem. One of the most famous failures in Bitcoin exchange history was MtGox [Decker, 2014]. MtGox was not properly implemented, had open security flaws, and ran insolvent, i.e. the exchange did not actually hold the bitcoins it claimed to. MtGox was a small company with the majority of the exchange programmed by an entrepreneur. It had a small team of lawyers for compliance, and support team which could not keep up with demand. Unfortunately when Bitcoin experienced explosive growth in 2013, MtGox's small team for maintaining the exchange simply couldn't hold up. The servers were experiencing hours of latency in trade delay, salami attacks were being executed which slowly were draining the bitcoin on the exchange, and the site was unusable at times. All of this lead to a price crash on Bitcoin and eventually the company filing for bankruptcy.

As the number of Bitcoin transactions grows year over year [Table 1], it is clear that there is a need for many exchanges, as to prevent one from ruining the entire ecosystem. A competition of exchanges provides users with the ability to choose the one they feel the most comfortable with and encourages exchanges to advertise what they can do to provide the highest level of security for users. All in all, the exchange is holding the investments of its users, and this should never be taken lightly as in the case of MtGox.

Table 1 Total Bitcoin transactions since inception

Date	Number of transactions
2010	0
2011	~30,000
2012	~2,100,000
2013	~10,000,000
2014	~55,000,000

3 Implementation and Results

We setup the experimental environment by storing our project in the Amazon Web Services cloud, in particular Amazon Elastic Compute Cloud (EC2). We utilize PHP including secure PHP login sessions and JavaScript to build the Bitcoin exchange. Two linux instances on Amazon EC2 cloud are setup to accommodate the web server / database server (as btc-public) and the Bitcoin daemon server (as btc-wallet). We then install a web server stack on btc-public and the bitcoin daemon

on btc-wallet and create and instantiate database tables. After setting up the secure PHP login system, we establish a connection to it via bitcoin daemon API, thereby implement withdrawals and deposits of Bitcoin.

Once the system setup is complete, we test run it for proper execution. We then continue with security analysis where a SSL certificate is setup, servers are secured by closing open ports and locking accounts and a series of security tests are applied.

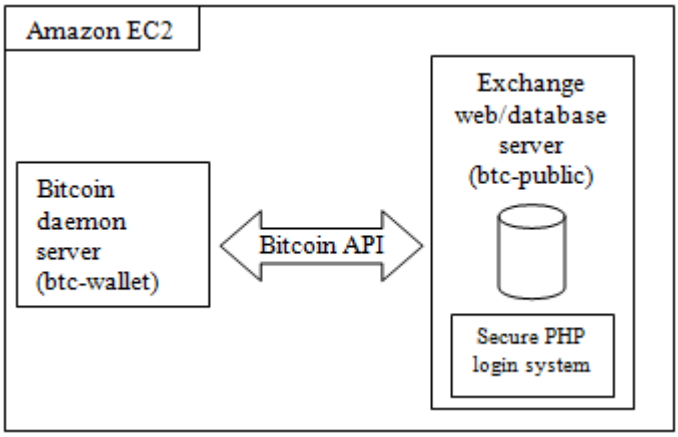


Fig. 1 Bitcoin exchange architecture in the Amazon EC2 cloud

3.1 Setup

Amazon Web Services (AWS) is used to host two virtual servers for the project. This solution works out very well for research and development projects, as the smallest AWS instance server only costs ~$0.01/hour, and it only charges when the server is running. This affords a cost efficient solution in a real cloud environment for little cost.

In order to connect to these servers, we need to maintain a private key file so that we will have an authenticated login. This file provides a strong authentication, as no third party can connect to the server without knowing the private key. During development, server ports together with traffic to IP addresses are deliberately set open. Once development is complete, servers are secured by implementing additional security countermeasures (section 4).

Figure 2 shows the complete Bitcoin daemon setup installation including the ~25GB blockchain. Bitcoin transactions are bundled into blocks. Each block contains its hash and the hash of the previous block. This creates the blockchain, which allows any node can verify any transaction. Block sizes vary, but the bitcoin algorithm allows “miners” to generate the appropriate hash to create the next block approximately every 10 minutes. The initial blockchain download can take a long time as every Bitcoin transaction needs to be downloaded in order to run a full Bitcoin node.

```
[ec2-user@ip-10-0-0-13 ~]$ bitcoind getinfo
{
    "version" : 90300,
    "protocolversion" : 70002,
    "walletversion" : 60000,
    "balance" : 0.04467332,
    "blocks" : 332386,
    "timeoffset" : -1,
    "connections" : 65,
    "proxy" : "",
    "difficulty" : 40300030327.89140320,
    "testnet" : false,
    "keypoololdest" : 1416954519,
    "keypoolsize" : 101,
    "paytxfee" : 0.00000000,
    "relayfee" : 0.00001000,
    "errors" : ""
}
```

Fig. 2 Bitcoin Daemon Setup

The next step is to setup the web server stack. For development, we employ the popular web stack XAMPP which combines Apache, MySQL, PHP, and Perl in one installation. Once this installation is complete, the Bitcoin exchange becomes available in the /opt/lampp/htdocs folder.

3.2 Secure Login

The next step is to setup the web server stack. For development, we employ the popular web stack XAMPP which combines Apache, MySQL, PHP, and Perl in one installation. Once this installation is complete, the Bitcoin exchange becomes available in the /opt/lampp/htdocs folder.

username	email	password	salt
test_user	test@example.com	00807432eae173f652f2064bdca1	f9aab579fc1b41ed0c44fe4ecdb
luke	luke.carr4@gmail.com	ff8a3cf89f014afeb059963584d6	9ad34ad4f4918100f904ec6d748
john	john@john.com	1d094bc2707429c77c8e64af0222	59db471e95b59f227597fd30b41

Fig. 3 Storing hased {password || salt} securely in the database

During a login attempt, a user entered password is hashed to verify it against the same user's hashed database record. If verified, the session variables are set and the user is allowed to access the Bitcoin exchange. If verification fails a log entry is made in the database to record the failed login attempt. To prevent against a brute force attack, once a user fails login 5 times, the account is locked for 2 hours.

Once session variables are set, a function called login_check() is invoked at the beginning of each secure page to verify that the current user is allowed to access it. The user password is revalidated at each page refresh, and a new unique verified session is generated for each page access to prevent session hijacking. Figure 4 displays the screenshot of a cookie

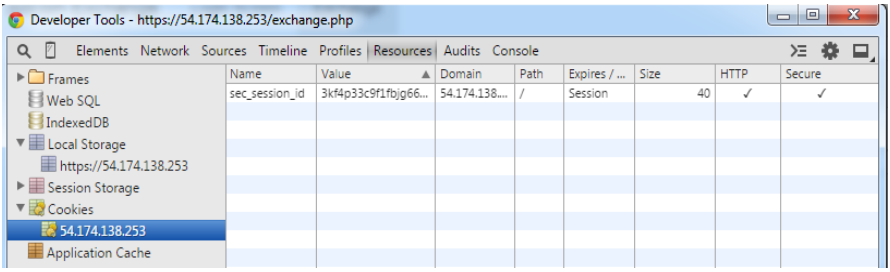


Fig. 4 A sample cookie for the Bitcoin exchange

As seen in Figure 4, the cookie expires after the session is complete, and it is encrypted. After ensuring that the login system is setup and fully functional, we continue with setting up the interface to the Bitcoin API.

3.3 Bitcoin API Interface

Once a full Bitcoin node is running on a machine, you receive access to the Bitcoin API which is what empowers Bitcoin. This API supports various services such as generating addresses, creating transactions, moving Bitcoins between accounts, etc.

Since the Bitcoins are stored on a separate server, a communication channel needs to be set up between the web server and the Bitcoin server. To ensure the integrity of this communication, we use a private key and a certificate so that the Bitcoin server knows that the web server is the only party who is allowed to make the API calls. Once this setup is complete, actual communication is handled by the JSON-RPC [Berriz, 2009] protocol. This allows the PHP logic to directly move coins between accounts, generate addresses for users to deposit with, and process withdrawals. The webserver and Bitcoin daemon server are in a virtual cloud to prevent public access to the Bitcoin daemon server. Bitcoin server traffic should be tunneled through the webserver to provide a connection to the blockchain. The webserver communicates to the bitcoin daemon server through a secure local connection to process Bitcoin deposits and withdrawals on behalf of users.

3.4 Trade Exchange

Our Bitcoin exchange scheme runs on one webpage (exchange.php) which is powered by Javascript and PHP. All variables, order books, and account activity modules update dynamically via Javascript when a change is detected to prevent the user from having to refresh the page. This is completed by creating an update function which has an Ajax call to a PHP script. The PHP script returns the requested information to the correct container on the screen, and the function is put on a timer to update again after a certain period of time.

Deposit and withdrawal of USD currency is very simple (for testing purposes): The user needs to input the amount and click on submit. USD is only used in this system as a dummy currency to trade against the BTC to provide exchange functionality. To comply with laws and regulations of the country where the Bitcoin exchange is hosted, the exchange should register with its local government to obtain a license to trade fiat. (if required).

Deposit and withdrawal of BTC use the above mentioned API. When a user clicks deposit BTC, an API call is made to the server requesting a Bitcoin address associated with the users account. It is important that each time the user makes a deposit they deposit into a new Bitcoin address. Although Bitcoin does provide a certain level of anonymity, since the full transaction ledger is available to view on the blockchain, if one can match an individual with an address he will be able to see where the individual is sending his coins. To prevent anyone from detecting that user X has sent Y amount of coins to an exchange, a new address is generated for each deposit which keeps any third party from determining whether or not user X is sending to an exchange or any arbitrary bitcoin service or website. The API command "bitcoind getaddressbyaccount x" will shows a long list of generated addresses used while testing the scheme.

3.5 Two-factor Authentication

One of the most important features a secure exchange should provide is two-factor authentication. On account creation, the account owner should provide a cell phone number which only the account holder has access to. This binds the users' account to a physical device which should be secured as it acts as the final line of defense in case of theft on the users' account.

Two-factor authentication is first used when the user attempts to log into their account. A secure code is sent via SMS to the users' cell phone, and the user must input the code to prove they are the account owner. This ensures that the users' account information and history is secure even if their password is stolen. Additionally, the user should have to authenticate via SMS code before any Bitcoin withdrawal. If a user leaves their computer unattended and logged into the exchange, their Bitcoin will be vulnerable to theft without an additional layer of two-factor authentication. By requiring the user to input an SMS code before withdrawal of Bitcoin, this proves that the original account holder is present when removing Bitcoin from the exchange.

4 Security Analysis

In order to facilitate development, the servers were initially set open to attacks on many default accounts (ftp, MySQL, etc) so that anyone who attempts can connect easily. After development was complete, all ports that are no longer necessary are closed to secure the servers. This is implemented easily using Amazon Web Services security groups, which allows one to attach a server instance to a security policy.

All traffic is permitted from the web server, as it is on a virtual private cloud connection. This allows the web server to communicate with the wallet server to facilitate API commands. The only port open to the internet on the wallet server is 8333 which is the port that the bitcoin protocol uses to communicate with other hosts on the peer-to-peer network. As previously mentioned, for additional security, bitcoin communication can be routed through the webserver to prevent any public interfaces to the wallet server. The btc-public server which hosts the web server and database only has the required ports open to allow users to connect to the website.

Closing unused ports ensures that the server protects itself against potential vulnerabilities. For maintenance and additional development, port 22 may also be opened and locked to a source IP address to allow SSH access.

SQL injection attacks are not possible in our Bitcoin exchange system due to the nature of the query builder that PHP offers. To execute a query, a statement object is created with the MySQL connection and a query is loaded into it by calling the prepare function. All variables are set as "?" character, the bind_param function should be called to bind variables to the "?" character in the query. PHP will not allow any character into the query which would facilitate a SQL injection.

All in all, the Bitcoin exchange system supports the following security principles:

- Principle of Least Privilege – The MySQL account which manages the Bitcoin exchange on the web server is set to have the minimum amount of privileges for the generic operations SELECT, UPDATE, and INSERT.
- Principle of Fail-Safe Defaults – Each time the user takes an action on the Bitcoin exchange, the PHP script checks to ensure that the user is still logged in and has access to the exchange. If for some reason the session is terminated or the user logs out, the user will not be able to access any of the scripts or functionality of the exchange.
- Principle of Fail-Safe Defaults – Each time the user takes an action on the Bitcoin exchange, the PHP script checks to ensure that the user is still logged in and has access to the exchange. If for some reason the session is terminated or the user logs out, the user will not be able to access any of the scripts or functionality of the exchange.
- Principle of Fail-Safe Defaults – Each time the user takes an action on the Bitcoin exchange, the PHP script checks to ensure that the user is still logged in and has access to the exchange. If for some reason the session is terminated or the user logs out, the user will not be able to access any of the scripts or functionality of the exchange.
- Principle of Open Design – *All components are implemented using publicly available encryption and hashing algorithms, there is no "security through obscurity" present in this exchange.*
- Principle of Separation of Privilege – In order to execute actions, the user must have a valid login session and cookie. These two checks provide proper integrity and authentication to the system. If the system only checked for a login, a user may be able to spoof a cookie or session ID in order to make invalid transactions.

- Principle of Least Common Mechanism – Each user executes in a separate session with separate database connectors. Due to the stateless nature of PHP scripts (users validated on each action, session IDs renewed on each action), there is not the possibility of sharing access to resources.
- Principle of Psychological Acceptability – All of the underlying encryption, hashing, session id regeneration, and security is hidden from the user. The user is only held responsible for his email address and password.

5 Conclusion

We present the design and implementation of a Bitcoin exchange scheme in the Amazon EC2 cloud with tested and verified security. Our results meet and exceed our initial expectations. Each component of this system is secure and actively defends against a wide variety of attacks. The servers themselves are protected from application vulnerabilities as all unused ports are closed. All communication from the public server to users, as well as communication between the public server and the wallet server are encrypted against a man-in-the-middle attack. The application itself is secure as user accounts and passwords are stored in a database not accessible via injection or remote connection.

If all online Bitcoin related services took security this seriously, attacks and theft could be prevented. Some developers don't understand the risks behind the code they write, which leaves wide open vulnerabilities to attackers. In an ecosystem where the reward for hacking is a pseudo-anonymous currency, the stakes are high for Bitcoin applications to secure their systems. All it takes is one hacker to gain access to a Bitcoin node server and they can completely drain it of all funds within seconds. One additional measure large exchanges can take is to store 90-95% of user funds offline [7] which would protect financial assets even in the worst case event of an attacker gaining access to the Bitcoin wallet server. Bitcoins are the most secure when they are removed from computing all together and stored on a physical medium (USB drive, Trezor hardware wallet, etc). An exchange does not need all of its Bitcoin assets on a Bitcoin daemon server at any given point in time. Ironically, the most secure method of storage is redundant offline devices stored in traditional bank lock boxes.

Our Bitcoin exchange can be easily extended to support deposit/withdrawal of multiple cryptocurrencies, and allow trading between them. To support fiat currency withdrawals and deposits would require one to follow specific laws on a per-country basis, rendering it unfeasible to implement.

The future for Bitcoin looks bright, and as long as emerging companies put as much emphasis on security as they do quarterly profits, MtGox2.0 can be avoided.

As for the future work, we plan on extending the security features of our exchange scheme by experimenting realtime attack scenarios and observing how our system responds. Additionally, due to the scale of data involved with Bitcoin transactions, we will research the feasibility of applying a MapReduce algorithm to our exchange using Hadoop.

References

1. Androulaki, E., Karame, G.O., Roeschlin, M., Scherer, T., Capkun, S.: Evaluating user privacy in bitcoin. In: Sadeghi, A.-R. (ed.) FC 2013. LNCS, vol. 7859, pp. 34–51. Springer, Heidelberg (2013)
2. Babaioff, M., Dobzinski, S., Oren, S., Zohar, A.: On bitcoin and red balloons. In: 13th ACM Conference on Electronic Commerce, pp. 56–73 (2012)
3. Berriz, G.F., Beaver, J.E., Cenik, C., Tasan, M., Roth, F.P.: Next generation software for functional trend analysis. Bioinformatics **25**(22), 3043–3044 (2009)
4. Decker, C., Wattenhofer, R.: Bitcoin transaction malleability and MtGox. In: European Symposium on Research in Computer Security, March 2014
5. Eyal, I., Sirer, E.G.: Majority is not enough: Bitcoin mining is vulnerable. In: Christin, N., Safavi-Naini, R. (eds.) FC 2014. LNCS, vol. 8437, pp. 431–449. Springer, Heidelberg (2014)
6. Fugger, R.: Money as IOUs in Social Trust Networks & A Proposal for a Decentralized Currency Network Protocol (2004)
7. Krieger, M.: Five Years Ago Today Bitcoin was Born– Read Satoshi's Original White Paper. http://libertyblitzkrieg.com/2013/10/31/five-years-ago-today-bitcoin-was-born-read-satoshis-original-white-paper/
8. Miers, I., Garman, C., Green, M., Rubin, A.D.: Zerocoin: Anonymous distributed E-Cash from bitcoin. In: IEEE Symposium on Security and Privacy (2013)
9. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System. white paper, October 2008
10. Reid, F., Harrigan, M.: An Analysis of Anonymity in the Bitcoin System, Security and Privacy in Social Networks, pp 197–223 (2013)
11. Stalder, F.: Failures and successes: Notes on the development of electronic cash. The Information Society (TIS) **18**(3), 209–219 (2002)
12. Vishnumurthy, V., Chandrakumar, S., Sirer, E.: KARMA: A secure economic framework for peer-to-peer resource sharing. In: Proceedings of the 1st Workshop on Economics of Peer-to-Peer Systems, Berkeley, CA, June 2003

Part III
Information Systems and Internet Technology

REST-Based Semantic Annotation of Web Services

Cleber Lira and Paulo Caetano

Abstract The advance in research conducted in the web services area has allowed organizations to integrate their applications, enabling the automation of business processes in several areas. However, it is necessary to provide solutions that favor the selection and integration of these web services at runtime effectively and efficiently. In the current web architecture many of these web services are designed according to the REST architectural style. Given this demand, this work presents a tool that aims to make the semantic annotation of restful web services. An example of use is also discussed, in which a WSDL document is semantically described.

Keywords Semantic annotation · Restful web services · Semantic web · Tool

1 Introduction

Semantic annotation is a process that allows the inclusion of data in other data set in a number of different domains (e.g. annotation of documents, wikis, blogs, tagging) and it can be conducted in a manual, semi-automatic or fully automatic way [1]. Semantic annotation can be used in documents which syntactically describe web services enabling these services to set invocation, discovery and composition tasks autonomously. The web services built using the .NET architecture and the J2EE framework do not have semantics due to the constraints in the current web architecture and conflicts that may exist in patterns of proprietary technologies [2]. The number of tools and applications available as web services has increased dramatically in recent years [3], e.g. from May 2010 to April 2012, BioCatalogue, a life science services repository, increased the number of biomedical web services from 1627 to 2290 [4]. Similarly, Seekda, a general-purpose

C. Lira(✉) · P. Caetano
Department of Computer Science, Salvador, Brazil
e-mail: cleberlira@gmail.com, paulo.caetano@pro.unifacs.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_25

service repository, found more than 28,000 web services in a period of 68 months [5]. Such number of services available on the Web represents a challenge for users who wish to find the web service to perform a specific task. Research on semantic web services has been greater in service-oriented architecture (SOA) that is based on SOAP, and few are the solutions proposed to provide semantics to services implemented as the REST architectural style [6]. The WSDL 2.0 language is designed to describe restful web services, having been a W3C recommendation since June 2007 [7]. Therefore, considering the importance of the incorporation of semantic annotation of restful web services syntactically described in WSDL documents and the reduced number of tools to support these activities, this paper presents a tool to assist software developers in the semantic annotation of these services.

This paper is organized as follows: Section 2 describes the concepts related to the semantic annotation of restful web services described in WSDL from the SAWSDL specification. In Section 3, related work is discussed. Section 4 presents the architecture and the services offered by the tool. In Section 5, we present the tool and its main features, a generated semantic document model and an example of semantic search for the requisition of such document. Section 6 presents a conclusion and future work.

2 Semantic Annotation of Restful Web Services with SAWSDL

To perform the semantic annotation, SAWSDL defines extension attributes that can be applied both to WSDL and XML Schema Definition elements. SAWSDL has no specified semantics, as they only allow to semantically annotate WSDL documents through concepts defined in an [8] ontology. WSDL 2.0 is used to syntactically describe web services based on the SOAP standard, and it also serves as a W3C recommendation for describing restful web services [7]. To define the semantic vocabulary in WSDL documents, this article considers the ontologies described in OWL language. OWL is the most commonly language used for the construction of ontologies in order to determine the semantic description on the web [9]. SAWSDL differs from other approaches because it performs the semantic annotation of web services on the same WSDL document [10]. Therefore, the WSDL document will be able to describe the syntactic and semantic features of web services. SAWSDL was designed as an extension of WSDL because it allows web services developers to enrich the descriptions of these applications including semantic information [11]. Specifically SAWSDL defines two extension attributes that are used in defining semantic annotation [17]:

1. modelReference**:** Used to specify the association between elements of a WSDL document and/or elements of an XML Schema Definition in a semantic model. This association is made through one or more URIs, that is, it is possible to associate an element to more than one semantic model. ModelReference can be

used, for example, to identify whether a particular web service meets the requirements of a customer. This extension attribute can annotate: WSDL elements (e.g. Interface, operations and faults) and WSDL type definitions (e.g. complex type definitions, simple type definitions, declarations of elements and attributes).

Figure 1 exemplifies the use of the modelReference attribute to annotate a WSDL document, specifically in the operation element, with the goal of providing the description of the operation of the service and the performance offered by the operation. Such operation and behavior are shaped by the ontology in line 2 which is assigned in the URI of the modelReference attribute.

```
1 ...
2 <wsdl:operation name="order" sawsdl:modelReference=http://www.w3.org/2002/ws/sawsdl/spec/ontology/rosetta#RequestPurchaseOrder">
3 <wsdl:input element="OrderRequest"/>
4 <wsdl:output element="OrderResponse"/>
5 </wsdl:operation>
6 ...
```

Fig. 1 modelReference to annotate an operation

2. SchemaMapping**:** to specify mappings between semantic data and the schema (XML Schema), two extension attributes are added to XML schema element declarations and type definitions, called liftingSchemaMapping and loweringSchemaMapping. These attributes are used to perform semantic adjustments to the messages exchanged among web services. The settings have the goal of transforming the data structure specified in the WSDL document into the format required by the associated semantic model and vice versa [16].

Figure 2 illustrates the use of the liftingSchemaMapping attribute in an XML Schema Definition document. In this case, XSLT is used as a language for specifying the mapping of elements in the XML Schema Definition for the concepts defined in the semantic model. Therefore, to perform the mapping of data contained in an XSLT file, a client tool retrieves the value provided by the URI in the attribute (sawsdl: liftingSchemaMapping) determined in line 4, and applies to the elements specified by the XML Schema Definition. In Figure 2, the mapping defined in the XSLT file determined by the URI in line 4 will be applied to Confirmed, Pending and Rejected elements, defined in lines 6, 7 and 8 of the XSD.

```
1  ...
2  <xs:element name="OrderResponse" type="confirmation" />
3  <xs:simpleType name="confirmation"
4  sawsdl:modelReference="http://www.w3.org/2002/ws/sawsdl/spec/ontology/purchaseorder#OrderConfirmation"
   sawsdl:liftingSchemaMapping="http://www.w3.org/2002/ws/sawsdl/spec/mapping/Response2Ont.xslt">
5    <xs:restriction base="xs:string">
6      <xs:enumeration value="Confirmed" />
7      <xs:enumeration value="Pending" />
8      <xs:enumeration value="Rejected" />
9    </xs:restriction>
10 </xs:simpleType>
11 ...
```

Fig. 2 Using liftingSchemaMapping to annotate a simple type

3 Related Work

During the research on literature, similar proposals were found for the semantic annotation of web services; however, the solutions differ with respect to the implemented requirements. Based on the identified proposals, the following requirements are considered in the implementation of the proposed tool: (R1) to realize the annotation of restful web services: web 2.0 applications are implemented following the REST architectural style. The research in semantic web services has been greater with service-oriented architecture, which is based on SOAP, and few are the solutions proposed to provide semantics to services implemented as the REST architectural style [6]; (R2) to realize the annotation of restful web services syntactically described with WSDL 2.0: During the analysis of the solutions found for the semantic annotation of restful web services, it was possible to identify tools that perform the syntactic description of these services through different specifications, such as SA-REST and WADL; however, WSDL 2.0 has been designed to describe restful web services to be a W3C recommendation since June 2007 [7]; (R3) to use pre-existing Ontologies on the web: during the analysis of the work it was found that the tools make use of pre-existing OWL ontologies on the web and it was found that most of the analyzed tools do not implement this functionality; (R4) to use technologies that are W3C recommendations: using W3C recommendations facilitates a supposed adoption of the tool by the software industry; (R5) to use ontological concepts defined in OWL. OWL is the most used language for the construction of ontologies in order to determine the semantic description on the web [9]; (R6) to maintain a semantic web services repository: these services may be used in the future to perform semantic search; (R7) to maintain an Ontology repository: to keep the ontologies that are used in a repository to process the annotations of a WSDL and/or XML Schema Definition document. The elements of these documents will be annotated using the semantic concepts of these ontologies which are stored in the local repository or of those existing on the web. This repository can also be used in the solution of an Ontology Recommender; (R8) the tool is extensible: the application architecture should be designed so as to enable the increase of new requirements, mitigating the maximum possible changes in its structure; (R9) to perform the annotation of web services built according to the SOAP standard. During the analysis of solutions, tools that perform semantic annotation of web services that are built according to the SOAP protocol have been identified, and in this case they will serve as the basis of comparison among the tools. Some of the requirements mentioned above are used in the reviewed related work, for example, the tool presented by [2] does not meet requirements (R1) and (R2) since it does not support the annotation of semantic restful web services. Also, the tool does not allow the use of online ontologies and thus does not meet requirement (R3). With regard to requirements (R4) (R5) (R6) (R7) and (R9) it can be stated that the tool meets them. Regarding requirement (R8) the tool does not support the growth of new developing formats. Regarding SWEET tool [14], the authors use Rests and MicroWSMO as a solution to provide

the semantic representation of services that are syntactically described in HTML documents; however, the services that have their exposed syntactic descriptions in WSDL documents are not semantically annotated and consequently do not support requirement (R2). Requirements (R3), (R5) (R6) (R7) and (R8) are not met by the tool. Requirement (R4) is partially met, because some technologies used in the solution are not W3C recommendations, such as WSMO. WSMO Studio [15] is a solution that integrates several tools to treat semantic web services. Requirements (R1), (R2), (R3) and (R5) are not met by the tool. The tool does not support WSDL 2.0, and thus does not provide support for restful web services, and the ontological vocabulary is determined by WSMO, making the requirement (R4) partially met. The tool does not meet the requirements (R6) (R7) and (R8). The RADIANTWEB application [3] is web-based and meets most of the requirements. Requirement (R4) is partially met, because the syntactic description of web services uses the standard WADL, which did not become a W3C recommendation. The tool does not meet requirements (R6) (R7) and (R8). IRIDESCENT [13] is, among the work reviewed, the most complete application that meets the proposed requirements. Requirements R6 and R7 have not been verified as being met by this tool. Table 1 shows the relationship between the identified requirements and the reviewed work.

Table 1 Evaluation of related works

	R1	R2	R3	R4	R5	R6	R7	R8	R9
SWS Editor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
[1] Anotação Semântica	No	No	No	Yes	Yes	Yes	Yes	No	Yes
SWEET	Yes	No	No	Partially	No	No	No	No	Yes
WSMO Studio	No	No	No	Partially	No	Yes	Yes	No	Yes
RADIANTWEB	Yes	No	Yes	Partially	Yes	No	No	No	Yes
IRIDESCENT	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes

Section 4 presents the architecture of the proposed tool which is designed to meet the above requirements, in order to support the semantic annotation of restful web services.

4 Tool Architecture

The proposed tool, SWS Editor1, has three (Figure 3) layers whose purpose is to perform the semantic annotation of restful web services: (1) The view layer acts as

[1] https://github.com/cleberlira/SWS-Editor

the interface between the consumer and the tool in which the functionality is provided for the purpose of establishing the relationship between the client and the services offered. The view layer may be constructed in any technology, for example, a mobile Android application. In this article, the view layer is represented by a windows forms application written in .NET; (2) The application layer whose responsibility is to provide the services responsible for providing the semantic annotation of restful web services. In this layer there is a component in the upper part, which represents the audit, and two representing the files used as input and output format of the files generated by the services offered by the tool. The semantic annotation core is represented by the following services: (i) Structural Validation, whose responsibility is the analysis of WSDL and XML Schema Definition documents (instance, training) in order to identify whether their structures were defined according to the definitions of the XML language of W3C and the obligation of the existing elements in the instance with the respective elements and attributes; (ii) Semantic Annotation, which aims to provide semantic annotation of restful web services syntactically described in WSDL using the SAWSDL specification; (iii) Obtaining Ontologies on the Web, whose responsibility is to find ontologies that are available on the Internet; (iv) Web Services on the Web, which aims to obtain the restful web services that are available on the Internet and (v) Ontology Recommender, whose responsibility is to provide an ontology recommender to perform semantic annotation of web services similar to those already annotated by the tool. In this service the user provides a string that represents the name of the ontology and another that describes its goal when obtaining the desired ontology; (3) and, finally, the data layer whose responsibility is to provide a repository

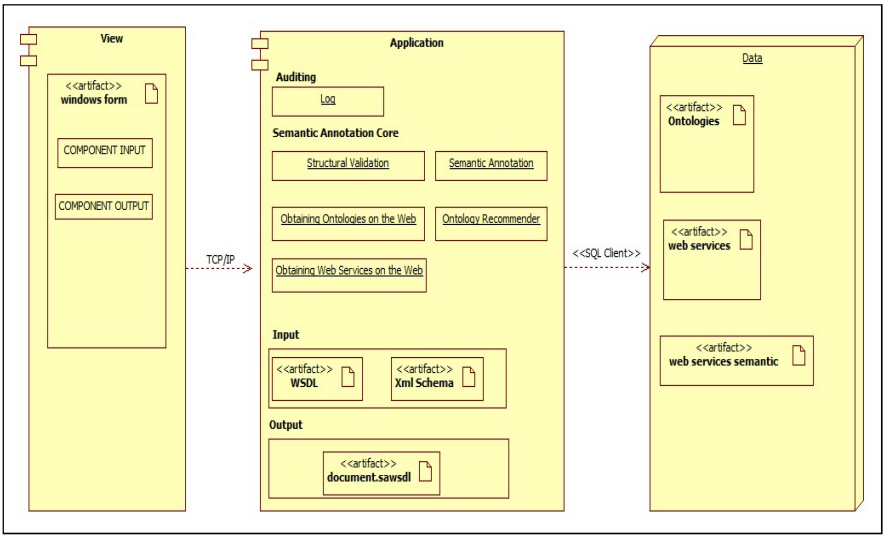


Fig. 3 Tool architecture

which is used by the services in which data is entered, i.e. (i) the inclusion of WSDL and SAWSDL files and Ontologies (ii) settings for authentication and authorization of access, (iii) treatment of error messages and (iv) audit information (objective of the service, use restrictions, version, developer, contact information). Communication between the layers, identified through the association of arrows, uses the TCP / IP protocol between the presentation layer and the application layer and the use of the SQL Server Native Client API between the application layer and the layer data.

Section 5 presents the steps required to perform the semantic annotation of restful web services syntactically described in WSDL.

5 Semantic Annotation for Restful Services With the Tool

To perform the semantic annotation of restful web services, the tool, as illustrated in Figure 4, consists of three main components: (i) Ontology, whose responsibility is to load the ontology, which will describe the web service; (ii) WSDL and XML Schema Editor, whose goal is to allow the user to edit the documents to be semantically annotated (the semantic annotation can also be performed using the context menu option, activated by right-clicking it when selecting the element of the WSDL or XML Schema Definition document) and (iii) WSDL / XML Schema, whose responsibility is to load the WSDL or XML Schema Definition document. The menu provides access to other options such as: Ontology Repository, which enables the user to access the stored ontologies allowing, also, the inclusion and removal of these documents. This repository will be available on the web through web services; Web Service Repository, which enables users to query both the web services syntactically described as well as the semantic web services and it also allows the removal and addition of new services. This repository will also be available on the web through web services.

To link the semantic concepts in the elements in a WSDL document, determined by ontologies defined in OWL, the SAWSDL specification is used. The following steps are taken to perform the semantic annotation of the web service and to make the document in the repository available: (i) select WSDL or XML Schema Definition documents; (ii) determine which elements of these documents will receive the semantic annotation; (iii) select the ontology that will determine the semantic vocabulary of the service (this selection can be performed either in the local repository (Ontology Repository) or in a repository of the web, or even through the ontology recommender (Ontology Recommender)); (iv) provide semantics to web services from what was determined in items (ii) and (iii); (v) save and make the restful web service available in the repository.

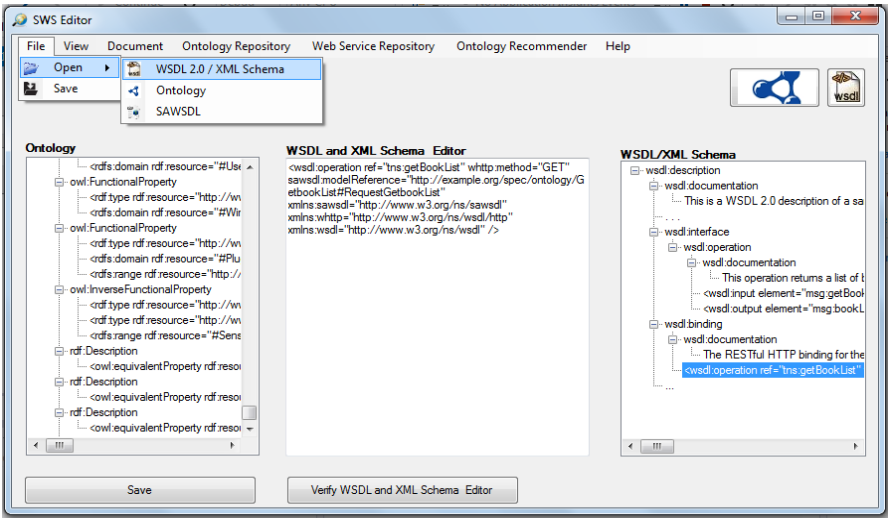


Fig. 4 Main screen of SWS Editor

Listing 1: Excerpt of a restful web service annotated with SAWSDL using the tool

```
01.<?xml version="1.0" encoding="UTF-8"?>
02.<wsdl:description
xmlns:wsdl="http://www.w3.org/ns/wsdl"
03.<wsdl:interface name="BookListInterface"
sawsdl:modelReference="http://example.edu.unifacs/ontol
ogies/bookstore.owl#Technologies">
04.<wsdl:operation name="getBookList"
sawsdl:modelReference="
http://example.edu.unifacs/ontologies/bookstore.owl#boo
kList#RequestbookList"
pattern=http://www.w3.org/ns/wsdl/in-out
style="http://www.w3.org/ns/wsdl/style/iri"
wsdlx:safe="true">
05.<wsdl:input element="msg:getBookList"/>
06.<wsdl:output element="msg:bookList"/>
07.</wsdl:operation>
08.</wsdl:interface>
09.<wsdl:binding name="BookListHTTPBinding"
type=http://www.w3.org/ns/wsdl/http
interface="tns:BookListInterface">
10.<wsdl:documentation>
.The RESTful HTTP binding for the book list service.
```

```
</wsdl:documentation>
11.<wsdl:operation ref="tns:getBookList" whttp:method="GET"
sawsdl:modelReference="
http://example.edu.unifacs/ontologies/bookstore.owl#GetbookLi
st#RequestGetbookList"/>
12.</wsdl:binding>
13.</wsdl:description>
```

Listing 1 shows the excerpt of the description of a restful service semantically annotated with SAWSDL using the proposed tool in this article. Line 3 uses the modelReference to categorize the BookListInterface interface element while line 4 associates the getBookList operation with the bookList class. Line 11 uses the modelReference by the GetbookList class to determine the behavior of the getBookList operation to be invoked from the GET operation.

With semantic web services it is possible to define the tasks of (i) invocation, (ii) discovery and composition (iii) of the services autonomously. Listing 2 shows an example of how to perform a search for a service using the semantic information provided in the WSDL document. In this case the search returns all the books that are in the technology category.

Listing 2: Example of Semantic Search

```
1. PREFIX books:
2.<http://example.org/categorization/books/technologies
#>
3.SELECT ?livro WHERE {
4.<http://example.edu.unifacs/ontologies/bookstore.owl#
Technologies> <http://example/categorization/books>
?livro }
```

6 Conclusion

The need for semantic annotations in restful web services was considered in this work, since Web 2.0 applications currently provide these services through their interfaces; however, these applications do not allow a software agent to automatically determine its meaning and to implement tasks such as invocation, discovery and composition automatically.

In this article, we presented a tool that semantically describes restful web services that are syntactically described in WSDL, besides implementing other requirements such as: (i) the use of existing ontologies and web services on the Internet; (ii) the creation of ontologies, web services and semantic web services repositories, which will be made available on the web, (iii) the creation of an ontology recommender based on the local repository. To perform the semantic annotation of restful web services we rely on the SAWSDL specification and the use of ontologies built in OWL language. The semantic annotation is facilitated by the tool because it provides many features for the user, such as an editor for the alteration of WSDL and XML Schema Definition documents and support to the context menu for the editing, inclusion and removal features of the semantics in such documents.

For future work, new features can be added, such as the possibility of adding semantics to WADL based documents and to microformats such as Rests and MicroWSMO. The Ontology Recommender service can be improved by similarity algorithms in order to provide a higher degree of automation in identifying the most appropriate ontologies in the description of restful web services.

References

1. Oren, E., et al.: What are semantic annotations Relatório técnico. DERI Galway (2006)
2. Budinoski, K., Jovanovik, M., Stojanov, R., Trajanov, D.: An application for semantic annotation of web services. In: 7th International Conference for Informatics and Information Technology, February 2010
3. Guttula, C.: Radiantweb: A Tool Facilitating Semantic Annotation Of Web Services (Doctoral dissertation, University of Georgia) (2012)
4. Goble, C.A., Belhajjame, K., Tanoh, F., Bhagat, J., Wolstencroft, K., Stevens, R., Nzuobontane, E., McWilliam, H., Laurent2, T., Lopez2, R.: BioCatalogue: a curated web service registry for the life science community. In: Nature Precedings (2009)
5. Seekda repository for Web services [Online] Available from http://www.seekda.com/. [Acessed December 2013]
6. Xavier, O.C.: ServiçosWeb Semânticos Baseados em RESTful Um Estudo de Caso em Redes Sociais Online (2011)
7. Mandel, L.: Describe REST Web services with WSDL 2.0. Rational Software Developer, IBM (2008)
8. kopecky, J., et al.: SAWSDL: semantic annotations for WSDL and XML schema. IEEE Internet Computing **11**(6), 60–67 (2007)
9. Panziera, L., Palmonari, M., Comerio, M., De Paoli, F.: WSML or OWL? A lesson learned by addressing NFP-based selection of semantic web services. In: Non Functional Properties and SLA Management (NFPSLAM 2010) Workshop (2010)
10. Prazeres, C.V.S.: Serviços Web Semânticos: da modelagem à composição. Tese de Doutorado. Universidade de São Paulo (2009)
11. Klusch, M., Kapahnke, P., Zinnikus, I.: Adaptive hybrid semantic selection of SAWSDL services with SAWSDL-MX2. Semantic-Enabled Advancements on the Web: Applications Across Industries: Applications Across Industries, p. 146 (2012)

12. Lausen, H., Farrell, J.: Semantic annotations for WSDL and XML schema. W3C recommendation, W3C (2007)
13. Stavropoulos, T.G., Vrakas, D., Vlahavas, I.: Iridescent: a Tool for Rapid Semantic Web Service Descriptions (2011)
14. Maleshkova, M., Pedrinaci, C., Domingue, J.: Supporting the creation of semantic restful service descriptions (2009)
15. Dimitrov, M., et al.: WSMO studio–A semantic web services modelling environment for WSMO. In: The Semantic Web: Research and Applications. Springer Berlin Heidelberg, pp. 749–758 (2007)
16. de Oliveira, D.M., Menegazzo, C.T., Claro, D.B.: Uma Análise Conceitual das Linguagens Semânticas de serviços Web focando nas Composiç oes: Comparaç ao entre OWL-S, WSMO e SAWSDL (2009)
17. Kopecky, J., et al.: SAWSDL: semantic annotations for WSDL and XML schema. IEEE Internet Computing **11**(6), 60–67 (2007)

Synthetical QoE-Driven Anomalous Cell Pattern Detection with a Hybrid Algorithm

Dandan Miao, Weijian Sun, Xiaowei Qin and Weidong Wang

Abstract Owing to more attention on quality of service (QoS) for subscribers, the mobile operators should shift evaluation standard from QoS to Quality of Experience (QoE). However, most researches in the field focus on end-to-end metrics, few of them consider synthetical QoE in the whole network. For mobile carriers, it is more significative to improve the overall system performance at the lowest cost. Therefore, the comprehensive evaluation of all users is more suitable for network optimization. As voice is still the basic service, we consider anomaly detection about voice service in this paper. Firstly, two synthetical QoE parameters, quality of voice (QoV) and successful rate of wireless access (WA), are considered to identify abnormalities of cells from the aspect of integrality and accessibility respectively. Then, we use a hybrid algorithm combining self-organizing map (SOM) and K-means to classify abnormal data points into several categories. After that, the data points for cells are treated as time series to compute the proportions in each anomaly model, which form anomalous cell patterns. To location where the exception happened accurately, the other 5 Key Performance Indicators (KPIs) are selected by association Rule according to the correlation between two synthetical QoE parameters. They are used to identify specific classes of faults. The experiment shows that the proposed method is effective to visualize and analyze anomalous cell patterns. It can be a guideline for the operators to perform faster and more efficient troubleshooting.

Keywords Data mining · SOM · KPIs · Synthetical QoE · Anomalous cell pattern

1 Introduction

Wireless mobile communication is currently shifting its focus from 2G/3G towards 4G/LTE, and now more people and institutions have begun to study the technology

D. Miao(✉) · W. Sun · X. Qin · W. Wang
Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei, China
e-mail: {ddmiao,sunwj}@mail.ustc.edu.cn, {qinxw,wdwang}@ustc.edu.cn

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_26

of 5G. This transition is not only related to the evolution of the access technology, but also to the development of many factors, including service provisioning/demands, user expectations and varieties of customers. In this context, the measurements of network performances are transferred from QoS to QoE. QoE is significant both in user and network perspective. However, most researches on QoE-aware management and control mainly utilize end-to-end QoS, such as packet loss rate, delay, jitter, etc. And subjective QoE estimation requires users' involvement and quantifies QoE in terms of a Mean Opinion Score (MOS), which may be influenced by types of participates and contents [1][2].

For mobile telecommunication carriers, they pay more attention on the integrated users' perception from the perspectives of macroscopic. It is no need and high-cost for them to ensure all users to achieve the optimum or monitor end-to-end QoS parameters and QoE of each user directly. Hence, for network optimization, it is more lucrative and suitable to consider synthetical and measurable QoE of a whole system.

KPIs are the first choice for operators to address this issue, since there are over a hundred possible measurements in the cellular network. The degradation of KPIs can indicate anomalies of network intuitively, which can reflect the status of users' application. In the past few years, research progress has been made in anomaly detection using KPIs. Single threshold methods, such as the univariate and the multivariate anomaly detection tests [3] were currently deployed by local cellular network operators. Their drawbacks were that detection results may rely on when outliers are present in the data set. Obviously, if the outliers are present in the training data, the statistical values will deviate from normal points. On this basis, In [4], Pichanun detected global anomalies as well as identified which KPIs of core network are abnormal with SOM. He used the codebooks of SOM to represent the statistical mean and variance. In [5][6], Gabriela used different univariate methods, such as Empirical Cumulative Distribution Function (ECDF), Autoregressive, Integrated Moving Average (ARIMA) models, etc, to model the KPI behavior and test data against the built models. But these statistic methods only judged whether the cell is abnormal or not. Different from these, Jaana [7] proposed SOM, together with a conventional clustering method, to visualize and group similarly behaving cells. The conjoint analysis dealt with multi-dimensional KPIs simultaneously. However, it used simulated traffic data. After that, Pasi applied the proposed approach in the analysis of degradations in signaling and traffic channel capacity of a GSM network at the cell level of the cellular network [8].

However, all KPIs analysis above only indicate cell-performance status instead of user's perspective. In this paper, our work mainly focus on the following points. 1) All the KPIs are collected from China Mobile Network which is a live current mobile network of TD-SCDMA system in China. 2) We select two synthetical QoE metrics, QoV and WA, to classify the overall state of network objects into normal/abnormal. They reflect the user experience in the whole networks from integrality and accessibility respectively. In this phase, a hybrid algorithm combining SOM and K-means is proposed to divide the abnormal data points into different anomalous categories. Then, we consider KPIs of cells as time series, and compute the percent of samples

generated in each anomaly cluster, which form anomalous cell patterns. 3) Utilizing the correlation between KPIs, another 5 KPIs are selected to location where the exception happened accurately.

The rest of this paper is organized as follows, the overall SOM based analysis process for KPIs are outlined in Section 2. In Section 3, we classify abnormal data into different anomaly models and conduct cause analysis of anomalous cell patterns. Finally, Section 4 gives some conclusions.

2 Proposed Approach

2.1 Self-organizing Map

The Kohonen's Self-Organizing Map is a widely used neural network algorithm [9]. It maps a high-dimensional data onto a low-dimensional grid, which provides the probability for its visualization property. It usually contains two steps in the training phase, including Neuron Competition and Neuron Inhibition.

Neuron Competition. The basic SOM consists of a regular grid of map units. Each map unit is denoted by i and represented by a prototype vector $\mathbf{m}_i$, whose dimension is equal to that of the input data. All the prototype vectors of map units are initialized randomly. Once one input data vector $\mathbf{x}$ is picked up from the measurement space and the corresponding best-matching (winner) unit (BMU) c can been determined. BMU minimizes the Euclidean distance between $\mathbf{x}$ and $\mathbf{m}_i$ (1), which is the unit closest to the input data sample $\mathbf{x}$.

$$c = \arg\min_i \|\mathbf{x} - \mathbf{m}_i\| \tag{1}$$

Neuron Inhibition. Next is followed by the application of an update rule, which is defined in (2), for the map unit locations in the measurement space. In this step, the units that are far from BMU are inhibited.

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t)h_{ci}(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] \tag{2}$$

Where, $\alpha(t) \in [0, 1]$ is the learning rate factor. It is a monotonically decreasing function of time, just like $\alpha(t) = \alpha_0(1 - t/T)$. $h_{ci}(t)$ is the neighborhood function of the algorithm. A widely used form is the Gaussian neighborhood function as follows in (3).

$$h_{ci}(t) = \exp\left(-\frac{\|\mathbf{r}_c - \mathbf{r}_i\|^2}{2\sigma^2(t)}\right) \tag{3}$$

In which, $\mathbf{r}_c$, $\mathbf{r}_i$ denote the coordinates of the winner unit c and an arbitrary unit i in the discrete output grid of SOM respectively. Therefore, $\|\mathbf{r}_c - \mathbf{r}_i\|$ is the distance

between unit c and i in the grid of SOM. $\sigma(t)$ is the standard deviation of the distances between adjacent points.

2.2 Network Analysis Using SOM

The whole system procedure for our analysis is shown in Figure 1. It can be divided into two parts: 1)the process to build K anomaly models; 2)analysis of anomalous patterns.

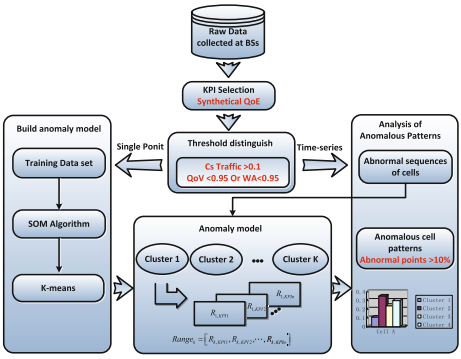


Fig. 1 A diagram to illustrate the phases of proposed analysis process

The Process to Build Anomaly Models. In this section, the hybrid algorithm combining SOM and K-means for the classification of anomaly models is presented. It consists of four steps: KPI selection, data preprocessing, SOM training and clustering analysis.

(1)KPI Selection. For each KPI, an objective value can be defined by the network operator in order to assess the acceptable performance of the network. Each object in the network has its own specific measurements to address different troubleshooting tasks. In this paper, we try to find the anomalous cell patterns and their causes from the viewpoint of synthetical QoE in a cell. All KPIs selected are shown in Table 1.

Synthetical QoE. Integrality and accessibility are the two most important points that voice users are concerned about. From these aspects, we choose QoV and WA as Synthetical QoE to reflect users' perception in the whole networks. Among them, QoV contains MOS, which is exactly the expression of the users' experience. And WA can show the overall performance of wireless access.

Other 5 KPIs. To location where the anomaly happens, we use association rules to choose other effective KPIs. Firstly, we quantize all the KPIs including synthetical QoE. Each level reflects the degree of anomalies roughly. When the level is far away from the normal ones, this KPI can regarded as anomaly. To simplify, each KPI only

Table 1 Key Performance Indicators

KPI	DFINITION
QoV	Voice quality in a cell
WA	Successful rate of wireless access
Cs Traffic	Traffic of calls per line
Drop	Number of TCH drops
Band35	Number of interference channels in Band 3-5
Band45	Number of interference channels in Band 4-5
RxQual67	Ratio of uplink RxQual in level 6-7

has two levels, anomaly or normal. We define a time point with N KPIs as $\mathbf{X(t)} = \{KPI_1, KPI_2, \cdots, KPI_N\}$. After quantizing KPIs, we only record abnormal KPIs, then $\mathbf{X(t)}$ will be changed as $\mathbf{AX(t)} = \{QoV, WA, KPI_1, KPI_3\}$. For datasets $\{AX(t), t = 1, 2, 3, \cdots\}$, association rules are used to pick out the KPIs that are related to QoV and WA. In our experiments, Drop, Band35, Band45 and RxQual67 in Table1 are selected. The Process is shown in fig. 2.

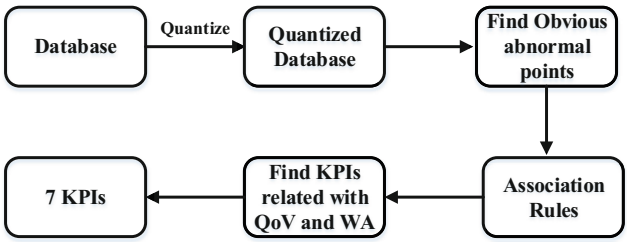


Fig. 2 The Process of Selecting Other 5 KPIs

(2)Data Preprocessing. SOM algorithm is a multivariate method that reflects the combination of variables, i.e., their joint distribution. Before it is applied, the data has to be preprocessed. In our experiments, we propose two assumptions to remove the samples that are not considered as the training data in the data preprocessing phase and one to determine the instantaneous anomaly later in analysis of anomalous cell patterns.

Assumption 1. When Cs traffic is less than 0.1, the cell is defined as super idle cell. In this case, QoV and WA may decrease to 0. We can not detect whether anomaly is caused by less number of calls handled or network performance. In general, these idle hours should be considered separately.

Assumption 2. Owing to the spare of anomalies, abnormal samples will be submerged into normal points if we analyze all data together. To address this issue, we apply inequality constraints with QoV $< th_1$ and WA $< th_2$ in order to filter the

Table 2 Dataset of Case Study

Location	Ningbo, China
Network Type	TD-SCDMA
Duration	01/12/2013 31/12/2013
Time granularity	1 hour
The Number of Cells	243
The Number of KPIs	64
Selected KPIs	7

normal samples. The threshold values are going to differ depending on scenes, in the paper, we set $th_1 = th_2 = 95\%$.

Assumption 3. When considering abnormal pattern for a cell, KPIs in a cell are time series sampled with interval of one hour. Cells can be treated as normal ones if the share of abnormal hours is very low. These anomalous points are usually regarded as instantaneous anomaly.

(3)SOM Training. After the subset of data is selected and preprocessed, SOM with M map units is trained. Each unit is an aggregated result of multi-dimensional KPIs. The trained SOM is adopted to visualize the multi-dimensional input data by the component plane representation of SOM.

(4)Clustering Analysis. After SOM training, the set of M codebook vectors are classified into several different numbers of clusters with k-means, in which the clustering process is carried out for the map units of SOM instead of the original data subset. We use DBI [11] to determine the optimal value of K. Each cluster is characterized by the joint distribution of KPIs. In Figure 1, we use the range of each KPI to depict cluster k, $Rang_k = [R_{k,KPI1}, R_{k,KPI2}, \cdots, R_{k,KPIn}]$. n is the number of KPIs selected. In general, different clusters describe different types of failures in wireless mobile network.

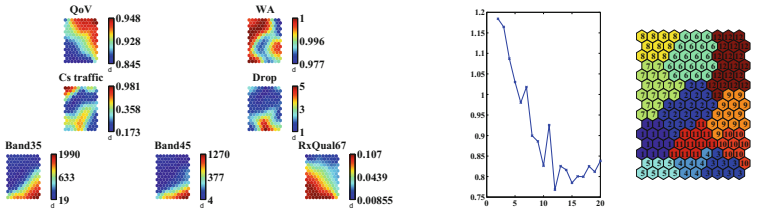
Analysis of Anomalous Patterns. The operators often monitor statistical KPIs with different sampling frequencies, which means indicators for a cell are time series instead of a single point. The cell pattern, computing the frequency of appearance in each type of clusters above for sequence data of a cell, reflects the characterization of the cell performance. In this paper, we only calculate proportions in five serious anomaly clusters to represent anomalous cell patterns.

3 Experiments: A Case Study

This section presents a study of deploying the above method in an operational environment. We apply the method on a KPI dataset collected from a live mobile networks. The detail information of dataset is described in Table2.

3.1 Anomaly Classification

According to *Assumption 1* and *Assumption 2*, abnormal data are determined as the input data of SOM. Then, the structure of data has been visualized using a SOM with 2-D hexagonal grid of size 15×10.



(a) som-visual (b) som-cluster

Fig. 3 (a)SOM of data samples representing possible network problems. (b)the corresponding clusters on top of SOM with k-means.

Figure 3(a) shows the component planes of SOM. The unit in the same position of component planes represents the same neuron of SOM. Different parts of each component plane show different values. The values of corresponding KPI are indicated by the color map and can be looked up by the color bar. For instance, the first component plane in Figure 3(a) represents the distribution of QoV. From the color map axis, the high values of QoV are denoted by crimson, which is located close to the top right corner of the map.

Codebook vectors of SOM are clustered by K-means. In Figure 3(b), we select $K = 12$ as the optimal number of clusters according to DBI whose trend is on the left of the figure. The map unit is labeled according to the cluster that it belongs to. And different clusters are shown with different colors. The properties of each data cluster is the joint distribution of KPIs, which can be analyzed combining the component plane representation shown in Figure 3(a). For example, $Rang_8 = [QoV \in [88\%, 94\%], WA \in [99.6\%, 100\%], \cdots, RxQual67 \in [0, 3\%]]$.

As discussed above, we analyze the anomalous cell pattern from the aspect of synthetical QoE in a mobile cell. During this process, QoV and WA are the indicators of abnormality in a cell, and other KPIs are used for analyzing causes of anomaly. In this paper, we pay more attention to five serious anomaly models, cluster 1, 3, 4, 5 and 12. Cluster 1 contains lower QoV, less than 90%. From Drop, Band35, Band45 and RxQual67 component planes in Figure 3(a), RxQual67 is up to more than 7%, but the number of interference channels and drops are relatively less, which demonstrates that the anomaly in cluster 1 may be caused by bit error.

QoV in cluster 3 is below 90%. RxQual67 is close to 10%, but unlike cluster 1, Band35 is larger. For these points, there may exist reverse link interference, which

causes uplink abnormality. Then, mobile terminals can not send the message to BS correctly. Then, high bit error rate and drops will emerge.

In cluster 4, it is easy to find obvious rise of call drops. QoV and WA both have slight variation, which suggest that the experts should examine surroundings that may affect channel conditions. And Band45 might be considered.

Anomaly in cluster 5 is similar to cluster 1, but the causes are different. Moderate interference and high BER may result in low QoV of cluster 5.

Cluster 12 is another type of abnormal model, in which WA is below examination criteria 98%. But the call drops, interference and BER are approximately zeros. When this happens, there may be something wrong in accessibility.

3.2 Analysis of Anomalous Cell Patterns

In Section 3.1, we divide the abnormal data into twelve categories and analyze the reasons for five serious anomalies briefly. In this section, KPIs of cells are considered as time series. The frequency of appearance in five serious abnormal clusters is computed. The vector containing these proportions over a time period is called anomalous cell patterns. Once determine the anomalous cell pattern, the causes of this anomaly may be clear.

By pretreatment, all the cells are broken into three groups: Idlecell, Normcell and ANormcell. Idlecell: Cs traffic is below 0.1 in 80% of hours. Normcell: the number of samples, whose Cs traffic is greater than 0.1, and QoV and WA are above 95%, exceeds ninety percent. ANormcell: failure to comply with above requirements are defined as abnormal cells.

In this paper, our analysis stresses the role of ANormcell with normal amounts of traffic. If there is no point belong to the five clusters for a cell, then we can regard it as a slight exception object. Through analysis, there are seven cells which generate more than 40% samples into five serious clusters, stating that they suffer from poor voice quality or wireless access. Figure 4 shows the percentages in five anomaly clusters of seven cells. Different anomalous cell patterns reflect different problems. Next, we will give two examples about Cell CDG081A and CDG904A to explain this problem.

In Figure 4, the histogram of cell CDG081A represents that samples in cluster 1 and 5 accounts for near 60%. In Figure 5(a), red dots denote the dot pairs of QoV and uplink RxQual in level 6-7, and blue line is a straight line fit. The linearity of fitting curve, in the form of $y = a*(1-x)$, shows that poor voice quality mainly results from uplink RxQual in level 6-7. When the cell is normal, the values of uplink RxQual and QoV are concentrated in a small range with small value a. The stronger the linear relationship, the wider the abnormal range. Some of the worse signal quality may be caused by interference in Band 4-5, which is shown in Figure 5(b). There also exist some situations with low QoV but less interference in Band 4-5. Other reasons, such as handover, fault or repeater stations, may cause low signal-noise ratio with an effect on RxQual.

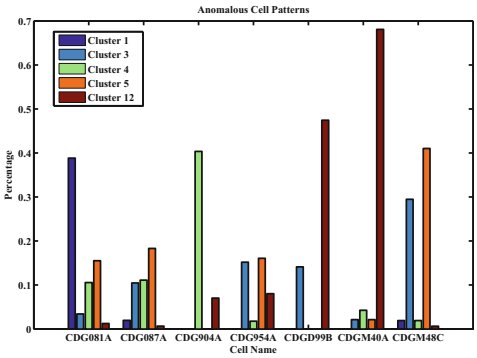


Fig. 4 The Percentage in Five Anomaly Clusters

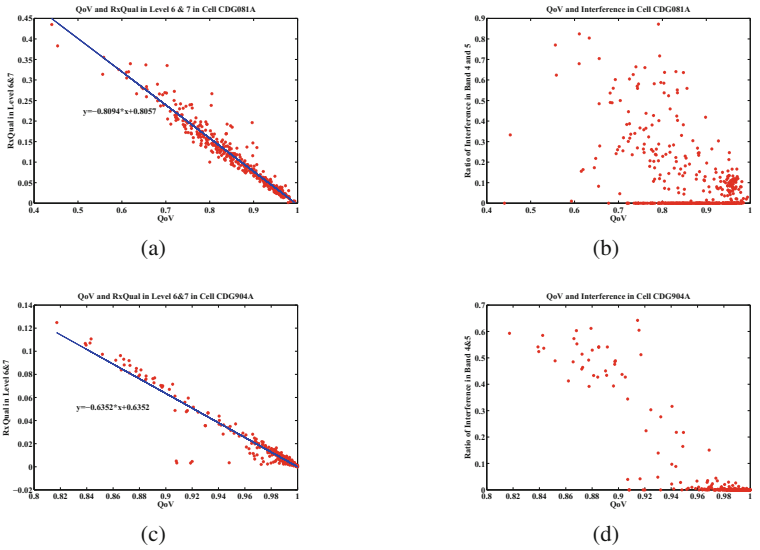


Fig. 5 (a)QoV vs RxQual67 in CDG081A. (b)QoV vs Band45 in CDG081A. (c)QoV vs RxQual67 in CDG904A. (d)QoV vs Band45 in CDG904A

Similarly, cell CDG904A generates more abnormal samples in cluster 4 and 12. According to the analysis in Section 3.1, channels with serious interference cause high ratio of uplink RxQual in level 6-7, which even results in some call drops. Different from CDG081A, the QoV values below 90% completely result from interference in Band 4-5 in Figure 5(d). The number of abnormal samples and the range of abnormal values in CDG904A are absolutely smaller than that in CDG081A. The parameter $a = 0.6352$ also reflects this phenomenon. Mainly because the area covered by CDG904A is surrounded by mountains, it adopts repeater station to expand coverage. This way increases interference in Band 4-5, which leads to a host of problems.

4 Conclusion

Most of papers focus on researching the mapping from end-to-end QoS to QoE or QoE-driven management. However, mobile operaters pay more attention to the overall system performance. In this paper, we select seven statistical KPIs, which can be monitored directly, to detect anomalous cell patterns. Among these, QoV and WA are selected to measure the synthetical QoE of a mobile cell, and others are used for cause analysis. Compared with traditional measurements, our method handles multi-dimensional KPIs simultaneously and detects anomalous cell patterns from a new perspective of synthetical QoE. Firstly, we use a unsupervised classification, combining SOM and k-means, to distinguish twelve anomaly clusters. The component planes of SOM visualize the range of KPIs in each category. Then, for time series of cells, the anomalous cell patterns, which are the frequency of appearance in five serious clusters, are exploited to represent the anomalies of cell behavior. At last, we discuss the possible reasons for anomalous cell patterns. Limited by available data, there are some defects for analysis of causes, but the whole effective analysis methods provide a good guide for anomaly detection and troubleshooting.

Acknowledgments

This research is supported by the National Basic Research Program of China (973 Program: 2013CB329004) and National High Technology Research and Development Program of China (863 Program: 2015AA01A706).

References

1. Siris, V.A., Balampekos, K., Marina, M.K.: Mobile quality of experience: recent advances and challenges. In: 2014 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), pp. 425–430. IEEE (2014)
2. Hoque, M.A., Siekkinen, M., Nurminen, J.K., Aalto, M., Tarkoma, S.: Mobile multimedia streaming techniques: Qoe and energy saving perspective. Pervasive and Mobile Computing **16**, 96–114 (2015)
3. Barreto, G.A., Mota, J.C.M., Souza, L.G.M., Frota, R.A., Aguayo, L.: Condition monitoring of 3G cellular networks through competitive neural models. IEEE Transactions on Neural Networks **16**(5), 1064–1075 (2005)
4. Sukkhawatchani, P., Usaha, W.: Performance evaluation of anomaly detection in cellular core networks using self-organizing map. In: 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008, vol. 1, pp. 361–364. IEEE (2008)
5. Ciocarlie, G.F., Lindqvist, U., Novaczki, S., Sanneck, H.: Detecting anomalies in cellular networks using an ensemble method. In: 2013 9th International Conference on Network and Service Management (CNSM), pp. 171–174. IEEE (2013)

6. Ciocarlie, G., Lindqvist, U., Nitz, K., Novaczki, S., Sanneck, H.: On the feasibility of deploying cell anomaly detection in operational cellular networks. In: 2014 IEEE Network Operations and Management Symposium (NOMS), pp. 1–6, May 2014
7. Laiho, J., Raivio, K., Lehtimaki, P., Hatonen, K., Simula, O.: Advanced analysis methods for 3G cellular networks. IEEE Transactions on Wireless Communications **4**(3), 930–942 (2005)
8. Lehtimäki, P., Raivio, K.: A SOM based approach for visualization of GSM network performance data. In: Innovations in Applied Artificial Intelligence, pp. 588–598. Springer (2005)
9. Kohonen, T., Maps, S.-O.: . Self-organizing maps. Springer series in information sciences, vol. 30 (1995)
10. Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J.: SOM toolbox for Matlab 5. Citeseer (2000)
11. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence **2**, 224–227 (1979)

A Mobile Group Tour Tracking and Gathering System

Chyi-Ren Dow, Yu-Yun Chang, Chiao-Wen Chen and Po-Yu Lai

Abstract This work proposes a group tour tracking and gathering system to let tour leaders and travel agencies keep the tour organized and ensure their safety, and reassure the tourists. This system provides several important functions, including fast roll call, ease of congregation, and prevention of members' involuntary separation from the group. We also design an algorithm for tourists to exchange and record their information. After they access to the Internet, they will send their information as well as the exchange information of other members to the database. Tourists will help each other to update the information, so that travel agencies and tour leaders can perform the tracking function.

Keywords Tracking · Gathering · Wearable device · Group mobility · Tourism system

1 Introduction

The well-developed wireless network technology in the world has become an Internet of Things (IoT), which connects every object to the Internet. The transmission speeds of mobile networks have increased and the transfer of big data can be completed in a shorter time. Smart wearable, smart city, smart environment, smart home and smart enterprises are top trends of IoT solutions. Traveling has become more and more popular in recent years. Foreign tourists arrival grew by 4.6% in the first half of 2014. The annual rate of tourist arrivals has grown each year. Tracking service is important for large groups and it is used to monitor and improve travel enterprises [1, 2]. To provide a tracking service, there are no information technology solutions to track and gather and thus to make the existing patterns become more convenient and efficient for group tour.

C.-R. Dow(✉) · Y.-Y. Chang · C.-W. Chen · P.-Y. Lai
Department of Information Engineering and Computer Science,
Feng Chia University, Taichung, Taiwan
e-mail: {crdow,m0428411}@fcu.edu.tw, cacawen@gmail.com, adoph2003@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_27

Tour leaders have to make sure the overall tourist satisfaction and safety in the whole tour. When they guide a group tour, there may have the following problems: (1) the members who travel with family pay more attention on safety and facilities. They may worry about other family members, such as children or seniors, to get lost with others in the unfamiliar environment. (2) Tourists present huge business opportunities, but none of the travel system can collate and analyze the detail of tourist records in an efficient way.

Especially, the international roaming is quite expensive. Even though foreign tourists could access the Internet by using Wi-Fi, the pubic Wi-Fi cannot cover all places. It is sometimes very difficult for foreign tourists to ask locals for directions when they get lost, so that to solve above problems will be more imperative for them. Mobile devices have four characteristics including convenience, ubiquity, personality and localization [6]. However, there is no service expressly designed for group tours. Wearable devices that are used alone or paired with smartphones are useful more for mobile interactions [5].

According to the above problems for group tours, we need a solution to achieve the tracking service by gathering the information from tourists using wearable technology. Besides, we have to make sure the information from tourists can be delivered. no matter they use what kinds of wireless technologies to access to the Internet. We focus on solving the obstacles during the travel. We use mobile networks or Wi-Fi and short-range communication with wearable technology to provide services of tourist-to-tourist and tourist-to-Internet to gather the information from tourists, and then tour leaders could track and gather their tourists to solve the above problems, and make travel safer and more convenient.

2 Related Work

Wong et al. [14] defined the group package tour as a trip, which was planned by a travel agency and offered at a single price. The group package tour operators include tour leader, restaurant, hotel, coach, optional tour, shopping, attraction, tour leader and others. Also, they pointed out that safety is of very important for package tour selection. Bowie [4] et al. identified important variables of tourist satisfaction: expectation, tourists' on-tour behavior and attitude, the perception of equity and the performance of tour leaders. Kennedy-Eden et al. [11] classified the mobile applications by the type of service for tourism into seven categories emerged, including social, navigation, mobile marketing, transaction, safety/emergency, information and entertainment. Emmanouilidis et al [7] considered the multiple taxonomy criteria of mobile guides' applications, which included mobile user interfaces, client architectures, context awareness, as well as offered highlighting functional, implementation and technological issues. They also pointed out that there are few literatures focused on user-centered evaluation of mobile guidance user experience. According to the above literatures, we can see that a tour leader is the master of the whole tour and has a very complicated role and requires powerful leadership skill and experience.

Huang et al. [10] based on [7] proposed a hiker tracking system, which used delay-tolerant, ad hoc, short-range and opportunistic communications techniques. Hikers will automatically exchange their IDs and locations when they pass by each other. When they reach the particular place, all of the information will be uploaded to the Internet. Thus, we can calculate the last known position of hikers. Ahas et al. [1] used the anonymous mobile position data of foreign tourists in Estonia to analyze their space consumption based on seasons. The mobile position data has great potential in recording invisible tourists' flows and provide a better way to measure tourism patterns. Ahas et al. [2] developed a tourism monitoring system that uses position data, which would be recorded automatically. A good temporal and spatial coverage of statistics and analyses helps enterprises to make strong business plan and assist investors in management. McKercher et al. [13] illustrated that using GPS technology to track and analyze the movement of tourists can reveal their patterns and behaviors in depth.

Wearable technologies are used to capture the physical and psychological status of human activities. Lara et al. [12] illustrated seven main issues in recognizing human activities when using wearable devices, including obtrusiveness, recognition performance, processing, selection of attributes and sensors, data collection protocol, energy consumption and flexibility. We focus on the issues of selection of attributes and sensors. According to the messages of sensors, we can know certain information of the wearer, such as surroundings, state of activity [9], location, health [16] and emotions [3],[8]. Perera et al. [15] pointed out that sharing data is the bigger picture of the IoT to valuable creation in open markets. However, the ownership, privacy and security of data should be safeguarded, and the anonymization of data would be the critical solution.

3 System Architecture

We propose a system for mobile group tourism, which can be applied on mobile networks or Wi-Fi, to track and gather tourists. In wireless sensor networks, data delivery models to the sink nodes consist of event-driven, query-driven, and continuous or hybrid combinations. The data delivery of our system is hybrid, which include contiguous (e.g., connection of subgroup members and tracking position of tourists) and event-driven (e.g., automatically rolling call, broadcast relating the information and emergency measures). It is assumed that all of the tourists wear a wearable device like a smart band, smart patch or smart badge. The wearable devices can access to the Internet through the Wi-Fi, and use short-range communications to connect with each other. Tour leaders can use smartphone to track tourists and tourist agencies can use management centers monitor service quality of the tour leaders. Therefore, the system can provide the real-time service for those users. During travel, the wearable device sends the GPS location and activity of tourists to the system through a Wi-Fi mesh router, and those data will be recorded to the database. The overview of the proposed system is shown in Fig. 1.

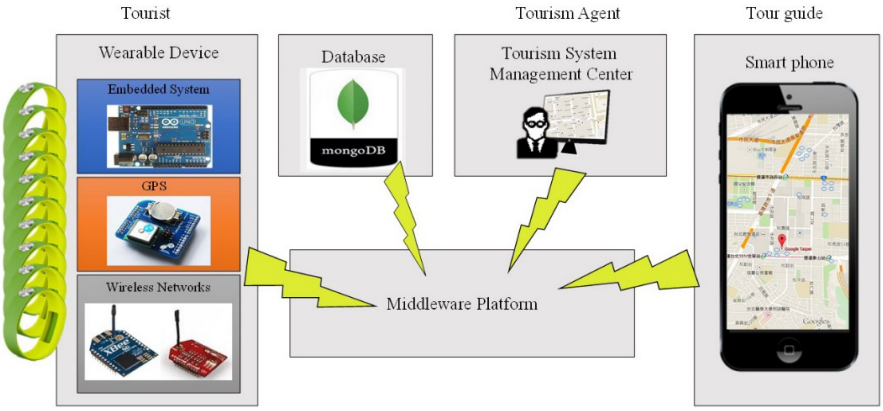


Fig. 1 Overview of the Proposed System

This system is divided into three subsystems, including a tourist subsystem, a tour leader subsystem and a management subsystem:

1. *Tourist Subsystem:* the position and activity information of tour leaders are collected and recorded by a contextual data processor. Dual-mode networks are used in wireless networks processor, which not only sends information of their positions and sojourn time to the database, but also link with other members during the traveling time. The visual and hearing notifications are presented in display and audio output. Three services are provided in this subsystem:
   (a) Subgroup State: the members of subgroup connect with each other. The wearable device will alarm subgroup members instantly when someone goes a far distance away from the subgroup,
   (b) Mobile Guide: mobile guide will send gathering time, direction and distance of appointed place to the tourists.
   (c) Emergency Measure: if tourists get into trouble, they can push the button of wearable devices, and then the tour leader will receive the emergency message immediately.
2. *Tour leader Subsystem:* contextual data processor and tourist management interface are included in this subsystem. The position of the tour leader is collected by the contextual data processor. Tourist management interface is visually presented in the APP of a smartphone. This subsystem provides two services:
   (a) Group Tracking: according to the GPS of tourist, the tour leader can handle the information of tourist. Thus, the guide can help tourists when they receive the risk message from the system.
   (b) Gathering: the APP provides an automatic roll call service to check tourists in order to reduce gathering time. When the tour leader changes the time or the gathering place, it would send the newest message to tourists.
3. *Management Center Subsystem:* middleware platform, group tour management interface and the database are included in this subsystem. The contextual date of tourists and tour leaders are collected and analyzed through the middleware

platform into database. It is a middleware for communications among tourists, tour leaders and travel agents. Group tours are tracked by the group tour management interface.

3.1 Tracking and Gathering Schemes

The basic idea of our algorithm is based on [10], to track the group tour. The group tour, which is guided by a tour leader, can be referred as follows:

$$G=\{L,TG\} \tag{1}$$

where G is a group tour, L is the tour leader and $T(G)$ is the tourists in the group tour. $T(G)$ can be represented as follows:

$$T(G) = \{t_1, t_2, t_3, \ldots, t_n\} \tag{2}$$

Table 1 Tracking and Gathering Tourists Algorithm

```
1.  Initialization:
2.  G • a group tour
3.  T(G) • the tourists in the group tour
4.  S_j • a subgroup
5.  t_i • a tourist
6.  CheckCounter • the counter of the check tourists
7.  UncheckList • the list of the tourists who are not
8.  here
9.  Automatic Roll Call:
10. for t_i in T(G)
11.     if t_i is here then
12.         CheckCounter++;
13.         else then
14.         Put t_i in UncheckList;
15.     end if
16. end for
17. if UncheckList != Null then
18.     for t_i in UncheckList
19.         Find the last known position of t_i from
20. database;
21.         Show the last known position of t_i on
22. mobile device;
23.     end for
24.     Show the CheckCounter and UncheckList on mobile
25. device;
26. else return All the members is here;
27. end if
```

where t_i is a member of the group tour and n is the number of members in the group tour. *T(G)* can be split up into the subgroups, which can be referred as follow:

$$S_1 \cup S_2 \cup \ldots \cup S_z = T(G) \tag{3}$$

$$S_1 \cap S_2 \cap \ldots \cap S_z = \emptyset \tag{4}$$

where S_k is a subgroup and z is the number of subgroups in a group tour.

$$S_k = \{t_{k1}, t_{k2}, t_{k3}, \ldots, t_{km}\} \tag{5}$$

where m is the number of members in the subgroup.

According to above process in wearable devices, tour leader can track all of the group members. Table 1 shows the tracking and gathering process in smartphone. Lines 1~6 are initialization. Automatic Roll Call takes the place of traditional to roll a call, in order to reduce gathering time. The members of group tour will be confirmed when they close to a certain distance (lines 8~15).

4 System Implementation and Prototype

We use Arduino, Raspberry Pi, GPS, Timer, Wi-Fi, Bluetooth, G-sensor, LCD and Magnetic Buzzer to build the system prototype. Arduino can sense and control the physical objects by a simple microcontroller board. We use it to build the prototype of wearable devices. According to above sensors, wearable devices detect the activity of tourists, such as subgroup connection state, position of tourists (e.g., latitude and longitude), movement state (e.g., tourist is moving or statics) and wireless state (e.g., connected or unconnected).

The message flow is shown in Fig. 2. A wearable device publishes the context message MSG1 to the middleware, and then the middleware will send MSG1 to other devices, which subscribe the wearable device. In other words, the tour leaders subscribe the members of group tour that they can keep the tour organized and ensure their safety by receiving context message from the members of group tour. Also, the travel agency can monitor all the group tours.

We use the open-source NodeMCU and Arduino hardware to build the system prototype shown in Fig. 3 (a) and (b), which presented on LCD, in order to let tourist know the direction and distance of gathering destination. When the gathering time is up, tourists will be notified by the sounds of Magnetic Buzzer.

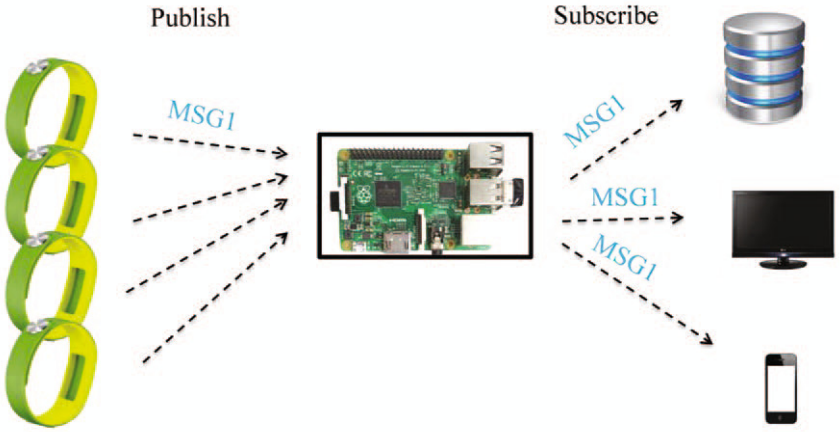
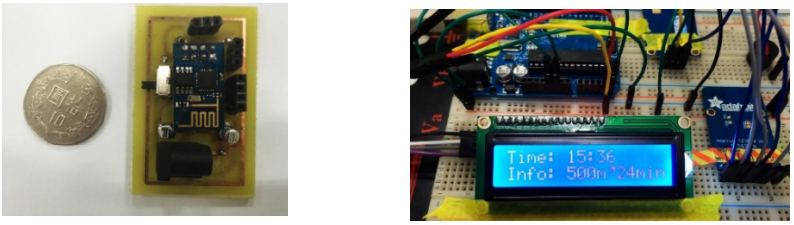


Fig. 2 The Message Flow



(a) Prototype Test of NodeMCU (b) Visible Information

Fig. 3 Prototypes of the Wearable Device

5 Experimental Results

After we classified 44 mobile applications of survey results from in the taxonomy of tourism mobile application by type of services. This research collected respondents, including 12 experts and 197 tourists. Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.

This section is divided into two analysis parts for tour leaders and tourists. We design the questionnaires using five-point Likert-scale (1=strongly unnecessary to 5=strongly necessary). We use descriptive statistics to compare the mean and ranks of mobile services and applications to understand what kinds are needed during the tour.

5.1 Tour Leaders

Twelve respondents from experts were analyzed. Table 2 shows the ranking of service in tourism. The first one is safety/emergency services. The second one is navigation services, the third is information services, the fourth is social services, and the last one is transaction services. We can see that all the mean values of services are more than three points, which means those services are really required for tour leader. The safety/emergency services are the most important services for tour leaders. The transaction services are the last one, but the 3.83 of mean value shows that many tour leaders also feel these services are necessary.

Table 2 The Services and Rankings by Tour Leaders

The Taxonomy by Type of Service	M	S.D.	R
Safety/emergency	4.72	0.547	1
Navigation	4.67	0.651	2
Information	4.50	0.640	3
Social	4.38	0.735	4
Transaction	3.83	1.403	5

5.2 Tourists

The total respondents of tourists are 197, which include 187 valid questionnaires and 10 invalid questionnaires, and the overall response rate was 95%. The characteristics of the respondents from tourists regarded to age, gender, education and occupation. Table 3 shows that information services are the most important services for tourists. The second one is safety/emergency services. The third one is social services, the fourth one is navigation services, and the final one is transaction services.

Table 3 The Services and Rankings by Tourist

The Taxonomy by Type of Service	M	S.D.	R
Information	4.47	0.664	1
Safety/emergency	4.40	0.717	2
Social	4.08	0.809	3
Navigation	3.81	1.153	4
Transaction	3.60	0.905	5

6 Conclusions

This system provides several functions to let tour leaders and travel agencies keep the tour organized and ensure their safety, and reassure tourists. Thus, the tourists can be prevented from getting lost from the group, and tour leaders will know the accurate location of tourists immediately if tourists do not gather on time. Tour leaders can take roll quickly, and travel agent monitors the tour leaders through an efficient monitoring system. Also, the huge business opportunities of tourist records could be collated and analyzed in an efficient way. The requirements of tour leaders and tourists are analyzed during the tour. Our results show that this system is required in tourism, as both the tour leader and tourists need safe mobile application services. From those reasons, we can see that our tracking and gathering system is essentially support for tour leaders to guide a tour. In the future, this system should reference our results to add more required functions, so that the system will be more useful during tours and abundant data from tourists can be collected.

Acknowledgment The authors would like to thank the National Science Council of Republic of China for financially supporting this research under Contract No. 104-2221-E-035-020-.

References

1. Ahas, R., Aasa, A., Mark, U., Pae, T., Kull, A.: Seasonal Tourism Spaces in Estonia: Case Study with Mobile Positioning Data. International Journal of Tourism Management **28**(3), 898–910 (2007)
2. Ahas, R., Saluveer, E., Tiru, M., Silm, S.: Mobile positioning based tourism monitoring system: positium barometer. In: Proceedings of the International and Communication Technologies in Tourism, Innsbruck, Austria, pp. 475–485, January 2008
3. Amyx, S.: Wearables Gushing with Emotions: a New Brand-Engagement Architecture. IEEE Transactions on Consumer Electronics Magazine **4**(1), 87–89 (2015)
4. Bowie, D., Chang, J.C.: Tourist Satisfaction: a View from a Mixed International Guided Package Tour. Vacation Marketing **11**(4), 303–322 (2005)
5. Brewster, S., Lumsden, J., Bell, M., Hall, M., Tasker, S.: Multimodal 'Eyes-Free' interaction techniques for wearable devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florida, USA, pp. 473–480, April 2003
6. Clarke, A.: Emerging Value Propositions for M-Commerce. Journal of Business Strategies **18**(2), 133–148 (2001)
7. Emmanouilidis, C., Koutsiamanis, R.A., Tasidou, A.: Mobile Guides: Taxonomy of Architectures, Context Awareness, Technologies and Application. Journal of Network and Computer Applications **36**(1), 103–125 (2013)
8. Grunerbl, A., Muaremi, A., Osmani, V., Bahle, G., Ohler, S., Troster, G., Mayora, O., Haring, C., Lukowicz, P.: Smartphone-Based Recognition of States and State Changes in Bipolar Disorder Patients. IEEE Journal of Biomedical and Health Informatics **19**(1), 140–148 (2015)

9. Hanai, Y., Nishimura, J., Kuroda, T.: Haar-like filtering for human activity recognition using 3D accelerometer. In: Proceedings of the IEEE Education Workshop, Marco Island, USA, pp. 675–678, January 2009
10. Huang, Y.T., Huang, J.H., Chen, L.J., Huang, P.: YushanNet: a delay-tolerant wireless sensor network for hiker tracking in Yushan National Park. In: Proceedings of the IEEE Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan, pp. 379–380, May 2009
11. Kennedy-Eden, H., Gretzel, U.: A Taxonomy of Mobile Applications in Tourism. E-Review of Tourism Research **10**(2), 47–50 (2012)
12. Lara, O.D., Labrador, M.A.: A Survey on Human Activity Recognition Using Wearable Sensors. IEEE Transactions on Communications Surveys and Tutorials **15**(3), 1192–1209 (2013)
13. McKercher, B., Shoval, N., Ng, E., Birenboim, A.: First and Repeat Visitor Behaviour: GPS Tracking and GIS Analysis in Hong Kong. International Journal of Tourism Space, Place and Environment **14**(1), 147–161 (2012)
14. Wong, C.K.S., Kwong, W.Y.Y.: Outbound Tourists' Selection Criteria for Choosing All-Inclusive Package Tours. Tourism Management **25**(5), 581–592 (2004)
15. Perera, C., Liu, C.H., Jayawardena, S.: The Emerging Internet of Things Marketplace From an Industrial Perspective: A Survey. IEEE Transactions on Emerging Topics in Computing (2015). doi:10.1109/TETC.2015.2390034
16. Perez, J.A., Leff, D.R., Ip, H.M.D., Yang, G.: From Wearable Sensors to Smart Implants Towards Pervasive and Personalized Healthcare. IEEE Engineering in Medicine and Biology Society **PP**(99), 1–13 (2015)

Model Based Evaluation of Cybersecurity Implementations

Aristides Dasso, Ana Funes, Germán Montejano, Daniel Riesco, Roberto Uzal and Narayan Debnath

Abstract Evaluation of Cybersecurity implementations is an important issue that is increasingly being considered in the agenda of organisations. We present here a model for the evaluation of Cybersecurity requirements. We start by establishing a set of security requirements in the form of a hierarchical structure to obtain a requirement tree, as it is prescribed by the Logic Score of Preference (LSP) evaluation method. Security requirements have been taken from the ISO/IEC 27002 standard. This requirement tree and an aggregation structure, built into a later step, form our Cybersecurity evaluation model, which allows to obtain a numerical final result for each system under evaluation. These final indicators, ranging into the interval 0..100, clearly show the degree of compliance of the systems under evaluation with respect to the desired requisites.

Keywords Cybersecurityevaluation · Cybersecurity implementation · Evaluation methods · Logic Score of Preference method · LSP method · Continuous logic

1 Introduction

Cybersecurity implementation, either in an existing working information system or in one even under development, can be a nightmare. In this process, not only physical details such as access policies to servers and installations must be taken into consideration but also software measures, such as firewalls, inline coding, etc., must be considered.

A. Dasso(✉) · A. Funes · G. Montejano · D. Riesco · R. Uzal
SEG, Universidad Nacional de San Luis, Ejército de Los Andes 950, San Luis, Argentina
e-mail: {arisdas,afunes,gmonte,driesco,ruzal}@unsl.edu.ar

N. Debnath
Winona State University, Winona, MN 55987, USA
e-mail: ndebnath@winona.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_28

To assess whether the correct steps in implementing security measures have been taken may be a difficult task. Numerous features should be taken into consideration, not only in information and software but also in computer installations, going from the development stages to fully working systems and from physical protection of hardware installations to the implementation of software gear and tools. Hence, the need of having available standards and tools to assess the degree of compliance with the standards is a must for people responsible of security implementation.

In this work, we present a model to assess Cybersecurity implementations. As part of the evaluation model, we have established a set of security requirements in a hierarchical form. We have taken these requirements from the ISO/IEC 27002 standard [16]. This is a well known and discussed standard that is part of set of standards on cyber security that have been constructed along the years with broad consensus and with contributions from a wide range of sources from academia, industry, etc. In Section 2, we give an overview of the state of the art in the field of evaluation of Cybersecurity implementations. Section 3 introduces the method applied for the development of the evaluation model proposed here. Section 4 explains the model proposed and, finally, in Section 5 we present our conclusions and future work.

2 Evaluation of Cybersecurity Implementations

Some work has been done on evaluation of methods, implementations, procedures, etc., of different areas of security. However, a great deal of these is based on testing, particularly on specific segments of security such as coding.

There are organizations that issue security implementation certificates, such as Card Industry [23] and Common Criteria [27, 28]. The later has received some criticism (see [18], [31]). Several commercial and government organizations use these certificates.

Private companies have also concerns about Cybersecurity and its evaluation, e.g. Oracle, that have a blog (Oracle Security Evaluations Blog) where it deals with subjects such as Common Criteria, FIPS and other security certificates [22], [3].

We can also find some governmental, educational and international institutions that also issue certificates or propose standards or tests for different parts of security implementations. CERT Division [4] of the Software Engineering Institute (SEI) of Carnegie Mellon University (CMU), "... was created in 1988 as the CERT Coordination Center in response to the Morris worm incident. ... is a national asset in the field of cyber security. We regularly partner with government, industry, law enforcement, and academia to develop advanced methods and technologies to counter large-scale, sophisticated cyber threats."

The National Institute of Standards and Technology (NIST) of USA have developed the Security Content Automation Protocol (SCAP) [20] that "is a synthesis of interoperable specifications derived from community ideas.". SCAP combines several standards, schemes, etc. that can be used to automate of various security measures.

IEEE also have its standards [25] and special initiatives "...with the aim of expanding and escalating its ongoing involvement in the field of cyber security", [14] and particularly the IEEE Center for Secure Design (CSD), that "...intends to shift some of the focus in security from finding bugs to identifying common design flaws." [13]

Other organizations have special security objectives, such as EVITA (E-Safety Vehicle Intrusion Protected Applications) whose objectives are "to design, verify, and prototype architecture for automotive on-board networks where security-relevant components are protected against tampering and sensitive data are protected against compromise.". EVITA is founded, at least partially, by the European Union [12].

The Department of Homeland Security of the USA has a Cyber Security Evaluation Tool (CSET®). "CSET helps asset owners to assess their information and operational systems cyber security practices by asking a series of detailed questions about system components and architecture, as well as operational policies and procedures. These questions are derived from accepted industry cyber security standards." [5].

Also the DoD, USA, have the Information Assurance Technology Analysis Center (IATAC), as part of the Defence Technical Information Center, that has published a report on the state of the art of "Measuring Cyber Security and Information Assurance" [2].

Some related work can be found in the literature. Weiß et al [30] studied a method to establish metrics for assessing security implementation based on evaluation of possible scenarios that could compromise the security of an installation. In A. Wang [29] there is a discussion about the foundations of what a security metrics should have. There is also a proposal of a metric that can be seen as a model for evaluating security. Barabanov et al. [1] have an overview of different standards, provisions, technologies, etc. that can be interesting as an introduction to the subject. Nicol et al. [21] proposed to adapt models for evaluation of system dependability to the security domain. A method called ADVISE created by LeMay et al. [19] permits the creation of security system models as well as it takes into consideration adversary attack preferences that represents how the adversary is likely to attack the system and the results of such attacks. Examples of evaluation of different ciphering techniques and their implementation, sometimes in hardware, although somewhat dated, can be found in the work of Elbirt et al. [11]. Also Wolf et al. [32] have evaluated, following some requirements, but in a rather subjective manner, the security of their implementation.

There is as well some concern about the training and credentials that Cybersecurity specialists should have, both for those embracing such a career as for the prospective employers, (see [24], [26]). Also the International Information System Security Certification Consortium, Inc., (ISC)® "is the global, not-for-profit leader in educating and certifying information security professionals throughout their careers. We are recognized for Gold Standard certifications and world class education programs." [15].

3 The Logic Score of Preference Method

In this section, we introduce the Logic Scoring of Preference (LSP) method [7, 8, 9, 10] employed in the development of our security evaluation model. This is a method for the realization of complex criterion functions and their application in the evaluation, optimization, comparison and selection of general complex systems.

Since it is a general evaluation method, it can be applied in the evaluation processes involved in the implementation of a security system in particular. The method is especially useful where the use of complex *and/or* decisions is necessary since is not a simple additive scoring method. As with many other methods, initially in LSP it must be clearly determined what the user requirements are, the main attributes of the system and their value preferences. These attributes are called *performance variables*. Each one of these variables is mapped into an *elementary preference* by defining and applying a corresponding *elementary criterion*, as the schema in Fig. 1 shows.

An elementary criterion is a function that transforms a performance variable (real value) into an elementary preference (value in [0,100]). An elementary preference represents the degree of fulfilment of a requirement, where 0 means that the requirement has not been fulfilled at all and 100 that it has been completely satisfied.

The elementary preferences, obtained from the transformation of each performance variable via the corresponding elementary criteria, are used as input to the LSP *Criterion Function* or *Aggregation Structure*. This function yields a single global indicator of the degree of fulfilment of the whole system requirements. It is built by aggregating elementary preferences by means of a set of operators. Aggregating preferences means to replace a group of input preferences by a single output preference, which denotes the degree of satisfaction of the evaluator with respect to the whole group of input preferences. Output preferences must be aggregated again until a single global preference is

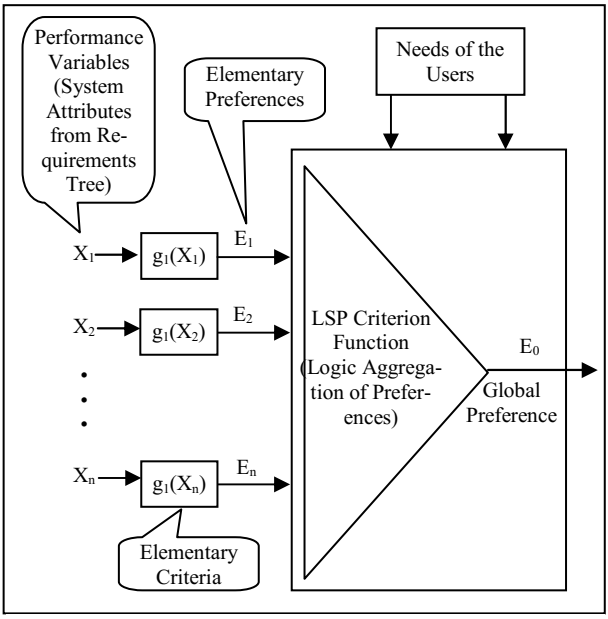


Fig. 1 An overview of the LSP evaluation process.

obtained. The global preference –obtained as output of the LSP Criterion Function– is the result of the combination of the elementary preferences, taking into account both the relative importance of each preference and the necessary logic relationship between them.

The calibration of the LSP Criterion Function obviously represents the most complex phase in the whole process. To calibrate a LSP Criterion Function it is necessary to take into account the needs of end users. Once the calibration of the LSP Criterion Function has finished, the evaluation of security implementation systems can start. It means that a set of values corresponding to the performance variables must be collected and provided as input to the model to obtain a global performance indicator of the degree of compliance of the security implementation system under evaluation.

4 The Evaluation Model of Cybersecurity Implementations

We present in this section part of the evaluation model created using the LSP method. This model is based on the ISO/IEC 27002, "Information technology - Security techniques - Code of practice for information security controls" [16] but, if necessary, it can be updated by the user adding, modifying or removing those requisites that are not suitable for its needs.

4.1 Cybersecurity Implementation Requirements

The LSP method prescribes as a first step to establish the requirements of the user. In our case, these requirements have been taken from the ISO/IEC 27002 standard, specifically from clauses 5 to 18 of the mentioned standard. The numbering of these clauses corresponds to the chapters in the standard; we have renumbered them starting from 1 (see Table 1).

Table 1 First level of the LSP Requirement Tree

1.	Information security policies
2.	Organization of information security
3.	Human resource security
4.	Asset management
5.	Access control
6.	Cryptography
7.	Physical and environmental security
8.	Operations security
9.	Communications security
10.	System acquisition, development and maintenance
11.	Supplier relationships
12.	Information security incident management
13.	Information security aspects of business continuity management
14.	Compliance

The ISO/IEC 27002 standard is a code of practice. It represents a good starting point for our model since is a well-proven and debated standard developed along several years by several organizations known in the field. However, the choice of

the requirements is a prerogative of the user. It has fourteen clauses, which contain thirty-five main categories and one hundred and fourteen controls. Clauses have objectives, while categories have controls as well as implementation guidance. Therefore, all have their execution and implementation defined in the standard. With these top fourteen clauses, we have built the first level of our requirement tree.

We show in Table 2 the partial requirement tree for Clause 5 "Access Control", where category 5.2.1 "User registration and de-registration" have been 'exploded' including four subcategories. Categories no further decomposed correspond to leaves in the requirement tree and they are transformed into elementary preferences by elementary criteria.

4.2 Elementary Criteria

Each control in the ISO/IEC 27002 standard has implementation guidance. We have considered control's guidance as a leaves in our requirement tree. Hence, they are performance variables in our model, as we show in Table 2.

During evaluation, performance variables are transformed into elementary preferences by means of the application of the defined elementary criteria. The definition of the correct elementary criteria is part of the work done when we develop the model. During model calibration, sometimes elementary criteria must be modified if it proves necessary to fine-tuning the model.

Table 2 Partial decomposition of Clause 5

- 5. Access Control
  - 5.1. Business requirements of access control
    - 5.1.1. Access control policy
    - 5.1.2. Access to networks and network services
  - 5.2. User access management
    - 5.2.1. User registration and de-registration
      - 5.2.1.1. Using unique IDs
      - 5.2.1.2. Disabling or removing unused IDs
      - 5.2.1.3. Identify and remove redundant IDs
      - 5.2.1.4. Redundant IDs are not reused
    - 5.2.2. User access provisioning
    - 5.2.3. Management of privileged access rights
    - 5.2.4. Management of secret authentication information of users
    - 5.2.5. Review of user access rights
    - 5.2.6. Removal or adjustment of access rights
  - 5.3. User responsibilities
    - 5.3.1. Use of secret authentication information
  - 5.4. System and application access control
    - 5.4.1. Information access restriction
    - 5.4.2. Secure log-on procedures
    - 5.4.3. Password management system
    - 5.4.4. Use of privileged utility programs
    - 5.4.5. Access control to program source code

Values assigned to each elementary performance can be assessed using appropriate metrics. They can also be obtained from experts opinions. During the process of elementary criteria definition, we have taken into consideration, among others, the ISO/IEC 27004 standard [17].

Equation (1) shows the elementary criterion *gRRIds* for guidance 5.2.1.3. "Identify and remove redundant IDs", where TUs denotes the total number of users and RIds the

number of redundant IDs found. The elementary preference is completely fulfilled when no redundant IDs are found in the system, it is no satisfied at all when the percentage of redundant IDs is equal or greater than 100 and a value between 0 and 100 is assigned in any other case.

$$g_{RRIds}(RIds, TUs) = \begin{cases} 100 \quad \text{if } RIds = 0 \\ 100 - RIds/TUs * 100 \\ \qquad \text{if } 0 < RIds/TUs * 100 \leq 100 \\ 0 \text{ if } RIds/TUs \quad * 100 > 0 \end{cases} \tag{1}$$

4.3 Building the Aggregation of Preferences

Once the requirement tree and the elementary criteria have been created, the aggregation of preferences can start. This process uses the elementary preferences corresponding to the performance variables in the requirement tree to build a new tree structure, the final aggregation structure or LSP Criterion Function.

The process starts by aggregating groups of related elementary preferences and generating, in this way, a number of aggregated preferences. Preferences are aggregated in new preferences by using Continuous Preference Logic (CPL) operators provided by the method. This bottom up process is repeated on the aggregated preferences until a single global preference is obtained. The aggregation structure created must reflect the user requirements that in this case correspond to those for the evaluation of the implementation of a security system.

If we want to aggregate n elementary preferences $E_1,...,E_n$ in a single preference E, the resulting preference E –interpreted as the degree of satisfaction of the n requirements– is expressed by a function (CPL operator) having the following properties:

1. The relative importance of each elementary preference E_i ($i = 1...n$) is expressed by a weight W_i,
2. $min(E_1, ..., E_n) \leq E \leq max(E_1, ..., E_n)$.

These operators are obtained from the instantiation of the weighted power mean:

$$E(r) = (W_1 E_1^r + W_2 E_2^r + ... + W_n E_n^r)^{1/r}, \text{ where}$$

$$0 < W_i < 100, \quad 0 \leq E_i \leq 100, \quad i = 1, ..., n, \; W_1 + ... + W_n = 1, \; -\infty \leq r \leq +\infty$$

The choice of r determines the location of $E(r)$. For $r = -\infty$ the weighted power mean reduces to the pure conjunction and for $r = +\infty$ to the pure disjunction, giving place to a Continuous Preference Logic (CPL). The range between pure conjunction and pure disjunction is usually covered by a sequence of equidistantly located CPL operators named: C, C++, C+, C+–, CA, C–+, C–, C– –, A, D– –, D–, D–+, DA, D+–, D+, D++, D. For a more detailed description of the technique see [8, 9].

The weights associated to each elementary preference are assigned by the user according to the importance that each preference has in the model being constructed. The same goes when choosing the different CPL operators. We have illustrated this point in Fig. 2 and Fig. 3 with part of the aggregation structures we have built for the requirement tree of Table 2. In these figures, circles represent CPL operators, rectangles correspond to elementary preferences and the weights are shown over the edges (arrows). Rounded rectangles in light grey do not form part of the aggregation structure, we have introduced them to indicate partial preferences corresponding to the aggregation of a set of items in the requirement tree.

Fig. 2 shows the Aggregation Structure for Category 5.2.1. "User registration and de-registration". Guidance 5.2.1.1. "Using unique IDs" and 5.2.1.2. "Disabling or removing unused IDs" have been aggregated using the strong conjunction operator CA since both are considered highly mandatory, i.e. they must be present, and their value should not be zero. Guidance 5.2.1.3. "Identify and remove redundant IDs" and 5.2.1.4. "Redundant IDs are not reused" are also considered important but not as much as the other two, so they have been aggregated with C–, a less conjunctive operator. The output preferences have been aggregated using the CA operator again.

In Fig. 3, the Aggregation Structure for Category 5.1. "Business requirements of access control" is shown. Since we have considered mandatory both variables, we have also used in this case a conjunctive GCD operator.

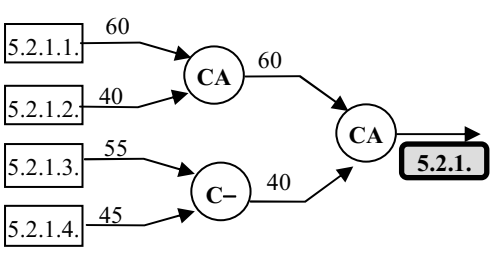


Fig. 2 Aggregation structure for Category 5.2.1. "User registration and de-registration".

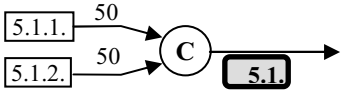


Fig. 3 Aggregation structure for Category 5.1. "Business requirements of access control".

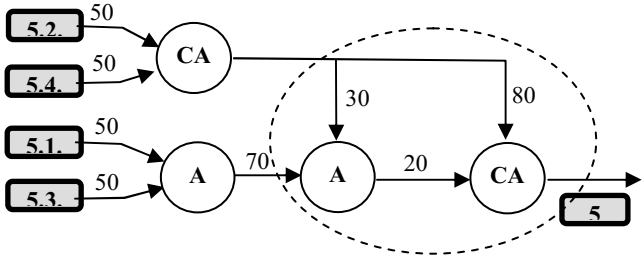


Fig. 4 Partial absorption structure for Category 5. "Access Control"

In general, security features are mandatory so most of the operators employed will be conjunctive. However, mandatory operators penalize the whole system when one of its inputs is not fulfilled. To avoid this situation, other GCD operators can be applied. For example, Fig. 4 shows the aggregation of preferences for Category 5. "Access Control" using a particular structure called partial absorption (circled with a dotted line in Fig. 4). Partial absorptions are useful when trying to join a mandatory preference m with a number of optional preferences. If the mandatory preference m is equal to zero then the result of the partial absorption is zero whatever the value of the optional preferences is. Otherwise, the output is the mean of the range $(m - \delta^-, m + \delta^+)$, where δ^- and δ^+ determine the weights to be used. Both δ^- and δ^+ are obtained from a pre calculated table (see [6] for more on this). In Fig. 4, both preferences 5.2 and 5.4 have been considered mandatory, while preferences 5.1 and 5.3 have been considered optional.

Through the presented examples of some of the model's aggregation structures, we have exposed the power and flexibility of the LSP method and its suitability for Cybersecurity evaluation. It is left to the person in charge of building the model, the responsibility of choosing the right operators, weights and to decide which requirements are mandatory, which are not and how strong the mandatory parts are.

5 Concluding Remarks

Assessing the functioning of a security system to verify how correctly implemented is and how much covers the security needs of an organization is a complex task, so it is important to have a model to assist us in this undertaking.

In this paper a model for the evaluation of cybersecurity implementation is presented. The model is based on the recommendations given by the ISO/IEC 27002 standard for security practices in Information Technology and it has been developed by following the LSP Method, a method for complex system evaluation, which is based on a Continuous Logic.

The presented model uses the above mentioned standard as a basis for a requirement tree and it allows obtaining a numerical indicator of the security requirements fulfillment.

Our model does not include cost as an item. Cost is obviously an important attribute to be well thought-out when implementing a system, however its evaluation is very complex since the different features to be considered are numerous and not trivial. Items as being taken into consideration here are not only equipment's cost, but also maintenance, amortization, storage, transportation, etc. This is another area that is being approached and where we expect to have results to show in future publications.

Since the model is based on a method that is very flexible it can be adapted to the different needs and requirements of the users of the model. This is an open area where there is room for improvement of the presented model.

References

1. Barabanov, R., Kowalski, S., Yngström, L.: Information Security Metrics: State of the Art. DSV Report series No 11 (2011). http://www.diva-portal.org/smash/get/diva2:469570/FULLTEXT01.pdf (Retrieved March 2015)
2. Bartol, N., Bates, B., Goertzel, K.M., Winograd, T.: Measuring Cyber Security and Information Assurance. State-of-the-Art Report (SOAR), Information Assurance Technology Analysis Center (IATAC), Defence Technical Information Center, Department of Defence, USA, May 8, 2009
3. Brickman, J.: Oracle Java Card Platform Implementation Completes EAL5+ Common Criteria Evaluation. Oracle Security Evaluations Blog. https://blogs.oracle.com/seceval/entry/oracle_java_card_platform_implementation (Retrieved March 2015)
4. CERT Division of the Software Engineering Institute (SEI). Carnegie Mellon University (CMU). http://www.cert.org/ (Retrieved March 2015)
5. Department of Homeland Security (DHS) Control Systems Security Program (CSSP). Cyber Security Evaluation Tool. https://ics-cert.us-cert.gov/sites/default/files/documents/DHS_CyberSecurity_CSSP-CSET-v4.pdf (Retrieved March 2015)
6. Dujmovic, J.J., Elnicki, R.: A DMS Cost/Benefit Decision Model: Mathematical Models for Data management System Evaluation, Comparison, and Selection, National Bureau of Standards, Washington, D.C., No. NBS-GCR-82-374, NTIS No. PB82-170150, 155 pages (1982)
7. Dujmovic, J.J.: A method for evaluation and selection of complex hardware and software systems. In: Proceedings of the 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG 1996, vol. 1, pp. 368–378 (1996)
8. Dujmovic, J.J.: Quantitative evaluation of software. In: Hamza, M.H. (ed.) Proceedings of the IASTED International Conference on Software Engineering, pp. 3–7. IASTED/Acta Press (1997)
9. Dujmovic, J.J., Bayucan, A.: Evaluation and comparison of windowed environments. In: Proceedings of the IASTED Interna Conference Software Engineering (SE 1997), pp. 102–105, 1997
10. Dujmovic, J.J.: Continuous Preference Logic for System Evaluation. IEEE Transactions on Fuzzy Systems **15**(6), December 2007
11. Elbirt, A.J., Paar, C.: An FPGA implementation and performance evaluation of the serpent block cipher. In: FPGA 2000 Proceedings of the 2000 ACM/SIGDA Eighth International Symposium on Field Programmable Gate Arrays, pp. 33–40. ACM, USA (2000). ISBN:1-58113-193-3. doi:10.1145/329166.3291
12. EVITA, E-safety Vehicle Intrusion Protected Applications. http://www.evita-project.org/ (Retrieved March 2015)
13. IEEE Computer Society Center for Secure Design. http://cybersecurity.ieee.org/center-for-secure-design.html (Retrieved March 2015)
14. IEEE Cybersecurity Initiative. http://cybersecurity.ieee.org/about.html (Retrieved March 2015)
15. International Information System Security Certification Consortium, Inc., $(ISC)^{2®}$. https://www.isc2.org/aboutus/default.aspx (Retrieved March 2015)
16. ISO/IEC 27002: Information technology - Security techniques - Code of practice for information security controls, 2nd edn., October 1, 2013. Reference number ISO/IEC 27002:2013(E)

17. ISO/IEC 27004: Information technology – Security techniques – Information security mangement - Measurement. Reference number ISO/IEC 27004:2009(E)
18. Jackson, W.: Under attack. GCN (2007). http://gcn.com/articles/2007/08/10/under-attack.aspx (Retrieved March 2015)
19. LeMay, E., Ford, M.D., Keefe, K., Sanders, W.H., Muehrcke, C.: Model-based security metrics using ADversary VIew Security Evaluation (ADVISE). In: Eighth International Conference on Quantitative Evaluation of Systems (QEST). IEEE, September 5–8, 2011. Print ISBN: 978-1-4577-0973-9
20. National Institute of Standards and Technology (NIST), Security Content Automation Protocol (SCAP). http://scap.nist.gov/index.html (Retrieved March 2015)
21. Nicol, D.M., Sanders, W.H., Trivedi, K.S.: Model-Based Evaluation- From Dependability to Security. IEEE Transactions on Dependable and Secure Computing **1**(1), 48–65 (2004). doi:10.1109/TDSC.2004.11. ISSN: 1545-5971
22. Oracle Security Evaluations Blog. https://blogs.oracle.com/seceval/ (Retrieved March 2015)
23. Payment Card Industry Security Standards Council. https://www.pcisecuritystandards.org/security_standards/role_of_pci_council.php (Retrieved March 2015)
24. Platt, J.R.: Raising the Bar for Cyber security Specialists. The Institute IEEE **39**(1), March 2015
25. Rozenfeld, M.: IEEE Standards on Cybersecurity. The Institute IEEE **39**(1), March 2015
26. Spidalieri, F., Kern S.: Professionalizing Cybersecurity: a path to universal standards and status. Salve Regina University. Newport, Rhode Island, USA. https://www.salve.edu/sites/default/files/filesfield/documents/Professionalizing-Cybersecurity.pdf (Retrieved March 2015)
27. The Common Criteria. http://www.commoncriteriaportal.org/ (Retrieved March 2015)
28. The National Information Assurance Partnership/Common Criteria Evaluation and Validation Scheme (NIAP/CCEVS) https://www.niap-ccevs.org/ (Retrieved March 2015)
29. Wang, A.J.A.: Information security models and metrics. In: ACM-SE 43 Proceedings of the 43rd Annual Southeast Regional Conference, vol. 2, pp. 178–184. ACM, USA (2005). ISBN:1-59593-059-0
30. Weiß, S., Weissmann, O., Dressler, F.: A Comprehensive and Comparative Metric for Information Security. http://www.ccs-labs.org/bib/weiss2005comprehensive/weiss2005comprehensive.pdf
31. Wikipedia. http://en.wikipedia.org/wiki/Common_Criteria#Criticisms (Retrieved March 2015)
32. Wolf, M., Gendrullis, T.: Design, implementation, and evaluation of a vehicular hardware security module. In: Information Security and Cryptology - ICISC 2011. 14th International Conference, Seoul, Korea, November 30–December 2, 2011. Lecture Notes in Computer Science, vol. 7259 (2012). Revised Selected Papers. ISBN: 978-3-642-31911-2 (Print) 978-3-642-31912-9 (Online)

Accelerating the Critical Line Algorithm for Portfolio Optimization Using GPUs

Raja H. Singh, Lee Barford and Frederick Harris Jr.

Abstract Since the introduction of the Modern Portfolio Theory by Markowitz in the Journal of Finance in 1952, it has been the underlying theory in several portfolio optimization techniques. With the advancement of computers, most portfolio optimization are done by CPUs. Over the years, there have been papers that introduce various optimization methods including those introduced by Markowitz, implemented on the CPUs. In the recent years, GPUs have taken the front seat as a technology to take computational speeds to new levels. However, very few papers published about portfolio optimization utilize GPUs. Our paper is about accelerating the open source Critical Line Algorithm for portfolio optimization by using GPU's, more precicely using NVIDIA'S GPUS and CUDA API, for time consuming parts of the algorithm.

Keywords Cuda · Markowitz · Stocks · Parallel computing · Pycuda

1 Introduction

Since its introduction in 1952 by Harry Markowitz, the Modern Portfolio Theory has become a solid foundation for the creation of various methods of portfolio optimization. The paper he wrote [1] provided mathematical model which took into account expected return and risk when constructing an optimal portfolio [2]. Other methods heavily used since the departure from Markowitz model such as "market-cap weighting" and "fundamental weighting" don't explicitly make assumptions about the risk and return parameters like the one used in Markowitz's theory. However, there has

R.H. Singh · L. Barford · F. Harris, Jr.(✉)
Department of Computer Science and Engineering, University of Nevada, Reno, USA
e-mail: rajas@nevada.unr.edu, lee_barford@ieee.org, Fred.Harris@cse.unr.edu

L. Barford
Keysight Laboratories, Keysight Technologies, Santa Clara, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_29

been a shift from market-cap weighted indexing to other schemes that take market anomalies into consideration rather than the "standard investment theory" that are very similar to the one Markowitz proposed; These modern methods tend to focus on minimizing the variance of the portfolio and focusing on portfolio diversification measures [2]. Since the model's introduction in 1952, the equity indexes have come full circle according to Kaplan [2]. This shift might have to do with the market slumps in the late 2000's or just be the cyclical nature of the market.

For this paper we utilized the open source (critical line) algorithm for portfolio optimization presented by Bailey and Prado in their paper [5]. The main reason for utilizing this algorithm besides the fact that it is open source was that it computed the same optimized portfolio as a VBA algorithm with a faster computation time [5] and it can handle any number of assets restricted only by the hardware limitations. The focus of this paper is to parallelize the bottle neck of the sequential algorithm to reduce the computation time for deriving an optimized portfolio.

2 Related Works

With the advancement in computing capabilities, there has been an increase in utilizing computers for optimization algorithms. However, though there is an abundance of literature highlighting various methods of optimization, very few if any provide an open source implementation of such algorithms. Most methods used by non-professionals are either too slow or rely too much on VBA/Excel for portfolio optimization. Though there are some open source implementations and guides on using VBA such as [4], most examples utilize a small set of assets and as the size of the co-variance matrix starts increasing dramatically, excel starts having issues with the size of the data.

We found two papers that address portfolio optimization on the GPUs [6] and [7]. The [6] written by Stchedroff used the Multi Directional optimization which is an example of the Direct Search method, to find an optimal portfolio from an asset population. The sequential code was executed on an unspecified CPU using multiple threads and the parallel code was written on an unspecified GPU. Stchedroff did mention using CUBLAS, BIAS and BLAS libraries for the matrix computations. According to his results he achieved a 10x speed up on the GPU vs CPU and a 60x speed up for the combined GPU and new algorithm approach. However the author didn't provide any psuedo code for the sequential or parallel implementations, and didn't specify any information on the CPU and GPU hardware.

The [7] written by Hu used simulated anealing to derive the optimal portfolio from randomly generated portfolios from asset population of twenty. He provided one psuedo algorithm for his portfolio selection process but didn't provide any psuedo code for the Cuda kernels. He mentioned a 4X speed up using two Nvidia GTX295 cards vs the sequential code executed on Intel i7 Quadcore. Besides these two papers we haven't found other papers that implemented any portfolio optimization algorithms on the GPU.

3 Portfolio Optimization

Concept: A vital assumption Markowitz's model makes is that "the investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing [1]." This concept is referred to the "expected returns-variance of returns rule." The idea is that for a given risk there should not be another portfolio that yields a greater or equal return with a risk that is less than or equal to the optimal portfolio. Another way Markowitz stated this is that for a given return there should not be a risk that is less than the risk of the optimal portfolio.

Model: Markowitz's model is based on portfolio returns, portfolio variances and the weights assigned to the assets included in the portfolio. We therefore summarize the ideas behind these components below.

The sum of the weights assigned to each asset have to sum up to one $\sum_{i=1}^{N} X_i = 1$. The Expected Return of the whole portfolio is expressed as: $E = \sum_{i=1}^{N} X_i \mu_i$. Where X_i is treated as a random variable representing the weight assigned to an asset and μ_i represents the expected return of an asset.

The risk in the model is expressed as the variance of the portfolio. However to understand the variance we first need to understand the co-variance between two assets. The co-variance between two assets is the expected value of the product of the deviations of the two assets from their mean. This equation can be represented as: $\sigma_{ij} = E[R_i - E(R_i)][R_j - E(R_j)]$, where $\sigma\sigma_{ij}$ is the co-variance between asset i and j.

We can use the fact that the variance of asset i is the co-variance of σ_{ii}. We can express the variance of the whole portfolio as:

$$V = \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} X_i X_j \tag{1}$$

From the co-variance matrix we are able to compute the correlation between the assets. An ideal portfolio usually contains assets with minimal correlation between each other [9] and where the variance of the assets with respect to other assets in the portfolio is more important than a particular asset's own variance alone [8].

4 Critical Line Algorithm

Markowitz's theory focuses on deriving an efficient portfolio that yields the maximum return for a minimum risk (or volatility). When several efficient portfolios are computed, they make up the Efficient Frontier. There can be no portfolios that exceed the frontier, the optimal portfolio (max return or min variance) will be on the frontier and all other portfolios will be inside the boundary of the curve. The portfolios on

the Efficient Frontier will dominate the rest of the portfolios [9]. Harry Markowitz referred to the the method of deriving the entire efficient frontier as the "Critical Line" Method. It is also known as Markowitz's Efficient Frontier. However we will call it Critical Line because it is referred to by that name in [1] and [5].

This method has changed over time from minor to major modifications such as including negative weight for short sales resulting in the problem taking on the form of an unconstrained problem. However, the algorithm provided by Bailey and Prada doesn't allow negative weights (associated with short sales), keeping the method as a constrained problem with inequality conditions and equality conditions as specified in their paper [5]. According to Bailey and Prada, there isn't an analytic solution to the optimal portfolio problem so it must be solved as an optimization problem.

Several works such as [9] suggest solving the efficient frontier using a gradient based approach. However, the approach is very slow and can lead to local maximal and minimal based on the starting position and the seed vector. Bailey and Prado point out that the gradient based optimization approach "require a separate run for each portfolio" in the efficient frontier; this might explain the slow computation for gradient based optimization algorithms. They continued to state that the Critical Line Algorithm is specifically designed for the inequality constrained optimization problem and it guarantees an exact solution after a set number of iterations. The Critical Line Algorithm also derives the whole frontier of efficient portfolios in one run vs multiple runs for each portfolio as experienced by gradient based approaches.

Various Approaches: Some literature points towards setting up the problem as a convex function and then using quadratic programming to solve the problem with linear constraints [9] or non-linear programming for convex functions with non-linear constraints. While others use heuristic approaches such as firefly algorithm [10] or genetic algorithms. Mixed integer programming is another approach which came up when we searched for various optimization methods of the portfolio.

Optimization Problem: The solution set is found by setting up the problem as a quadratic problem subject to linear constrains in inequalities and one linear constraint inequality [5]. In order to understand the model, it is first important to know the variables Bailey and Prado used for the model in [5]:

- N = 1,2,3 contain the indices for the asset universe
- ω is the row vector of the weights assigned to the assets (a optimizer variable that needs to be solved)
- l is the row vector with the lower bound for the weights $\omega_i \geq l_i$, for all i in N
- u is the row vector with the upper bound for the weights $\omega_i \leq U_i$, for all i in N
- F $\subseteq$ N represents the "assets that don't lie on their boundaries" where $l_i \leq \omega_i \leq \mu_i$
- B $\subseteq$ N represents the "assets that lie on their boundaries" where $1 \leq k \leq n$ and B $\cup$ F = N
- the covariance matrix $\sum$, row vector ω, and the mean vector μ are modeled below:

$$\sum = \begin{bmatrix} \sum_F & \sum_{FB} \\ \sum_{BF} & \sum_B \end{bmatrix}, \mu = \begin{bmatrix} \mu_F \\ \mu_B \end{bmatrix}, \omega = \begin{bmatrix} \omega_F \\ \omega_B \end{bmatrix}$$

where $\sum_F$ is the covariance matrix for Free assets, $\sum_{FB}$ is the covariance between free assets F and assets lying on the boundary B, $\sum BF$ is the transpose of the $\sum_{FB}$, μ_F is the row vector for the means of the free weights, μB is the row vector for the means for the assets lying on the boundary, ω_F is the row vector for the weights assigned to the free assets and ω_B is the row vector for the weights assigned to the assets lying on the boundary.

The solution to the problem is obtained by minimizing the Lagrange Function with respect to the weights in the row vector ω and its multipliers γ and λ. The Lagrange Function is shown in Equation 2

$$L[\omega, \gamma, \lambda] = \frac{1}{2}\omega' \sum \omega - \gamma\left(\gamma' 1_n - 1\right) - \lambda\left(\omega'\mu - \mu_p\right) \tag{2}$$

Bailey and Prado used a modified version of the function in Equation . They stated due to the conditional inequalities they would not be able to solve for the variables or would be required to use the Karush-Kuhn-Tucker conditions. The Karush-Kuhn-Tucker conditions are used in mathematical optimization to set up first order conditions so that the solution is optimal in non-linear programming [11]. Instead they used the divide and conquer method where they converted the constrained problem into an unconstrained problem by using the "two mutual funds theorem." They modified the Lagrange Multiplier shown in Equation 2 to the format shown in Equation 3, changing the constrained problem to an unconstrained problem and computed the efficient frontier by deriving a convex combination between any two neighboring turning points. Due to the length of the algorithm we have not provided it in the paper, please refer to [5] for the complete algorithm.

$$\begin{aligned} L\left[\omega, \gamma, \lambda\right] = \frac{1}{2}\omega_F' \sum_F \omega_F + \frac{1}{2}\omega_F' \sum_{FB} \omega_B \\ + \frac{1}{2}\omega_B' \sum_{BF} \omega_F \frac{1}{2}\omega_B' \sum_B \omega_B - \gamma(\omega_F' 1_k + \omega_B' 1_{n-k} - 1) \\ -\lambda\left(\omega_F'\mu_F + \omega_B'\mu_B - \mu_p\right) \end{aligned} \tag{3}$$

5 Hardware

The algorithms were executed on the CUBIX by the CUBIX Corporation. The box has 8 NVIDIA GTX-780 cards, 2 Intel Xeon E5 2620s and 65.9 Gigabytes of Ram. We didn't use multiple GPUs.

6 Data Collection

The data gathered for the trials was in the same format as specified in their paper [5]. The assets were collected from Yahoo and contained daily returns for each asset for a little over two years. We only used stocks as the assets for the portfolio optimization to keep the data gathering process simple. We originally tried to start off at one hundred assets and increment the number of assets by fifty up to five-hundred assets. We dropped any asset that had missing value(s) for daily returns. This resulted in the number of assets per file not being increments of 50 and the asset universe varying in size based on the number of assets that were dropped.

The end results was we had files with the following number of assets: 86, 118, 139, 160, 174, 257, 310, 346, 390, 463, and 501. This allowed us to test the Critical Line Algorithm on various asset sizes. We also used the 10 asset file used by Prado and Bailey to do their own testing, this allowed us to confirm that the parallel algorithm didn't provide erroneous results. Once data was collected, we computed the average daily return for the each asset and computed the co-variance matrix for the asset set.

7 Sequential Execution

The sequential code was slightly modified. Bailey and Prado wrote all the methods within a class. However due to minimal amount of comments and nature of Python language we decided to bring all the functions outside of the class so that we could modify and test each function as we needed. We also modified the main driver slightly so that we could profile and test the code using various asset numbers. However, all the modifications we made left the computations and the control flow intact.

The code was profiled to both time the program completion and determine which part(s) of the program contribute the most to the completion time. This allowed for a more precise target to parallelize using CUDA. After profiling the code it was determined that the majority of the time was spent in the reduce matrix function. The purpose of the function is to derive a sub matrix for a set of assets from a larger matrix.

The getMatrices function calls the reduceMatrix function. The reduceMatrix function contains two for loops, with the first for loop extracting specified columns and appending them to an intermediate matrix. The second for loop extracts specified rows from the intermediate matrix and appends them to the return matrix. The loops rely heavily on the numpy.append function which calls the numpy.concatenate function. This function joins two arrays together. However, to accomplish this task it requires creating a new array large enough to hold all the values. With small size arrays this task is very fast and the overhead associated with it is trivial, but as the size of the matrices and the number of calls to the function increase, the overhead associated with this task takes a toll on the overall completion time. This was made evident by the profiling of the code and led us to parallelize the reduceMatrix function.

8 Parallel Execution

Memory Transfer: The most significant overhead associated with parallel computing (GPU specific) comes from the transfer of data from the host to device and back again. In order to minimize the impact of this overhead we created two large matrices when we first initialized the CLA class. The first matrix contained the covariance matrix for the complete asset population in the device memory (deviceCovar matrix) and the second matrix (intermMatrix) was a square matrix based on the asset population size was used to extract and hold all the necessary columns from the deviceCovar matrix; it acted as the temporary matrix that held the specified columns of the deviceCovar matrix. From the intermMatrix we extracted the specified rows for the reduced returnMatrix. By allocating a large bank of memory and then transferring the needed data to the device once during the initialization phase we reduced overhead associated with memory allocation and data transfer significantly. Once the data and the needed memory were on the device, we could utilize the two matrices repeatedly without incurring the overhead of transferring data from the host to the device repeatedly. However, we did have to transfer the reduced matrix from the device to the host at the end of the reduceMatrix function.

reduceMatrix Function: We created a new reduceMatrix function we named gpuReduceMatrix because we utilized the original reduceMatrix function for reduction of the row vectors. As stated previously, with a small asset population the overhead associated with the reduceMatrix function is very trivial. Keeping that in mind we created a branching instruction where we specified a threshold value that had to be met before we utilized the GPU version of the code else we executed the sequential code written by the Bailey and Prado. The threshold referred to the size of the deviceCovar matrix. If we set the threshold to zero then every time the gpuReduceMatrix was called it executed the parallel code. On the other hand if we specified a threshold then the matrices to be reduced had to meet the specified threshold size before the parallel code was executed. If the threshold was not met then the gpuReduceMatrix function used the same code as the sequential version. In the parallel code section of the gpuReduceMatrix function, we wrote two kernels that replaced the for loops that were executed in the sequential code. These two kernels are discussed later.

Cuda Kernels: To parallelize the reduceMatrix function we wrote two cuda kernels and called them in the gpuReduceMatrix. The first one (columnAppend) appended the specified columns from the deviceCovar matrix to the intermMatrix. The second one (rowAppend) appended the specified rows from the intermMatrix to the gpuReturnMatrix. Inside the columnAppend kernel each block mapped to a column of the deviceCovar matrix based on its blockId and each block thread appended an element of that column to the intermMatrix. Inside the rowAppend kernel each block mapped to a specified row of the intermMatrix and each thread copied an element of a row to the gpuReturnMatrix.

Since the max size of the asset population was less than the max number of blocks the GPU card allowed us to launch, we didn't stride the blocks for columns or rows. Meaning each block only mapped one column in the columnAppend kernal or row in the rowAppend kernel. However, we did stride the the threads so each thread would be able to append more than one element in each kernel if the number of elements in the column or row exceeded the max number of threads we could launch. Instead of iterating through a loop to append a single column/row per iteration we were able to append all the columns simultaneously and then all the rows simultaneously by launching the appropriate kernels with specified number of blocks and threads. Algorithm 1 and Algorithm 2 show the code for the cuda kernels.

Algorithm 1. columnAppend Kernel

```
function APPENDCOLUMN(dest, src, posList, numColsAppendMatrix, padding, size,
realColumns)
    tid ← threadIdx.x
    appendPos ← blockIdx.x
    extractionPosDataMatrix ← posList[blockIdx.x]
    numColsDataMatrix ← realColumns
    for mimickedIndex < size do
        globalTid += tid × numColsDataMatrix + extractionPosDataMatrix
        stride ← blockDim.x×i + tid
        index ← stride × (numColsAppendMatrix + padding) + appendPos
        rowPos ← (index ÷ realColumns)
        mimickedIndex ← index  (rowPos × padding)
        if mimickedIndex ≥ size then
            break
        end if
        dest[index] ← src[globalTid]
    end for
end function
```

getMatrices Function: After we parallelized the reduceMatrix function we determined we had to modify the getMatrices function. Each instance of the getMatrices calls the reduceMatrix four times: twice to reduce two NXN matrices (both covariance matrices) and twice to reduce two row vectors (one for average returns of the assets, and the other for the weights of those assets). We determined through trial and error that parallelizing the calls for the row vectors was not efficient and led to an increase in computational time. We modified the getMatrices function to call the GPU version of the reduceMatrix function (gpuReduceMatrix) only for NXN square matrices and keep using the original sequential version of the reduceMatrix for the row vector reductions.

9 Results and Analysis

All of the GPU implementations had a speed up over the sequential trial after a given threshold. It makes sense that once the number of assets grow, the computations

Algorithm 2. rowAppend Kernel

```
function ROWAPPEND(destMatrix, srcMatrix, rowList, numCols, padding, sizeOfInputMatrix)
    tid ← threadIdx.x
    rowPosDest ← blockIdx.x
    rowPosSrc ← rowList[blockIdxX]
    destIndex ← rowPosDest × numCols+tid
    srcIndex ← rowPosSrc × (numCols+padding)+tid
    while iterator < numCols do
        if iterator > numCols then
            break
        end if
        matrix[destIndex] ← inputMatrix[srcIndex]
        srcIndex += stride
        destIndex += stride
        iterator += stride
    end while
end function
```

should be shifted over to the GPU. However, one of the interesting procedures we tried was a hybrid approach where prior to a certain threshold, the CPU would run the computations after which threshold the GPU would take over. This threshold was the number of elements in the co-variance matrix (rows × columns). After trying different thresholds we were able to see that the optimal CPU/GPU combination was at fifty elements in the co-variance matrix. That is when the number of elements in the matrix was below fifty the CPU reduced the matrix otherwise the GPU reduced the matrix for the co-variance matrices.

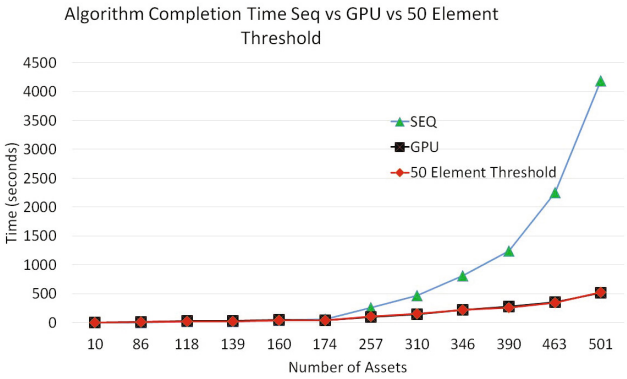


Fig. 1 Completion Time(s) for Sequential, Parallel and 50 Element Threshold

Graph in Figure 1 shows the completion times for the sequential vs parallel vs the threshold method. The 50 threshold method runs slightly faster than the GPU reduceMatrix however after 257 assets, the GPU version runs slightly faster. It is easier

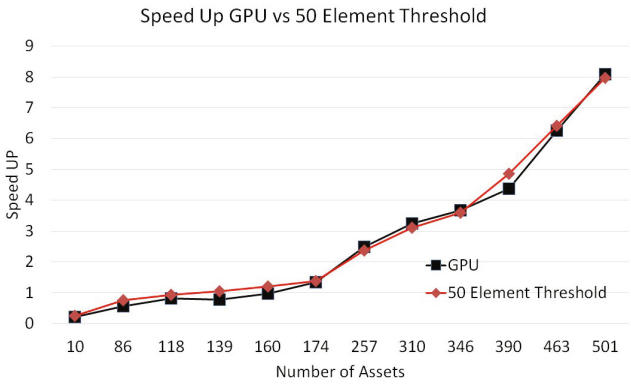


Fig. 2 Speed Up Parallel vs 50 Element Threshold

to see speed up between the reduceMatrix running on the GPU for all covariance matrix reductions and the threshold approach in the graph in Figure 2. The sequential implementation runs at a much faster time for a small number of assets, however once the number of assets exceeds 139, the 50 element threshold method starts getting speed up with the GPU version attaining speed up at 174 assets and above. However, please note at around 500 assets, the GPU method has a slightly higher speed up than the threshold method. This was tested multiple times, and it seems that after 500 assets, when the GPU does all the reductions for the covariance matrix, it will be faster than both the sequential and the 50 element threshold method.

10 Conclusion and Future Work

In this paper we used GPUs to speed up the Critical Line Algorithm to optimize portfolios using Markowitz's Efficient Frontier presented by Bailey and Prado. Prior to parallelizing the bottleneck of the Critical Line Algorithm on the GPU, we did have an inkling that there would be speed up. We were not sure about the extent of speed up. However after the results, we were able to determine that the speed up tends to grow in relation to the number of assets. The most speed up we were able to attain with an asset population of 501 asset was around 8X.

However, we believe as the size of the asset population increases, the speed up we attain in relation to the sequential completion time will grow drastically. The method in which the Critical Line Algorithm has been optimized, it can handle 65535 assets on one GPU before it will require additional GPUs for the matrix reduction. The purpose of this paper was to speed up the Critical Line Algorithm presented in [5] and to hopefully encourage others into utilizing the power of GPUs for computational finance and publishing the results.

There are several modifications that can be made to the Critical Line Algorithm, such as including negative weights for short sales. The speed up achieved was the result of parallelization of a small part of the code. However, as the number of assets increase, we believe the other parts of the algorithm will start adding to the time and could be viable for parallelization.

Acknowledgment We would like to thank D.H. Bailey and M. Lopez de Prado for releasing their algorithm. This material is based in part upon work supported by: The National Science Foundation under grant number(s) IIA-1329469, and by Cubix Corporation through use of their PCIe slot expansion hardware solutions and HostEngine. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or Cubix Corporation.

References

1. Markowitz, H.: Portfolio Selection. The Journal of Finance **7**(1), 77 (1952). JStor. Web, March 28, 2015
2. Kaplan, P.: Back to Markowitz. Morningstar (2014). http://corporate.morningstar.com/US/documents/Indexes/Back-To-Markowitz-2014.pdf (accessed: May 5, 2015)
3. Balch, T., Romero, P.: What Hedge Funds Really Do: An Introduction to Portfolio Management
4. Kwan, C.C.: A Simple Spreadsheet-Based Exposition of the Markowitz Critical Line Method for Portfolio Selection. Spreadsheets in Education (eJSiE) **2**(3), Article 2 (2007)
5. Bailey, D.H., Lopez de Prado, M.: An Open-Source Implementation of the Critical-Line Algorithm for Portfolio Optimization. Algorithms **6**, 169–196 (2013)
6. Stchedroff, N.: Portfolio Optimization. Wilmott **2013**, 52–57 (2013). doi:10.1002/wilm.10184
7. Hu, J.: Stock Portfolio Optimization Using CUDA GPU (unpublished)
8. Rubinstein, M.: Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. The Journal of Finance **57**, 1041–1045 (2002). doi:10.1111/1540-6261.00453
9. Hilpisch, Y.: Python for Finance: Analyze Big Financial Data. O'Reilly Media, Beijing (2014)
10. Bacanin, N., Tuba, M.: Upgraded firefly algorithm for portfolio optimization problem. In: UKSim-AMSS 16th International Conference on Computer Modelling and Simulation (2014)
11. Cottle, R.: William Karush and the KKT Theorem. Documenta Mathematica, Extra Vol., 255–269 (2010)
12. Kloeckner, A.: Welcome to PyCUDA's Documentation! — PyCUDA 2014.1 Documentation (2015). http://documen.tician.de/pycuda/ (accessed: March 28, 2015)

Maximum Clique Solver Using Bitsets on GPUs

Matthew VanCompernolle, Lee Barford and Frederick Harris Jr.

Abstract Finding the maximum clique in a graph is useful for solving problems in many real world applications. However the problem is classified as NP-hard, thus making it very difficult to solve for large and dense graphs. This paper presents one of the only exact maximum clique solvers that takes advantage of the parallelism of Graphical Processing Units (GPUs). The algorithm makes use of bitsets to reduce the amount of storage space needed and take advantage of bit-level parallelism in hardware to increase performance. The results show that the GPU implementation of the algorithm performs better than the corresponding sequential algorithm in almost all cases; performance gains tend to be more prominent on larger graph sizes that can be solved using more levels of parallelism.

1 Introduction

A clique, or complete sub-graph, is a subset of the vertices in an undirected graph, such that there exists and edge between every pair of vertices in the subset. The Maximum Clique Problem (MCP) is one of the most studied NP-hard problems which consists of finding the largest possible clique in a graph. For any large or dense graph, MCP often cannot be solved in a reasonable amount of time even on high end machines due to its exponential time complexity. Despite the difficulty of the problem, finding a maximum clique has many useful applications in areas such as bioinformatics, computer vision, robotics, patterns in telecommunications traffic, and more [2, 11]. For this reason and the fact that many real-world graphs

M. VanCompernolle · L. Barford · F. Harris Jr.(✉)
Department of Computer Science and Engineering, University of Nevada, Reno, USA
e-mail: mvancomp@nevada.unr.edu, lee_barford@ieee.org, Fred.Harris@cse.unr.edu

L. Barford
Keysight Laboratories, Keysight Technologies, Santa Clara, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_30

do not elicit worst-case behaviour for MCP, new algorithms and optimizations to the problem continue to be a popular field of research [8].

One emerging trend in research is the use of Graphical Processing Units (GPUs) for general purpose computing to develop high performance applications. The release of NVIDIA's CUDA, a general purpose parallel computing platform released in 2006, enabled programmers to finally take advantage of the parallelism and high computational power of NVIDIA's GPUs while enabling a low learning curve for developers familiar with standard programming languages [5].

This paper describes the implementation of a maximum clique algorithm on GPUs that utilizes bitset representations to reduce memory requirements and increase performance. The implementation does many things that most maximum clique solvers to not attempt, despite the popularity of the field of research. The algorithm presented is one of very few maximum clique solvers that runs on GPUs, makes use of recursion on the GPU, and supports systems with multiple GPUs.

The rest of the paper is structure as follows: Section 2 covers background information necessary to better understand the proposed algorithm and summarizes related maximum clique algorithms. Section 3 explains the new maximum clique algorithm in detail. The results, comparisons between the sequential and parallel GPU algorithm, and an analysis of the results are discussed in section 4. Finally, conclusions and possible future improvements are stated in Section 5.

2 Background

Basic Maximum Clique Algorithm: A basic method for searching for a maximum clique in a graph $G = (V, E)$, where V is a finite set of vertices and E is a finite set of edges that are unordered pairs (u, v) where $u, v \in V$, is a branch and bound, depth first search method. The basic implementation explained in [14] uses two global sets: one to store the current clique Q and another to store the largest clique found so far, Q_{max}. A candidate set R is also stored throughout the search to keep track of which vertices may be added to Q. When the algorithm starts, Q is empty and R contains all vertices in the graph, $R = V$. The search process beings by selected a vertex p from R and adding the vertex to Q, forming a new current clique, and then proceeds to calculate a new candidate set R that is the intersection of the current R and the vertices adjacent to p. This process occurs recursively until R is empty, which means that Q is a maximal clique and cannot grow any larger, or $|Q| + |R| <= |Q_{max}|$ which means that the Q cannot grow larger than Q_{max}. When Q is determined to be maximal, it is compared against Q_{max}, and if $|Q| > |Q_{max}|$, $Q_{max} = Q$. The search then backtracks in order to find other maximal cliques, removing p from both Q and R and then selected a new p from R and repeating the process until the search space has been exhausted.

Approximate Coloring: The basic MCP algorithm is much too slow to find the maximum clique for large or dense graphs in a reasonable amount of time. One way to increase the performance of the algorithm is to further prune the search space in order to avoid unnecessary searching. One method of doing so is to use approximate coloring of the graph vertices, such as the ones used in the MCQ, MCR, MCS, and BBMC algorithms [6, 11, 14, 15].

Bitset Representations and BBMC: Another alteration to the basic MCP implementation is the use of bitsets to represent which vertices are in a set. The BBMC algorithm along with its many variations such as BBMCI, BBMCR, BBMCL, and BBMCS all use bitsets to represent whether a vertex from the graph is in a set or not [9–11, 13]. The use of bitsets allows the algorithms to take advantage of bit-level parallelism in hardware where a single computer instruction can operate on a number of bits the size of a word on a processor at a time (usually 32 or 64 bits in a word). Algorithms using bitsets therefore have two main advantages: the use of a bitset instead of typical storage implementations reduces the memory needed by a factor equal to the size of a word on the system, and common operations can be done in parallel using bit-wise operations.

Parallel MCP Algorithms: A significant way to speed up the search for the maximum clique in a graph is to take advantage of any available parallelism in hardware. The majority of parallel implementations for MCP have been written to take advantage of multiple CPU cores. The state-of-the-art parallel exact maximum clique finder in [7] was designed for large sparse graphs to exploit characteristics of social and informational networks and is parallelized on a shared memory system, but could also be used on a distributed memory architecture. The implementation achieved approximately linear speedup with respect to the number of processors used, while sometimes achieving super linear speedup due to increased pruning caused from workers sharing upper bounds with each other.

A very similar parallel approach was proposed in [4] that searches for the maximum clique in a graph using multiple threads. The implementation is based off of the sequential BBMC algorithm that uses bitsets and is implemented in C++ using C++11 threads and shared memory parallelism.

There are also algorithms that attempt to utilize the GPU to solve MCP, though there are not every many. One implementation used a neural network and CUDA to search for the maximum clique in a graph using a GPU [1]. The algorithm was limited, only being able to handle graphs with 800 vertices maximum due to memory constraints.

3 Parallel GPU Implementation

Parallel Method: The BBMC implementation on the GPU, which will be referred to as BBMCG, has two distinct dimensions of parallelism. The first dimension uses block-wise parallelism to divide the search space of the graph into subtrees in order to traverse the search space in parallel. This method was inspired by the work queue distribution method proposed in [4], where a work queue is initially populated by splitting the search space immediately below the root. In BBMCG, each block is initialized with a current clique Q that contains one vertex from the graph, indicated by a 1 at the position for the vertex in the bitset. The corresponding initial candidate set R for each block is a set containing all the adjacent vertices to the vertex in Q. BBMCG therefore creates a number of blocks for the algorithm that is equal to the number of vertices in the graph. Each block then traverses its own search tree recursively utilizing recursion on the GPU based on the BBMC algorithm described in [6, 11, 12]. A single global variable is shared between the blocks to indicate the current maximum clique size found. If a block finds a clique in its search space larger than the current maximum, it atomically updates the global maximum using CUDA's built in atomicMax function and stores the maximum clique's bitset in a designated space for the block in global memory. By traversing the search space in parallel and communicating the largest clique found, more of the search space is potentially pruned as larger cliques are found faster.

BBMCG's second dimension of parallelism is thread-wise parallel computations for bit-wise operations. CUDA lacks a bitset class, so bitsets are implemented using arrays of unsigned integers. In a bitset, a single bit is needed for each vertex in the graph. The number of unsigned integers needed to represent a bitset is calculated by dividing the number of vertices n in the graph by the number of bits in an unsigned integer in CUDA and rounding up. Since unsigned integers are typically 32 bits long, one unsigned integer is needed to represent 32 vertices in the graph. The adjacency matrix which requires a bitset for each vertex in the graph could then be represented as a 2-D array of unsigned integers, but CUDA generally prefers working with contiguous 1-D arrays. The adjacency and inverse adjacency matrices are stored as 1-D arrays of unsigned integer mappings of the 2-D array equivalents.

Any bit-wise operation that is not sequential in nature is parallelized by performing the operations using multiple threads per block. The number of threads used per block is equivalent to the number of unsigned integers needed to represent a bitset for the graph. BBMCG maps a single thread to each unsigned integer of a bitset in order for the threads to perform bit-wise operations on subsections of the entire bitset in parallel. In the base BBMC algorithm, the majority of computations are bit-wise operations on bitsets such as finding the first set bit, getting the number of set bits, intersecting bitsets, copying bitsets, and setting or clearing a bit in a bitset. All of these operations besides setting and clearing a bit can be executed in parallel to increase performance. For example, to calculate the number of bits that are 1 in a bitset, each thread executes the function $\_\_popc$ on its corresponding unsigned integer from the bitset, which is a built in CUDA function that counts the number of bits set to 1 in

an unsigned int. Results of each thread are then stored in array in shared memory, and the threads calculate the sum of all the set bits using an efficient parallel vector reduction algorithm on the array.

Memory Allocation and Management: Although the BBMCG is able to run on graph sizes larger than those in [1], due to the use of bitsets, some memory limitations were still encountered. One of the primary memory consumers in BBMCG is the storage space needed to store the run-time stack for recursion. Each recursive call in the search needs memory to store local variables, arrays to store the coloring of vertices to improve pruning, and several bitsets to store the current clique and candidate set. The maximum number of recursive calls directly corresponds to the size of the maximum clique in the graph, so the algorithm needs to allocate enough memory to reach the maximum depth to find the maximum clique in worst case scenarios. The amount of memory needed is then multiplied by the number of blocks used, which all make their own recursive calls as they search. Faster forms of memory, such as local registers and shared memory (limited to 48 KB on modern cards) are not large enough to store data at each level of recursion for all blocks. Even constant memory memory is limited to 64 KB and is too small to store the adjacency matrices as suggested in [1]. A single adjacency matrix for a graph size of 700 is approximately the maximum size that can be stored in constant memory, which would severely limit the problems BBMCG can run on.

BBMCG makes heavy use of global memory to address the memory limitations of shared and constant memory, but still uses registers to store non-array based data. Although global memory is much slower than shared and constant memory, it is necessary in order to run the algorithm on reasonable sizes. Global memory needed for the for the entire problem is pre-allocated before launching the kernel. The memory is allocated as several 1-D arrays that each store a single type of data for all of the blocks in the Kernel. Each block calculates the starting position in the 1-D array for its portion of memory and points to it using a pointer. Each block's section of memory in the array is then broken down further into subsections for levels of recursion where each subsection is big enough to store the associated data for a single level. When a block makes a recursive call, a new pointer is created that is offset the size of one subsection of data from the current pointer position, pointing to a new subsection of space reserved to store data for the new level of recursion. The 1-D array is created to be large enough so that each block has enough memory to store data for each level of recursion all the way down to the maximum calculated level. One disadvantage is that the amount of memory allocated may be much larger than what is actually used, because memory is allocated for a number of recursive calls equal to the number of vertices in the graph for each block. The smaller the maximum clique is in the graph, the less reserved memory is actually used. CUDA supports dynamic memory allocation which would only use the amount of memory that is needed by the algorithm, but it is extremely slow and not worth the storage benefits.

Preprocessing and Post-processing: Several preprocessing steps need to happen before the BBMCG kernel is launched on the GPU. First the vertices of the graph need to be initially ordered based on the desired ordering method. The initial search space subtrees for each block must be generated from the graph as a pair of bitsets holding a current clique and candidate set. On the CPU the adjacency and inverse adjacency matrices are represented using a bitset class, so they must be converted to arrays of unsigned integers to be able to be passed to the GPU. Before launching the kernel, all of the global memory must be allocated for various types of data such as current cliques, candidate sets, vertex color arrays, the adjacency matrices, local maximum cliques, and more. After memory is allocated and data is correctly represented, data for the adjacency matrices and block subtrees have to copied to global memory on the GPU. Lastly, since the algorithm supports the use of multiple GPUs, threads must be created to launch a BBMCG kernel on each GPU. The multiple GPU variation of BBMCG is primarily used to increase the amount of available memory and will be explained in the discussion section. In the post-processing stage, each thread copies the local maximums of the blocks to CPU memory and finds the largest clique among the blocks.

4 Results and Discussion

Sequential Implementation and Test Environment: The sequential version of BBMC used in the results and the preprocessing portion of BBMCG were both written in C++11. The sequential implementation is a port of the BBMC algorithm provided in [6], which provides explanations of many different sequential maximum clique solvers and their implementations in Java. It is important to note that our C++ version of BBMC is slightly slower than the Java implementation it was based on. The C++11 version is nearly identical to the Java implementation, but uses vectors from the standard template library in place of the ArrayList class in Java and Boost's dynamic_bitset class in place of Java's BitSet class. Differences in these class implementations may be the cause of the performance difference, but further investigation is needed.

The computer used to gather both sequential and parallel GPU results is composed of a Intel(R) Core(TM) i7 CPU 4790K @ 4.00 GHz, two NVIDIA GTX 970's with 1664 CUDA cores running at 1050 MHz and 4 GB of memory each, and 16 GB of DDR3 system memory. Randomly generated graphs and a subset of the DIMACS benchmark graphs [3] were used to compare the sequential BBMC and parallel BBMCG algorithms.

Results: The randomly generated graphs are used to show trends for the sequential and parallel algorithms as graph size and edge density change. Figure 1 compares the running times on randomly generate graphs with 200 vertices and edge probabilities p. In the graph each pair of vertices has a probability of p to be an edge in the

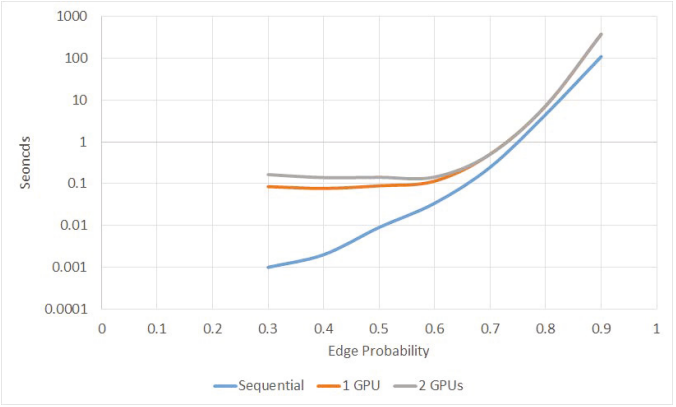


Fig. 1 Growth rate of BBMC and BBMCG run times with respect to graph density on randomly generated graphs with 200 vertices.

graph. The results show that at a size of 200 vertices, the sequential algorithm outperforms BBMCG in all cases, but the two have similar growth rates with respect to increasing edge probability. It can also be seen the preprocessing stage for BBMCG takes approximately 0.1 seconds, and is slightly higher when running on multiple GPUs than on a single one. It is the case that preprocessing always take approximately that much time, and therefore is negligible in nontrivial problems. A final observation that can be made is that BBMCG runs nearly identical on multiple GPUs as it does on a single GPU for nontrivial problems.

Figure 2 compares the running times of BBMC and BBMCG on various graph sizes. It can be seen that the growth rate for the BBMCG algorithm with respect to graph size is much lower than the growth rate of the sequential algorithm. The parallel algorithm outperforms the sequential on all graph sizes above 400 vertices. Once again, the multiple GPU implementation performs almost identical to the single GPU implementation in almost all cases.

We ran BBMC and BBMCG algorithms to find the maximum clique in a subset of the DIMACS benchmark graphs. Results show that on 20 of the 24 selected graphs, BBMCG outperforms the sequential algorithm. The four graphs where speedup does not occur all have less than 400 vertices, which is consistent with the results shown in Figure 1 and Figure 2. Typical speedup is between the 1.4 to 4.0 range, with a few more extreme outliers. The results presented use minimum-width ordering for the initial ordering, which tends to perform best on the majority of graphs for BBMC and BBMCG. Experiments found that when using non-increasing degree or MCR methods for initial ordering the results were comparable.

Analysis: The results were consistently positive on graphs of 400 vertices or more. This seems to be the threshold where the BBMCG algorithm can start taking advantage of the parallel processing power of the GPU. As the size of the graph increases,

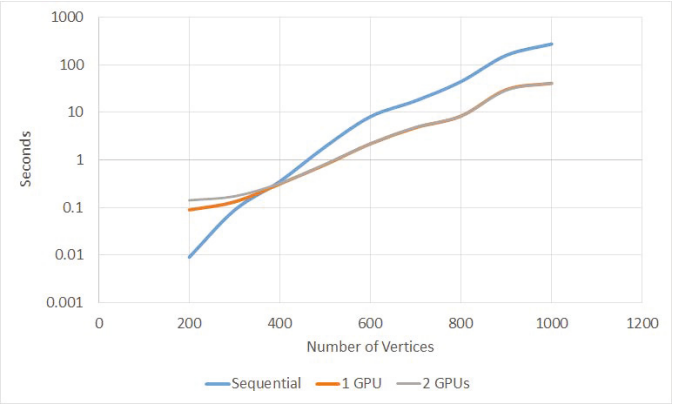


Fig. 2 Growth rate of BBMC and BBMCG run times with respect to graph size on randomly generated graphs with 0.5 edge probability.

both block-wise and thread-wise parallel components increase in parallelism. The number of blocks used in BBMCG is equivalent to the number of vertices in the graph. Therefore, larger graphs have more ways to divide up the search space and consequently can optimally use more blocks to search different subtrees. The number of subtrees used to split the search space into was experimented with, and the ideal number of blocks to use almost always turned out to be the number of vertices in the graph, not more or less. A performance gain is much more likely on larger graphs, as many different portions of the graph are all searched simultaneously with the possibility of containing the maximum clique. Searching the graph subtrees simultaneously means that large cliques can be found from subtrees that would not be searched sequentially until later in the search, further pruning the graph and removing unnecessary searching. This can sometimes lead to extreme speedup, such as in the case of san400_0.9_1 where BBMCG obtained 2,3076.9 speedup over BBMC.

Thread-wise parallelism is a much more consistent method of increasing performance. Block-wise parallelism relies on a non-deterministic search method in an attempt to randomly guess good search paths and find large cliques faster in the graph, and will in most cases, but the strategy depends heavily on the characteristics of the graph being searched. Thread-wise parallelism speeds up many forms of bit-wise operations in the algorithm and does so regardless of the characteristics of the graph. As the size of the graph increases, the number of unsigned integers needed in a bitset to represent the graph increases. Since the number of threads in a block is equivalent to the number of unsigned integers in a bitset, it means that more threads can be used as the graph size increases. In order for a block to make use of 32 threads, which is the number of threads in a warp, the graph must have at least 993 vertices. Since the GPU runs threads in groups the size of warps, most of the graphs cause BBMCG to run idling threads. Only 4 of the 24 DIMACS graphs make use of a full warp. As a result, the majority of graphs tested are not running optimally on the GPU. Another negative aspcect of thread-wise parallelism is the fact that not all

operations can be parallelized, such as setting a bit, causing all but 1 thread to idle at these occurrences.

One severe limitation of the BBMCG algorithm is the lack of load balancing. When a block finishes searching its subtree, it runs out of work and terminates. It is very unlikely that each block is given an equal amount of work when the kernel is initialized, so it is very possible the most blocks only work for a fraction of the run time while a few blocks search the most difficult subtrees. A consequence of this is that the amount of parallelism reduces as the algorithm runs and performance is decreased.

The multiple GPU implementation of BBMCG suffers from the same load balancing issue and is the reason that it performs identically to a single GPU implementation. In the multiple GPU implementation, the initial subtrees for the blocks are divided among the multiple GPUs. Unified memory is used to shared the current maximum clique size between the GPUs and each GPU updates the maximum size as they find it. Unified memory has to be communicated between GPUs, causing a slight slow down. Although the additional GPUs do provide additional computational resources, it is unlikely that they have an even work load on initialization of the kernel. The benefits of finding and sharing large cliques faster across GPUs is limited by the lack of load balancing and slow communication, yielding almost no performance benefits. The block that has to search the most difficult subtree has to do it alone regardless of the addition of more GPUs. Multiple GPUs do provide additional memory however, and therefore do have some use to BBMCG. A maximum graph size of around 1500 vertices can run on a single GPU, but a maximum of around 2500 vertices can run on two GPUs.

BBMCG is also slowed down by its reliance on global memory to run. Global memory is many times slower than other available forms of memory, such as shared memory. Although shared memory is used when necessary in bitwise operations and in the coloring process, it is used seldom elsewhere. Another memory related issue is the maximum stack size CUDA allows. Even at the maximum size, it was determined that the maximum level of recursion that could be reached before memory was exhausted was approximately 500 levels. This limits BBMCG to only being able to search graphs that have a maximum clique size of about 500 maximum until it run out of memory.

5 Conclusions and Future Work

In this paper, a maximum clique algorithm for GPUs, BBMCG, was presented based off of one of the leading sequential algorithms, BBMC. The algorithm addresses memory limitations for maximum clique solvers on GPUs by using bitsets to reduce memory requirements. Two dimensions of parallelism are utilized on the GPU to divide up and search the graph's search space in parallel and to improve computational speed by parallelizing bit-wise operations. The result achieved is moderate speedup on all but small graph sizes and occasionally significant speedup on some problems.

It is promising that speedup is consistently achieved regardless of the reliance on slow global memory, a lack of a load balancing system, and small graph sizes not utilizing all of the threads in a warp. Future work will attempt to address the issues to improve performance further. Preliminary work on a load balancing system suggests that it could be a source for a large boost in performance and would have the benefit of enabling multiple GPU systems to provide further speedup. Another possibly significant alteration would be to integrate a managed memory pool system in global memory with the goal to eliminate the inefficient allocation of excess memory to blocks caused by not knowing how much memory is needed to solve the problem.

Acknowledgements This material is based in part upon work supported by: The National Science Foundation under grant number(s) IIA-1329469, and by Cubix Corporation through use of their PCIe slot expansion hardware solutions and HostEngine. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or Cubix Corporation.

References

1. Cruz, R., Lopez, N., Trefftz, C.: Parallelizing a heuristic for the maximum clique problem on gpus and clusters of workstations. In: 2013 IEEE International Conference on Electro/Information Technology (EIT), pp. 1–6. IEEE (2013)
2. Howbert, J.J., Roberts, J.: The maximum clique problem (2007). http://courses.cs.washington.edu/courses/csep521/07wi/prj/jeff_jacki.pdf (accessed: May 9, 2015)
3. Johnson, D., Trick, M.: Cliques, coloring, and satisfiability, dimacs series in disc. Math. and Theoret. Comput. Sci. **26** (1996)
4. McCreesh, C., Prosser, P.: Multi-threading a state-of-the-art maximum clique algorithm. Algorithms **6**(4), 618–635 (2013)
5. NVIDIA Corporation: Cuda c programming guide (2015). http://docs.nvidia.com/cuda/cuda-c-programming-guide/ (accessed: May 9, 2015)
6. Prosser, P.: Exact algorithms for maximum clique: A computational study. Algorithms **5**(4), 545–587 (2012)
7. Rossi, R.A., Gleich, D.F., Gebremedhin, A.H., Patwary, M., Ali, M.: Parallel maximum clique algorithms with applications to network analysis and storage. arXiv preprint arXiv:1302.6256 (2013)
8. Rossi, R.A., Gleich, D.F., Gebremedhin, A.H., Patwary, M.M.A.: A fast parallel maximum clique algorithm for large sparse graphs and temporal strong components. CoRR abs/1302.6256 (2013). http://arxiv.org/abs/1302.6256
9. San Segundo, P., Matia, F., Rodriguez-Losada, D., Hernando, M.: An improved bit parallel exact maximum clique algorithm. Optimization Letters **7**(3), 467–479 (2013)
10. San Segundo, P., Rodriguez-Losada, D.: Robust global feature based data association with a sparse bit optimized maximum clique algorithm. IEEE Transactions on Robotics **29**(5), 1332–1339 (2013)
11. San Segundo, P., Rodríguez-Losada, D., Jiménez, A.: An exact bit-parallel algorithm for the maximum clique problem. Computers & Operations Research **38**(2), 571–581 (2011)
12. San Segundo, P., Rodríguez-Losada, D., Matía, F., Galán, R.: Fast exact feature based data correspondence search with an efficient bit-parallel mcp solver. Applied Intelligence **32**(3), 311–329 (2010)
13. San Segundo, P., Tapia, C.: Relaxed approximate coloring in exact maximum clique search. Computers & Operations Research **44**, 185–192 (2014)

14. Tomita, E., Kameda, T.: An efficient branch-and-bound algorithm for finding a maximum clique with computational experiments. Journal of Global Optimization **37**(1), 95–111 (2007)
15. Tomita, E., Seki, T.: An efficient branch-and-bound algorithm for finding a maximum clique. In: Discrete Mathematics and Theoretical Computer Science, pp. 278–289. Springer (2003)
16. Trefftz, C., Santamaria-Galvis, A., Cruz, R.: Parallelizing an algorithm to find the maximal clique on interval graphs on graphical processing units. In: 2014 IEEE International Conference on Electro/Information Technology (EIT), pp. 100–102. IEEE (2014)
17. Xiang, J., Guo, C., Aboulnaga, A.: Scalable maximum clique computation using mapreduce. In: 2013 IEEE 29th International Conference on Data Engineering (ICDE), pp. 74–85. IEEE (2013)

CUDA Implementation of Computer Go Game Tree Search

Christine Johnson, Lee Barford, Sergiu M. Dascalu and Frederick C. Harris Jr.

Abstract Go is a fascinating game that has yet to be played well by a computer program due to its large board size and exponential time complexity. This paper presents a GPU implementation of PV-Split, a parallel implementation of a widely used game tree search algorithm for two-player zero-sum games. With many game trees, it often takes too much time to traverse the entire tree, but theoretically, the deeper the tree is traversed, the more accurate the best move found will be. By parallelizing the Go game tree search, we have successfully reduced the computation time, enabling deeper levels of the tree to be reached in smaller amounts of time. Results for the sequential and GPU implementations were compared, and the highest speedup achieved with the parallel algorithm was approximately 72x at 6 levels deep in the game tree. Although there has been related work with respect to game tree searches on the GPU, no exact best move search algorithms have been presented for Go, which uses significantly more memory due to its large board size. This paper also presents a technique for reducing the amount of required memory from previous game tree traversal methods while still allowing each processing element to play out games independently.

1 Introduction

Computer Go has been a popular research topic for over 50 years [14] due to its classification as a PSPACE-hard problem, meaning it is at least as hard as the most difficult problems in the PSPACE class, and it has not been proven to be solved using

C. Johnson · L. Barford · S.M. Dascalu · F.C. Harris, Jr.(✉)
Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA
e-mail: c_johnson72@hotmail.com, lee_barford@ieee.org, Fred.Harris@cse.unr.edu

L. Barford
Keysight Laboratories, Keysight Technologies, Santa Rosa, CA, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_31

a polynomial amount of space [11]. The characteristic of this problem that attributes to its exponential time complexity is the number of possible games is $\approx 10^{171}$ [1], and a program needs to check all of these games to determine which one yields the best score for a player. Thus, running multiple of these checks in parallel is a reasonable way of reducing the computation time.

Graphics Processing Units (GPUs) have become popular devices for parallel computing due to their low price to performance ratio. NVIDIA has allowed programmers to take advantage of their GPUs for general-purpose computing with the CUDA platform. CUDA C has been one of the most successful languages ever designed for parallel computing [10], making it an optimal choice for reducing program execution time by accelerating computation with the GPU.

The remainder of this paper is structured as follows. Section 2 provides the necessary background information, such as the rules for playing Go and game tree search methods. Section 3 references related research on the Go game tree search and parallel game tree searches. Section 4 discusses the GPU implementation of the Go game tree search. Section 5 compares the sequential and parallel results, and explains analyses of those results. Finally, Section 6 concludes and discusses future work.

2 Background

Although researchers have been studying the complexity of Computer Go since 1962, the game of Go was invented in China around 2300 B.C.. The first program that was capable of beating a novice player was written in 1968 [14]; however, there has yet to be a program written that can beat any expert player [3]. Such a program is challenging for a number of reasons: the search space of all possible moves is very large, it is difficult to construct long-term strategies, and determining the end of the game is not intuitive in a program [13]. The latter two reasons are the most difficult to address, so most research is aimed at speeding up the necessary computation needed to search the possible move space.

Go is often compared to Chess, due to them being similar with regards to being two-player strategy games [1]. However, Chess is considered a *solved* game because a Chess program was written that successfully defeated the world champion [4]. The search space of Chess is estimated to be $\approx 10^{70}$ where the search space of Go is estimated to be $\approx 10^{171}$ [1]. Additionally, Chess has a distinct *end game*, such as a checkmate state, where Go does not [13], as mentioned previously.

2.1 Rules

Go has many rules for specific board states; however, only the basic rules and those relevant to the presented research will be listed. The following rules are referenced from [13, 14].

Standard Go is played on a 19x19 board, creating 361 possible *intersections*, or places for players to place their stones. There are two players, one black and one white, who take turns placing their stones on the intersections of the game board. The black player has 181 stones and white has 180, the sum of them equaling the number of intersections. Stones of the same color that occupy consecutive intersections on the board are called *blocks*. The number of *liberties* for a stone is the number of neighboring intersections that are not occupied by any stones. A stone can have a maximum amount of four liberties. When the opponent's stones occupy all of a stone's liberties, the stone is considered *captured* by the opponent, and is removed from the board. Blocks of stones can be captured as well if the entire block is surrounded by the opponent's stones.

The black player always goes first, and *komi* applies, meaning the white player receives a small amount of extra points on the final score since he or she is at the disadvantage of going second.

Objective. The objective of the game is to control the most territory. Stones are strategically placed to protect stones from being captured, and to capture the opponent's stones.

Scoring. There exist two variants of scoring methods: *territory scoring*, which count the surrounded territory plus the number of captured stones, and *area scoring*, which count the surrounded territory plus the *alive* stones on the board. For this research the latter is used. Contrary to standard Go rules but for simplicity, alive stones are just those that have not been captured. *Surrounded territory* is defined as blocks of empty intersections adjacent to stones of a single player.

End Game. The game is over when both players pass consecutively because they can no longer make any moves. Since it is difficult to detect when humans would surrender, in this research the game ends after a number of moves equal to the number of stones has been made.

2.2 Game Tree

Like most two-player games, Go program implementations typically use a game tree to store all possible moves that can be made from the remaining set of moves to choose from. The aim of game tree traversal algorithms is to allow the player to make the best move. This is achieved by traversing the tree to observe how each combination of moves would end a game, allowing the player to choose the move that will yield the best score. In the tree, the root represents the current game state, and all other node represents possible game states. Consecutive game states are reached by making a move in the game [8]. The children of each node represent all remaining moves that can be made after the move represented by the node, and some evaluation function is used at each leaf to determine the score for the player by ending the game with this combination of moves [7]. No algorithm exists that can traverse the entire game tree for Go in a reasonable amount of time [7, 14].

The following are descriptions for the most common methods for game tree traversal, which all use Depth-First Search [8]. Each consecutive method is an enhancement on the previous, leading up to the method used in this research.

Minimax: Minimax tree search or variations of it are commonly used in two-player games because they attempt to maximize the score of the player while also trying to minimize the score of the opponent. All the even levels in the tree represent moves for the current player, and thus contain the moves at which the player wants to maximize its score. All the odd levels in the tree represent moves for the opponent, or moves where the player wants to minimize the score [8]. Starting at the leaf nodes and choosing best and worst nodes on the even and odd levels, respectively, yields the best possible move [14]. The NegaMax implementation has a single maximize function that it uses the negation of when minimizing [7].

Alpha-Beta Pruning: Alpha-Beta Pruning is an enhancement to Minimax that yields the same move, but finds it more efficiently by pruning away entire branches of the tree not worth exploring, thus reducing the search space [7]. A variable α represents the worst possible score the maximize player, and a variable β represents the worst possible score for the minimize player. The variables are updated if better values for each are found. Any node with a value below the current α will not be chosen by the maximizer, so they are pruned. Similarly, any nodes with values greater than β will not be chosen by the minimizer, so they are pruned [14].

Principal Variation Search: Principal Variation Search (PVS), also known as NegaScout, is an enhancement to Alpha-Beta pruning that finds the same move more efficiently by reducing the size of the search window, thus pruning more of the tree branches. The algorithm assumes the tree is ordered in such a way that the best move out of the children of a node will be the left-most child. The path that leads to the best move is the *principal variation* [9]. At each node, if the first child of that node does not fall into the minimal search window, then the current node is pruned. In order for this algorithm to be most efficient, the best move should be explored first, followed by the next best move, etc. Iterative-deepening and transposition tables are enhancements that address move-reordering [7]. If the tree is poorly ordered, each sub-tree that is better than its elder siblings must be searched again, making it potentially less efficient than the Alpha-Beta pruning [9].

Iterative-Deepening: Iterative-deepening is an enhancement to Depth-First Search, which sets the max depth for traversal to some shallow level initially, and re-traverses the tree after increasing the depth each time [14]. This is beneficial for a number of reasons; one of them being that information from previous iterations can be stored so those levels do not actually need to be traversed again in each iteration [13]. A second reason is this allows us to essentially convert a Depth-First Search into a Breadth-First Search without the high memory requirements of BFS.

3 Related Work

For over 50 years, Go has been an intriguing research topic. Research has been performed involving the use of neural-networks for supervised learning [14] as well as Monte-Carlo methods, which choose the best move out of a set of randomly selected paths in the game tree to obtain decent results [5]. This research focuses on related work done to optimize game tree search for an exact best move using variations of Alpha-Beta Search.

Van der Werf *et al.* [13] solved Go for any possible opening move on a 5x5 board in 2002. The authors incorporated the five goals for their heuristic function discussed previously in Section 2. Assumptions were made that may only hold for smaller sized boards. The best move was found 23 levels deep in the game tree in approximately 4 hours. Future work included solving Go on 6x6 and 7x7 boards.

Strnad *et al.* [12] implemented PV-Split on the GPU for the game Reverse, also known as Othello in 2011. The authors designed an iterative algorithm, since GPU recursion was not supported on NVIDIA cards at that time. Future work included implementing transposition tables and iterative deepening to improve the algorithm.

Elnaggar *et al.* [7] did comparisons of the variations of game tree traversal, such as Minimax, Alpha-Beta Search, and Principal Variation Search. Results for a subset of the methods were presented; implemented sequentially and in parallel using MPI in 2014. Future work included using the OpenCL or CUDA libraries for GPU implementations.

Li *et al.* [8] designed a parallel algorithm for game tree traversal for Connect6 and Chess that was ran on the GPU in 2014. As opposed to using a variation of Alpha-Beta Search, the authors use a node-based parallel algorithm; meaning sets of nodes rather than sub-trees are assigned to each GPU processor to avoid the complexities involved with tree splitting. Future work included running the presented method on multiple GPUs, and on larger search spaces such as the Go game tree to observe the effectiveness and efficiency of the method.

Elnaggar *et al.* [6] designed a robot that can autonomously play Checkers with a human in 2014. The Checkers game tree was traversed using an iterative-deepening Alpha-Beta Search that was ran on the GPU. The novel feature of the presented method is the GPU kernel recursion, which was recently made available by NVIDIA in CUDA 5.0 on devices of Compute Capability 3.5 or higher [2]. Future work included using the presented method to solve more complex games like Go and Stratego.

4 GPU Method

The deployed GPU implementation is a parallel version of the Alpha-Beta algorithm called PV-Split, which was first presented by [9]. This algorithm was chosen because it allows for high parallelism opportunities, and the data dependencies are minimal and manageable on the GPU. The following is a more detailed description of the algorithm.

4.1 PV-Split

If the game tree is naively divided into sub-trees that are processed in parallel, it is likely that redundant nodes or nodes that would pruned in the sequential algorithm will be explored due to the search window not being set properly [9]. PV-Split was designed based upon how the sequential Alpha-Beta algorithm searches the tree with optimal move-ordering. As stated previously, optimal move-ordering is defined as having the moves in the tree ordered from best to worst, implying the principal variation is the left-most path. PV-Split starts by traversing the left-most path sequentially to initialize the Alpha-Beta search window. Once the cut-off level for this iteration has been reached, the remaining siblings in each level are processed in parallel as the algorithm recursively goes back up the tree [9]. Figure 1 shows how PV-Split would divide the game tree among multiple processing elements.

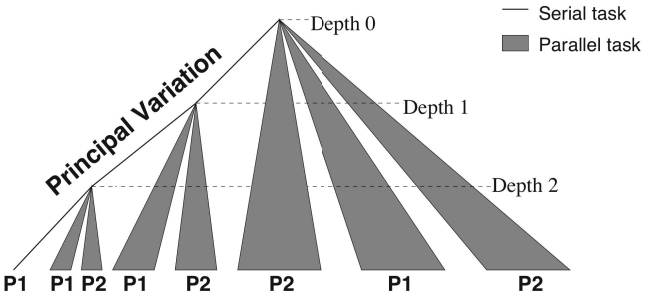


Fig. 1 PV-Split Search using Two Processing Elements.

The minimal search space is achieved when the best α value for each level is possessed by each processing element [9]. Alpha-beta windows are stored globally for each level of the tree, so if one processing element finds the best move in a level, it atomically updates the window for that level and the rest of the moves can be pruned from the search space [12]. Similar to PVS, PV-Split is most efficient with optimal move-ordering, thus iterative-deepening is used to reorder the moves at each iteration.

The GPU implementation uses two recursive device functions to perform the traversal of the left-most path and the tree splitting. The pseudocode for the device functions is shown in Algorithms 1 and 2.

PV-Split suffers from one main disadvantage: the sub-trees are not guaranteed to be of equal size due to the pruning, and PV-Split does not enforce any type of load-balancing. If one processing element takes longer to complete its sub-tree, the remaining processing elements are forced to wait until it completes, lessening the efficiency of the system.

Algorithm 1. PV-Split

```
1: function PVSPLIT(side, remainingLevels, alpha, beta)
2:     if remainingLevels = 0 then
3:         return evaluateLiberties(side)
4:     end if
5:
6:     find the left-most child
7:     make the move
8:     score ← -PVSPLIT(-side, remainingLevels - 1,
9:                      -beta, -alpha)
10:    if score > alpha then
11:        alpha ← score
12:        updateAlphaBeta(side, depth, score)
13:    end if
14:    undo move
15:    if alpha ≥ beta then
16:        return alpha
17:    end if
18:
19:    divide all threads among the remaining siblings
20:    while move < totalMoves do
21:        make move
22:        score ← -TREESPLIT(-side,
23:                           remainingLevels - 1,
24:                           -beta, -alpha)
25:        if score > alpha then
26:            alpha ← score
27:            updateAlphaBeta(side, depth, score)
28:        end if
29:        undo move
30:        if alpha ≥ beta then
31:            return alpha
32:        end if
33:
34:        move ← next available move
35:
36:    end while
37:
38:    return alpha
39: end function
```

4.2 Memory Optimization

Processing sub-trees in parallel allows for the play out of multiple games simultaneously, thus requiring multiple game boards. A common method for maintaining independent games is to create a copy of the current board state for each processing element [6]. This method works well for small game trees, but becomes infeasible with large trees such as the one required for Go. The infeasabililty stems from the fact that the maximum width of the Go game tree is 361!, requiring the maximum amount

of blocks and threads with striding to process the nodes in parallel. A thread's board copy can be reused as it strides, so the maximum amount of board copies is bounded by the total number of threads, which is 67,107,840 threads for the architecture of the device on which the experiments were performed [2]. Storing this many copies of the board exceeds the available device global memory.

In order to reduce the amount of memory required, there were no additional copies of the game board created. A single copy of the board was used, and each thread was allocated a key that was used to determine if a move had been processed by the thread. Each key is represented by a bit. The keys are stored in an array of integers, each integer holding 32 keys. A second integer array of the same size was used to determine if an occupied intersection was occupied by a white or black stone. Therefore, the total memory used is that required to store the game board and two bits for each thread, which reduces the memory from the previous method by a factor of over 1,400 when using the maximum amount of threads.

5 Results and Analysis

Results were collected for the situation where the black player has placed the first stone, and the game tree is searched for the best move for the white player. PVS was implemented in C++ and PV-Split was implemented in CUDA C. Sequential and parallel timings were collected for a 5x5 board and the standard 19x19 board. The system used for the experiments has the following specifications: Intel i7 4790k running at 4.0 GHz, 16 GB DDR3 memory, NVIDIA GeForce GTX 970.

Figures 2 - 4 show the execution times and speedups for the two board sizes

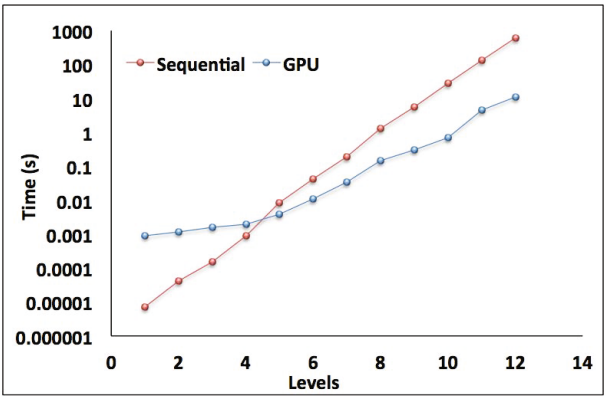


Fig. 2 Execution times for a 5x5 board.

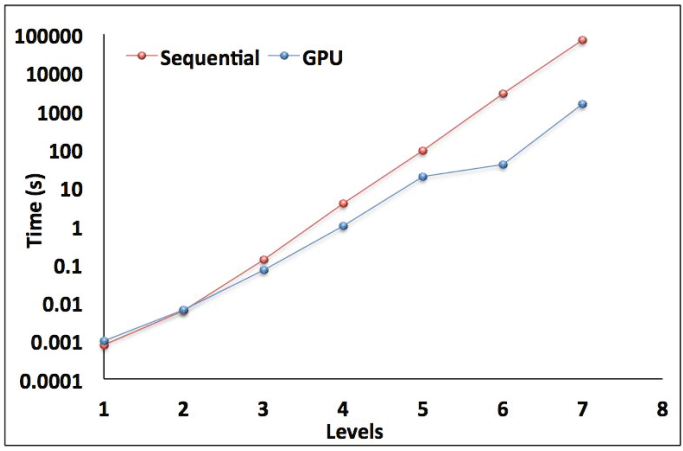


Fig. 3 Execution times for a 19x19 board.

Algorithm 2. Tree-Split

```
1: function TREESPLIT(side, remainingLevels, alpha, beta)
2:     if remainingLevels = 0 then
3:         return evaluateLiberties(side)
4:     end if
5:
6:     divide available threads among remaining moves
7:     while move < totalMoves do
8:         make move
9:         score ← -TREESPLIT(-side,
10:                            remainingLevels - 1,
11:                            -beta, -alpha)
12:        if score > alpha then
13:            alpha ← score
14:            updateAlphaBeta(side, depth, score)
15:        end if
16:        undo move
17:        if alpha ≥ beta then
18:            return alpha
19:        end if
20:
21:        move ← next available move
22:
23:    end while
24:
25:    return alpha
26: end function
```

The graph in Figure 2 shows that it is only beneficial to move the search computation to the GPU once the number of levels exceeds four for the 5x5 board. Figure 3 shows it is beneficial to move the computation to the GPU once the amount of levels

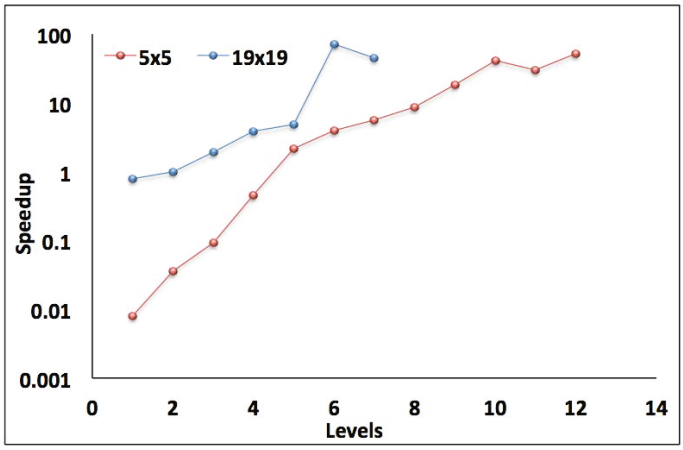


Fig. 4 Speedups for the 5x5 and 19x19 boards.

exceeds two for the 19x19 board. It is expected that the level threshold at which it is beneficial to move the computation to the GPU is lower for the 19x19 board since the breadth of the 19x19 game tree is larger.

For the GPU implementation, various combinations of block and thread amounts were executed to determine the optimal combination for both board sizes. The combinations were tested for deeper levels of the trees, but the same combination was used for all levels for the results for that board size. For the 5x5 board, 4096 blocks and 32 threads per block were used, and for the 19x19 board, 512 blocks and 32 threads per block were used. As the number of threads passed these values, the execution time would start increasing.

Increased execution time, when a result of an increase in the number of threads, can typically be attributed to either thread scheduling or memory throughput. To further understand the reasoning for the increased execution time, the CUDA profiler [2] was used to analyze the differences in various memory usages as the number of threads increased.

As expected, the results for 1 and 32 threads are the same since a warp will always schedule 32 threads. However, as the number of threads increases from 32 to 128, the number of blocks that can be simultaneously executed on a Streaming Multiprocessor reduces from 20 to 5. In this situation, the number of blocks that can be executed is limited by the number of registers since the recursive kernel has high register usage. Each GPU model has a set limit on the number of registers that can be used by each block, so if this amount is exceeded then some of the blocks must wait to be scheduled. Increasing the number of threads increases the number of times the kernel is executed, thus increasing the register usage. As a result, the decreased warp occupancy theoretically reduces the number of elements processing simultaneously and also reduces the efficiency at which the GPU can perform latency hiding, both attributing to an increase in execution time [2].

6 Conclusion and Future Work

In this paper, we have presented a CUDA implementation of the recursive PV-Split algorithm for searching the Go game tree. Running the game tree search in parallel allows for a reduction in computation time so more accurate *best moves* can be found by reaching deeper levels of the tree in a less amounts of time. The results for the standard board size show speedup for all levels greater than 2, with a high speedup of ≈72x at ply 6. In addition to reducing the computation time, a memory optimization method was introduced that addressed the issue of having to store independent copies of the game board for each processing element on the GPU.

In the future, we would like to make optimizations to the current GPU implementation, such as reducing the amount of register usage so a higher number of threads can be used efficiently, and taking advantage of shared memory to promote faster memory accesses. We would also like to run the game tree search on multiple GPUs.

Acknowledgment This material is based in part upon work supported by: The National Science Foundation under grant number(s) IIA-1329469, and by Cubix Corporation through use of their PCIe slot expansion hardware solutions and HostEngine. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or Cubix Corporation.

References

1. Comparision between chess and go. http://users.eniinternet.com/bradleym/Compare.html
2. Cuda toolkit documentation. https://docs.nvidia.com/cuda/index.html
3. The mystery of go, the ancient game that computers still can't win. http://www.wired.com/2014/05/the-world-of-computer-go/
4. Campbell, M., Hoane, A.J., Hsu, F.H.: Deep blue. Artificial Intelligence **134**(1), 57–83 (2002)
5. Chou, C.W., Chou, P.C., Doghmen, H., Lee, C.S., Su, T.C., Teytaud, F., Teytaud, O., Wang, H.M., Wang, M.H., Wu, L.W., et al.: Towards a solution of 7x7 go with meta-MCTS. In: Advances in Computer Games, pp. 84–95. Springer (2012)
6. Elnaggar, A., Gadallah, M., Aziem, M.A., Aldeeb, H., et al.: Autonomous checkers robot using enhanced massive parallel game tree search. In: 2014 9th International Conference on Informatics and Systems (INFOS), pp. PDC–35. IEEE (2014)
7. Elnaggar, A.A., Aziem, M.A., Gadallah, M., El-Deeb, H.: A comparative study of game tree searching methods. International Journal of Advanced Computer Science and Applications (IJACSA) **5**(5) (2014)
8. Li, L., Liu, H., Wang, H., Liu, T., Li, W.: A parallel algorithm for game tree search using GPGPU (2014)
9. Marsland, T.A., Campbell, M.: Parallel search of strongly ordered game trees. ACM Computing Surveys (CSUR) **14**(4), 533–551 (1982)
10. Sanders, J., Kandrot, E.: CUDA by example: an introduction to general-purpose GPU programming. Addison-Wesley Professional (2010)
11. Sipser, M.: Introduction to the Theory of Computation. Cengage Learning (2012)

12. Strnad, D., Guid, N.: Parallel alpha-beta algorithm on the GPU. In: 2011 33rd International Conference on Information Technology Interfaces (ITI), Proceedings of the ITI, pp. 571–576. IEEE (2011)
13. van der Werf, E.C., Van Den Herik, H.J., Uiterwijk, J.W.: Solving go on small boards. ICGA Journal **26**(2), 92–107 (2003)
14. van der Werf, E.C.D.: AI techniques for the game of Go. UPM, Universitaire Pers Maastricht (2005)

Negotiation and Collaboration Protocol Based on EbXML Intended to Optimize Port Processes

Vinícius Eduardo Ferreira dos Santos Silva, Nunzio Marco Torrisi and Rodrigo Palucci Pantoni

Abstract This paper proposes the replacement of the EDI (Electronic Data Interchange) technology, used in ports for information exchange between ports and ships, by more effective, flexible and open alternative. The proposal is motivated to design a negotiation protocol based on EbXML (Electronic Business using eXtensible Markup Language) to reduce interaction among actors in uses cases flows of the system. The methodology was based on system analysis techniques and documents of the port of Santos (Brazil). The results were validated using uses cases flows by reducing the number of interactions among the different entities or actors of the system. As conclusion, the study confirms that the substitution of technology is feasible and can be applied to other port information systems with the advantage of optimizing port processes.

Keywords EbXML · EDI · Electronic trading · Corporate interoperability · Port system model

1 Introduction

Conceptually, a port can be defined as an area located on the ocean border, sea, lake or river, for the berthing of boats and ships [1].

V.E.F. dos Santos Silva · N.M. Torrisi · R.P. Pantoni
Department of Computer Science, Federal University of ABC (UFABC), São Bernardo, Brazil
e-mail: {vinicius.eduardo,nunzio.torrisi}@ufabc.edu.br

R.P. Pantoni(✉)
Department of Informatics, Federal Institute of Education, Science and Technology of São Paulo (IFSP), Sertãozinho - SP, Brazil
e-mail: rpantoni@ifsp.edu.br

S. Latifi (ed.) , *Information Technology New Generations*, Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_32

The purpose of a port is to provide services for supporting ship loading, unloading, temporary storage of cargo - in some cases, specially designated terminals for passengers. In this context, modernization and automation of the ports has been developed, which has greatly increased amount of cargo with agility processes [2].

The global port system has invested heavily in Information Technology (IT), since the implementation of information systems has become synonymous with productivity and competitiveness.

The ports have been traditionally adopted a standard for information exchange technology among the port and its partners (administrative and commercial focus). This standard is called EDI (Electronic Data Interchange) [3]. There are specializations of the standard, for example, EDIFACT (Data Interchange For Administration, Commerce and Transport), that is the international standard EDI developed under the United Nations.

However, the EDI technology has some disadvantages:

- Proprietary technology. Only solutions that use this technology can be interoperable;
- Processes that use this technology need human intervention, generating delays in the process, lower productivity and increased costs [4, 5];
- High cost to implement, low participation of small and medium enterprises and difficulties in consolidating a worldwide standard [6].

In view of the disadvantages mentioned, there is an open and interoperable technology called EbXML (Electronic Business using eXtensible Markup Language), which is intended for message exchanges (e-business).

Thus, the aim of this work is to study and design a part of port information system of Santos Port (Brazil), based on replacing the technology to EDI-DTE (specific EDI format from this port) for EbXML. Such substitution is intended to simplify or optimize port processes and enables the integration of systems and their functionality, facilitating the electronic exchange of data among all actors involved in the processes. The challenge of this work is to attest the improvement of port processes by replacing technology.

2 EbXML

EbXML is a family of XML (eXtensible Markup Language) based on standards sponsored by OASIS and UN/CEFACT. It is a modular suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet. Using EbXML, companies have a standard method to exchange business messages, conduct trading relationships, communicate data in common terms and define and register business processes [7].

The information included in this section was obtained in the standard EbXML Specification [8].

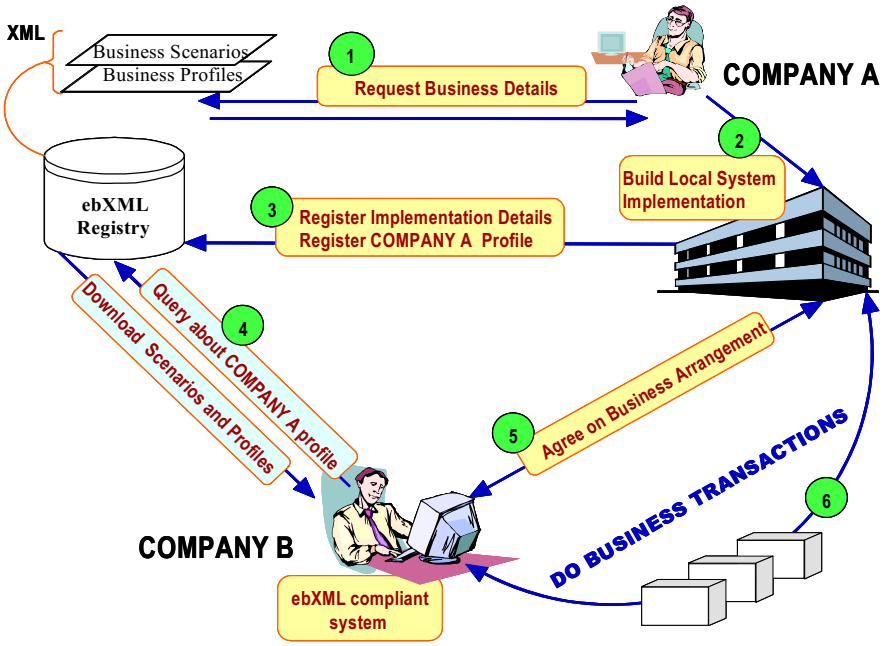


Fig. 1 Two companies conducting eBusiness using EbXML (Source: [8]).

Fig. 1 shows a high-level use case scenario for two Trading Partners, first configuring and then engaging in a simple business transaction and interchange. In the context of this work, the two entities can be the port and ship, for example. This model is provided as an example of the process and steps that may be required to configure and deploy EbXML Applications and related architecture components. These components can be implemented in an incremental manner. The EbXML specifications are not limited to this simple model, provided here as quick introduction to the concepts [8].

The conceptual overview described below introduces the following concepts and underlying architecture [8]:

1. A standard mechanism for describing a Business Process and its associated information model.
2. A mechanism for registering and storing Business Process and Information Meta Models so they can be shared and reused.
3. Discovery of information about each participant including:
   - The Business Processes they support.
   - The Business Service Interfaces they offer in support of the Business Process.
   - The Business Messages that are exchanged between their respective Business Service Interfaces.
   - The technical configuration of the supported transport, security and encoding protocols.

4. A mechanism for registering the aforementioned information so that it may be discovered and retrieved.
5. A mechanism for describing the execution of a mutually agreed upon business arrangement which can be derived from information provided by each participant from item 3 above (Collaboration Protocol Agreement – CPA).
6. A standardized business Messaging Service framework that enables interoperable, secure and reliable exchange of Messages between Trading Partners.
7. A mechanism for configuration of the respective Messaging Services to engage in the agreed upon Business Process in accordance with the constraints defined in the business arrangement.

In Fig. 1, Company A (port, ship or any other related actor) has become aware of an EbXML Registry that is accessible on the Internet (Fig. 1, step 1). Company A, after reviewing the contents of the EbXML Registry, decides to build and deploy its own EbXML compliant application (Fig. 1, step 2). Custom software development is not a necessary prerequisite for EbXML participation. EbXML compliant applications and components may also be commercially available as shrink-wrapped solutions [8].

Company A then submits its own Business Profile information (including implementation details and reference links) to the EbXML Registry (Fig. 1, step 3). The business profile submitted to the EbXML Registry describes the company's EbXML capabilities and constraints, as well as its supported business scenarios. These business scenarios are XML versions of the Business Processes and associated information bundles (e.g. a sales tax calculation) in which the company is able to engage. After receiving verification that the format and usage of a business scenario is correct, an acknowledgment is sent to Company A (Fig. 1, step 3) [8].

Company B (port, ship or any other related actor) discovers the business scenarios supported by Company A in the EbXML Registry (Fig. 1, step 4). Company B sends a request to Company A stating that they would like to engage in a business scenario using EbXML (Fig. 1, step 5). Company B acquires an EbXML compliant shrink-wrapped application.

Before engaging in the scenario Company B submits a proposed business arrangement directly to Company A's EbXML compliant software Interface. The proposed business arrangement outlines the mutually agreed upon business scenarios and specific agreements. The business arrangement also contains information pertaining to the messaging requirements for transactions to take place, contingency plans, and security-related requirements (Fig.1, step 5). Company A then accepts the business agreement. Company A and B are now ready to engage in e-business using EbXML (Fig. 1, step 6) [8].

3 Related Works

In the port of Marseille (France), according to [3] is used an information system for tracking goods, containers and ships. The connection is supported by an integrated and modular transfer communication platform, database and applications

shared by the elements of the transport chain through EDI technology. The port system is based on PROTIS - management of information flows for goods exporting and importing - and ESCALE - scale management of ships and port operations associated with the billing of services provided by the port. The purpose of the PROTIS system consists of the network provision and an application that facilitates the management of information flows relating to export and import of goods in terms of community port of Marseille added to the notice of call of the ship. The COMPAS system, which is used in the Spanish ports of Valencia and Las Palmas, enables electronic transfer of manifests of companies to ports and customs via EDI through a single window, before the arrival of ships in order to expedite the process of goods. Manifests are validated through customs that, via EDI, informs the agent of the consignee company about the result of the validation. This system uses the following EDI messages: a) request for mooring; b) notification of dangerous goods; c) booking; and d) shipload note.

The Dakosy system is the result of a project initiated by Hamburg city transport administration in order to increase efficiency and competitiveness of the port. To achieve this purpose, it has invested in infrastructure as well as equipment to automate port procedures. In terms of middleware technology, the project used EDIFACT combined with XML technology. Related services are monitoring of banking transactions, monitoring of dangerous goods and cargo handling [9].

In the port of Yokohama was implanted a port information system with EDIFACT and XML technologies to enable the exchange of information, faster access to cargo handling returns and facilitate cargo delivery through customs and control vessels. Additionally, the Port Authorities aimed to promote the reduction and provide accident prevention and the preservation of the environment. Among these procedures is the use of large-scale satellite communications, secure way of transmitting / receiving files / messages and one of the main ways of communicating with vessels. It is also done using digital cameras to record in real time port activities [1].

In [10] study, conducted in Korea, it was proposed business models for data exchange with content based on EbXML, which allow trading partners to perform trading business with operational safety. It suggests that it should continue research on the EbXML application system focusing on negotiation contents, digital exchanges and advanced security model, using Web service security. This research is supported by the Ministry of Culture, Sports and Tourism (MCST) in Korea and the creative Content Agency of Korea (KOCCA), using the technology of Culture (CT) in a development program started in 2011.

In [11] was studied the use of EbXML in the Hong Kong maritime port. This proposal is presented as an alternative to EDI, on the one hand because it incorporates positive experiences of many years of use of EDI leveraging the investments made, on the other hand, because it has the advantage of transforming the sector of maritime transport in a single market, global and cohesive. The main part of this work aimed at the automatic reception of dangerous goods manifest in XML format sent by shipping agencies.

The work [6] refers to the study and proposal for the Cabo Verde ports in Africa and it is reported as the most similar work to this proposal. Only differs from this work to use to exchange information Web Services, rather than EbXML, and be related to ports and realities of Cabo Verde. The author presents Web Services as an alternative to the integration of information systems, although Web Services may contain detectable flaws and problems. In [12], it was developed a tool named WS-FIT, which is a communication fault injector for testing Web Services calls that use the SOAP protocol. The tool provides an instrumented SOAP library that allows modifications in the messages exchanged between providers and Web Service clients, bypassing the difficulties using encryption with the lowest level of nozzles.

It is important to emphasize that the different technologies has been used in sea ports for different purposes. EbXML itself is used for ship loading, downloading, temporary storage of shiploads, among others processes, as internal port processes as integration of information. Therefore, other technologies may be used for the same purposes, such as EDI and Web Services.

This study differs from others in the literature to present system optimization results (reduction of interactions among system actors) through the analysis of use case flows using EbXML as an alternative to EDI.

4 Methodology

In this section it is presented the methodology applied in the work according to established purposes. In general, the development of this work occurred mainly through the reading of technical documents and especially technical specifications analysis of the port, standard EbXML and information technologies.

The EbXML reference documents were critical to the understanding of technology, such as [7, 8].

To understand the port processes, visits and meetings were held with Bonded Terminals of the Santos port, which released the document [13], which has the necessary information for use cases developed, addressed in this paper in the next section.

Based on operation studies of the Santos port, it was modified a part of the port system to be possible to replace the EDI-DTE technology to EbXML. This modification allowed improvements made in the system, enabling some port processes to be optimized.

Based on these modifications, use cases flows were analyzed, based on the methodology for software development RUP (Rational Unified Process) from IBM [14].

For the implementation it was used Microsoft .NET Framework 4.5.2 platform with the CSharp language. For the tests and EbXML Repository simulation it was used Microsoft SQL Server.

For the EbXML implementation in the system, a negotiation protocol including an XML Schema specific was created. For more information, the EbXML specification [8] can be consulted.

5 Results

The results are analyzed using use cases flows inspired by RUP form, as mentioned in the methodology section.

This technique was chosen to allow visibility of system parts where there were modifications in reason of replacement of EDI technology EbXML. The initial objective was to analyze and compare the number of interactions among different entities or actors in the system to verify the possible reduction of interactions through technology replacement.

For a better understanding of the use cases, system actors and their initials (used in the text of use cases – Tables 2 to 6) and descriptions are presented in Table 1.

Table 1 System Actors

Actor	Description
Customs Authorities (CA)	Customs authorities are an official government agency that controls the movement of inputs and outputs for the goods overseas or from responsible, including the collection of relevant taxes.
Shipping Agency (SA)	Agency of cargo in foreign trade.
CODESP	It Brazilian state company, incorporated as a joint stock company, created on 7 November 1980 to replace the Companhia Docas de Santos, who was privatized. Responsible for management of the Santos port.
Bonded Enclosure (BE)	Are considered Bonded Enclosure: I) From primary zone, patios, warehouses, terminals and other places for the movement and storage of goods imported or for export; II) Of the secondary zone, warehouses, inland customs stations, warehouses, terminals or other units for the storage of goods under the conditions of the previous item.
Port Operator (PO)	Pre-qualified legal person performing operations in the port area of the organized point. Receives at its pier (port) the ship.

Table 2 presents a comparison of the use case flows called “Transmit mooring data”.

Table 2 Use Case Comparison – "Transmit mooring data"

Precondition	
1. There are ships for mooring.	
Basic Flow - EDI	**Basic Flow - EbXML**
1.CODESP records in the system the list of RAPs (Request for Mooring and Priority) containing the goods arrival information informing the vessel code (Lloyd) and data arrival of the goods (mooring forecast). 2.CODESP sends the information of the RAPs to the PO. 3.PO sends the information of the RAPs to CA. 4.CA sends the information of the RAPs to BE. 5. SA transmits release response of the goods to the system. 6.CODESP consults whether the goods are released in the system. 7. CODESP records the execution information of the ship's mooring. 8.CODESP records the ship undocking information. 9. BE records cargo delivery information. 10. BE informs the delivery record charge for CA. 11.SA sends manifest to CODESP and PO. 12.CODESP and PO register manifest information in the system. 13.CODESP and PO inform the CA on the manifest record. 14. CA consults information delivery shipload and manifest in the system to perform the DT (Transfer Declaration) closure. 15.CA communicates closure of DT to CODESP and PO. 16. End of use case.	1.CODESP records in the system the list of RAPs (Request for Mooring and Priority) containing the goods arrival information informing the vessel code (Lloyd) and data arrival of the goods (mooring forecast). 2. System transmits mooring information to CA, BE and PO. 3. CA transmits response to release of the goods to the system. 4. System transmits information release of the goods for CODESP. 5. CODESP records the execution information of the ship's mooring. 6. CODESP records the ship undocking information. 7. BE records cargo delivery information. 8. System transmits delivery record information load to CA. 9. SA sends the manifest for the system. 10. System responds to the CODESP, CA and PO, informing the receiving of manifest. 11. CA consults the information of delivery of cargo and manifest in the system to perform the DT closure. 12. CA transmits DT closure information response to the system. 13. System transmits DT closure information for CODESP and PO. 14. End of use case.
Alternative Flow - EDI	**Alternative Flow - EbXML**
5.1. In case of denial of release, CA indicates the shipload lock, which must remain available for inspection purposes and physical inspection. 5.2. If the shipload is released returns to the item 5.	3.1. In case of denial of release, CA indicates the shipload lock, which must remain available for inspection purposes and physical inspection. 3.2. If the shipload is released returns to the item 3.

In Table 2, observe that there is reduction among actor's interactions. The actors SA, BE, PO and CA receive directly from CODESP data mooring using EbXML protocol, that enables communication simultaneously to the different entities in the system, as opposed to EDI, who works with point to point data communication.

Table 3 presents the second use case, called "Create and transmit RAPs."

Table 3 Use Case Comparison – "Create and Transmit RAPs"

Precondition	
1. Mooring data is up to date.	
Basic Flow - EDI	**Basic Flow - EbXML**
1. SA analyzes and sends data of the RAPs to CODESP. 2. CODESP creates the RAPs in the system and updates the fields: ship code, voyage number and estimated arrival of the ship in port. 3. CODESP sends the information of the RAPs to the PO. 4. PO sends information of the RAPs to CA. 5. CA sends the information of the RAPs to EB. 6. End of use case.	1. SA analyzes and sends data of the RAPs to the system. 2. System creates RAPs using data sent by the SA. 3. System communicates CODESP the creation of the RAPs. 4. CODESP accesses the system and updates the RAPs in the following fields: ship code, voyage number and estimated arrival of the ship in port. 5. CODESP requests in the system the transmission of the RAPs updated to CA. 6. End of use case.
Alternative Flow - EDI	**Alternative Flow - EbXML**
No alternative flow.	1.1 In the case of data inconsistency, the system sends message to SA and notifies the user.

In the use cases described in Table 3, it is noticed that the dependence of actors was reduced in "Basic Flow – EbXML" in comparison with "Basic Flow – EDI". Besides, in the basic flow of EbXML use case, BE (actor) could be eliminated. Finally, the system with EbXML executes natural validation (through XML Schemas) that prevents the information from being transmitted inconsistently.

Table 4 shows the third use case under consideration, called "Close DT".

Table 4 Use Case Comparison – "Close DT"

Precondition	
1. Use case "Create and transmit RAPs" was executed.	
Basic Flow - EDI	**Basic Flow - EbXML**
1.CA receives RAPs data from PO. 2. "DT and complementary DT" are automatically closed by system after receiving RAPs data. 3. CA sends the "DT closing" and "DT complementary" information to PO. 4. PO sends the "DT closing" and "DT complementary" information to CODESP. 5. PO updates in the system the status of untied shipload and/or containers for "Authorized Delivery". 6. PO sends authorization of delivery of untied shipload and/or containers to BE. 7. BE sends receiving answer to PO. 8. PO modifies the delivery status to "Done". 9. End of use case.	1. CA receives RAPs data from PO. 2. "DT and complementary DT" are automatically closed by system after receiving RAPs data. 3. CA sends the "DT closing" and "DT complementary" information to the system. 4. System communicates CODESP and PO about "DT closing" and "DT complementary". 5. System updates the status of untied shipload and/or containers for "Authorized Delivery". 6. System sends authorization of delivery of untied shipload and/or containers to BE. 7. BE sends receiving answer to PO. 8. System modifies the delivery status to "Done". 9. End of use case.
Alternative Flow - EDI	**Alternative Flow - EbXML**
7.1. In the case of response information is different from "Received": 7.2. BE warns the PO to undertake changes in shipload information related to closed DTs. 7.3. PO updates the system shipload information as well as status update to "Cancelled". 7.4. The operation is terminated.	7.1. In the case of response information is different from "Received": 7.2. System notifies PO requesting to take changes in load information related to closed DTs. 7.3. PO updates the system shipload information as well as status update to "Cancelled". 7.4. The operation is terminated.

As mentioned in "Basic Flow – EbXML" (Table 4), steps 3 and 4, there is decrease of user interventions. The system simultaneously performs several updates and broadcasts messages from the data sent by other actors.

In the case of "Basic Flow – EDI" (Table 4), step 5, it is necessary for the PO to update the system with some information. In the "Basic Flow – EbXML", it is possible to automate the system due to the simultaneous transmission of EbXML information, which also enables a data consistency checking.

Table 5 Use Case Comparison – "Deliver Manifest"

Precondition	
1. Shipload data is up to date.	
Basic Flow - EDI	**Basic Flow - EbXML**
1. SA sends manifest data to PO. 2. PO sends the manifest data to CODESP. 3. CODESP creates manifests in the system. 4. CODESP informs the PO about the creation of manifest in the system for mooring control purposes. 5. End of use case.	1. SA sends manifest data to the system. 2. System creates the manifest automatically. 3. System announces the creation of the manifest to CODESP and PO. 4. End of use case.
Alternative Flow - EDI	**Alternative Flow - EbXML**
No alternative flow.	No alternative flow.

As showed in Table 5, there is reduction of manual labor and interaction among the actors, illustrated by elimination of items 1, 2 and 3 of the "Basic Flow – EDI", that have been replaced by items 1 and 2 of the "Basic Flow – EbXML". The user advantage is the information transmission without his intervention. This is possible due data consistency checking and simultaneous data communication.

Table 6 Use Case Comparison – "Cancel Form"

Precondition	
1. There is shipload to dispatch.	
Basic Flow - EDI	**Basic Flow - EbXML**
1. BE informs to PO about the cancellation of the form. 2. PO updates the system with the information about cancellation. 3. PO informs to CODESP with the information about cancellation. 4. End of use case.	1. BE informs to PO about the cancellation of the form. 2. System records the cancellation of the forms. 3. System informs about the cancellation of forms to CODESP and PO. 4. End of use case.
Alternative Flow - EDI	**Alternative Flow - EbXML**
No alternative flow.	No alternative flow.

According Table 6, in the "Basic Flow – EbXML", there is reduction of the manual labor of actors, as exposed by items 2 and 3. The system itself informs the actors about information records and does not require human intervention. As modifications described in other use cases, this fact occurs because of data consistency checking and simultaneous data communication.

6 Conclusions

In the use cases described in this paper was possible to evaluate process improvement in the port.

In attested cases, it was evident that the tool use case flows proposed in RUP is a great tool to describe the usability modification made by replacing a data communication technology (middleware).

The study confirms that the substitution of EDI technology by EbXML is feasible and can be applied to other port information systems with the advantage of optimizing the user operations or improving port processes.

Acknowledgment The authors gratefully acknowledge the academic support and research structure from the UFABC (Federal University of ABC) and IFSP (Federal Institute of Education, Science and Technology of São Paulo). Furthermore, the financial support of IFSP, FAPESP (São Paulo Research Foundation) and CNPq (National Council for Scientific and Technological Development).

References

1. Kurosawa, R.S.S.: Information Systems Analysis Applied to Port Management. Thesis (PHD in Naval Engineering) — University of São Paulo, São Paulo (2012)
2. Torres, L.F.R.: Analytical and Operational Study of Electronic Data Superhighway: An Electronic Management Model for the Brazilian Ports. (PhD in Energy Engineering and Electrical Automation) – University of São Paulo, São Paulo (2007)
3. Conférence des Nations Uies sur le Commerce et le Développment Conférence des Nations Uies sur le Commerce et le Développment (CNUCED): The modern management of ports. Port management certificate, Edition of 9th May 1997, revision May 2005
4. Beam, C., Segev, A.: Automated Negotiations: A Survey of the State of the Art. Wirtschaftsinformatik **39**(3), 263–268 (2001)
5. Baarslag, T., Hindriks, K., Hendrikx, M., Dirkzwager, A., Jonker, C.: Decoupling negotiating agents to explore the space of negotiation strategies. In: Marsa-Maestre, I., Lopez-Carmona, M.A., Ito, T., Zhang, M., Bai, Q., Fujita, K. (eds.) Novel Insights in Agent-based Complex Automated Negotiation. LNCS, vol. 535, pp. 61–83. Springer, Japan (2014)
6. Carvalho, J.G.M.: System port information management. 2010. 103 f. Dissertation (Msc in Electronic Engineering and Telecommunications) - Department of Electronics, Telecommunications and Informatics, University of Aveiro (2010)
7. EbXML website. http://www.ebxml.org/
8. EbXML Specification, EbXML Technical Architecture Specification v.1.0.4 (2001)
9. Posti, A., Häkkinen, J., Tapaninen, U.: Promoting information exchange with a port community system – case Finland. In: Kersten, W., Blecker, T., Jahn, C. (eds.) International Supply Chain Management and Collaboration Practices, vol. 4, pp. 455–471. BoD, Germany (2011)

10. Shin, D., Shin, D.: Secure Business Transaction Models For Trading Game Contents Based On EbXML Applying Web Service Security Standards. International Journal of Software Engineering & Its Applications **7**(3) (2013)
11. Center for E-commerce Infrastructure Development (CECID), The University of Hong Kong (2002). http://www.cecid.hku.hk/newsletterVOL1Sep2002.php?
12. Looker, N., Munro, M., Xu, J.: WS-FIT: a tool for dependability analysis of web services. In: IEEE 28th Annual International Computer Software and Applications Conference, vol. 2, pp. 120–123. IEEE Press, New York (2004)
13. Brazilian Association of Terminals and Enclosures (ABTRA): Transfer Request - DTE System. Operating manual. Santos (2008)
14. Kruchten, P.: Rational Unfied Process best practices for software development teams. Rational Software, Canada (2001)

Reversible Data Hiding Scheme Using the Difference Between the Maximum and Minimum Values

Pyung-Han Kim, Kwang-Yeol Jung, In-Soo Lee and Kee-Young Yoo

Abstract Qu et al. proposed a PPVO scheme based reversible data hiding technique in 2015. Their scheme obtains the predicted error value by using context pixels, and uses current pixel, predicted error value and block pixels that except for the current pixel. The PPVO scheme has a complex process because of embedding and extraction methods. In this paper, we propose an improved reversible data hiding scheme by using the difference values between the maximum value and minimum value. In our scheme, the difference values are used for a secret message embedding and extraction. Also, our scheme satisfies the characteristic of reversible data hiding. In experimental results, the embedding capacity and image quality of the proposed scheme are superior in comparison with Qu et al.'s PPVO scheme.

Keywords Reversible data hiding · Steganography · Difference value · Minumum-value-standard · Maximum-value-standard

1 Introduction

With the development of computer hardware and network technique, information transmission via Internet is widely using in our life. However, the attacker is easy access to the private information in the internet environment and it can modulate the important information. Thus, there exist some security problem to vulnerable

P.-H. Kim · K.-Y. Jung · K.-Y. Yoo(✉)
School of Computer Science and Engineering, Kyungpook National University,
Daegu 702-701, South Korea
e-mail: k2jbks90@infosec.knu.ac.kr, yook@knu.ac.kr

I.-S. Lee
School of Electronics and Engineering, Kyungpook National University,
Daegu 702-701, South Korea

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_33

by the attacks such as interruption, modification and fabrication. To overcome these problems, the safe information exchange methods are necessary to protect the private information from malicious attackers. These information exchange methods are classified as follows: cryptography and data hiding. Data hiding is a method that embeds secret message into digital multimedia such as videos, audios, images [1–5]. In the image data hiding method, the original image is called the cover image, and stego image is created by embedding the secret message into the cover image. Data hiding method has a problem that during embedding secret message, the cover image is distorted. This distortion is a very serious problem in sensitive areas such as military, medical or artwork. Therefore, the reversible data hiding method is proposed to solve the problem of data hiding [6–13].

The reversible data hiding method recovers the lossless cover image after extracting secret messages. Representatively, the histogram shifting and the difference expansion(DE) methods are famous methods. The histogram shifting scheme generates the histogram based on the frequency of the pixel values, and the secret message is embedded into the peak point [14]. This scheme is creating an image of good quality, but the embedding capacity is limited. The DE method is proposed by Tian in 2003 [15]. This scheme embeds the secret message by expanding difference values for increasing the hiding capacity. However, the image quality is decreased according to changes of all pixel values.

In 2013, a new reversible data hiding scheme was proposed by Li et al. [16]. This scheme uses the pixel value ordering(PVO) to calculate the prediction error value. Also, the secret message is embedded by using the prediction-error expansion(PEE). However, this scheme has a problem that the prediction error value is limited by 1 or -1. To overcome this problem, Peng et al. proposed an improved scheme [17]. Their scheme embeds a secret message by using the prediction error value 0 besides 1 or -1. Even though the several methods was proposed, these methods ignored many pixels in the cover image. In order to solve these problems, Qu et al. proposed an improved scheme [18]. Their scheme uses the context pixels for obtaining prediction error value, and the secret message is embedded by using prediction error value. Although Qu et al.'s scheme has better performance than previous schemes, their schemes has some problems. Their scheme is a bit complex in the embedding and extraction process, and ought to remember the location of pixels after ascending order for cover image restoration. To overcome these problems, we propose an improved scheme. The proposed scheme calculates the difference value between the maximum value and minimum value, and uses the difference value during embedding and extraction procedure for recovering lossless cover image. The constitution of this paper is as follows. In Section 2, PVO, IPVO and PPVO schemes are introduced. Section 3 explains the proposed scheme. Experimental results and analysis are given in Section 4. Section 5 mentions conclusions and future works.

2 Related Works

2.1 PVO Scheme

In 2013, Li et al. proposed a new reversible data hiding scheme based on the PEE and PVO schemes. Their scheme obtained the prediction error value by using a pixel value ordering, and they embedded the secret message by utilizing prediction error expansion. When the pixels are denoted as b_1,b_2,...,b_n and the sorted pixels are defined as $b_{\sigma(1)}$,$b_{\sigma(2)}$,...,$b_{\sigma(n)}$, maximum prediction-error value is calculated by Eq. (1).

$$PE_{max} = b_{\sigma(n)} - b_{\sigma(n-1)} \tag{1}$$

If a stego image pixel value is $\tilde{b}$, the secret bit s is embedded by using Eq. (2).

$$\tilde{b} = \begin{cases} b_{\sigma(n)} & \text{if } PE_{max} = 0 \\ b_{\sigma(n)} + s & \text{if } PE_{max} = 1 \\ b_{\sigma(n)} + 1 & \text{if } PE_{max} > 1 \end{cases} \tag{2}$$

2.2 IPVO Scheme

Li et al.'s scheme embeds the secret message by using the prediction error values 1 and -1, whereas not using the prediction error value 0. To improve this weakness, Peng et al.'s proposed an improved scheme that uses the prediction error value 0 in 2014. Their scheme is a little different compared with Li et al.'s scheme in calculation process of prediction error value because of using the Eq. (3).

$$d_{max} = x_u - x_v \tag{3}$$

where,

$$\begin{cases} u = min(\sigma(n), \sigma(n-1)) \\ v = max(\sigma(n), \sigma(n-1)) \end{cases} \tag{4}$$

Due to the Eq. (3), Peng et al.'s scheme has more prediction error value 0 than Li et al.'s scheme. Based on this characteristic, a new prediction error method is proposed for using the prediction error value 0. The secret message is embedded by using the Eq. (5).

$$\tilde{x} = \begin{cases} x_{\sigma(n)} + b \text{ if } d_{max} = 1 \\ x_{\sigma(n)} + 1 \text{ if } d_{max} > 1 \\ x_{\sigma(n)} + b \text{ if } d_{max} = 0 \\ x_{\sigma(n)} + 1 \text{ if } d_{max} < 0 \end{cases} \tag{5}$$

2.3 PPVO Scheme

Recently, a high-fidelity reversible data hiding scheme was proposed based on pixel-based pixel value ordering by Qu et al. Their scheme obtains the predicted error value by using context pixels, and uses the predicted error value and embedding algorithm in order to generate the stego image. In Fig. 1, the context pixels are shown.

X	C_1	C_4
C_2	C_3	C_6
C_5	C_7	C_8

Fig. 1 The context pixels of pixel X in PPVO scheme

The CN defined as a number of pixels except for X in the context pixels. If CN value is 8, then context pixel vecter $C = C_1,...,C_8$. Qu et al.'s embedding algorithm is shows as follows. In the embedding algorithm, maximum value and minimum value are defined as Max and Min, respectively. VC mean that all pixels have the same pixel values except for current pixel.

1) Embedding Algorithm

Input: Current pixel value x, predicted value $\hat{x}$ and secret bit b
Output: Stego pixel value $\tilde{x}$

```
if Max ≠ Min
  if x ≤ Min
    if x = x̂
      x̃ = x − b;
    else if x < x̂
      x̃ = x − 1;
  else if x ≥ Max
    if x = x̂
      x̃ = x + b;
    else if x > x̂
      x̃ = x + 1;
else
  if VC = 254
   x̃ = x + b;
  else
    if x = x̂
      x̃ = x − b;
    else if x < x̂
      x̃ = x − 1;
return x̃
```

Although Qu et al.'s scheme utilizes more pixels than previous schemes, their scheme has more complexity in the embedding and extraction process, and ought to remember the location of pixels after ascending order for cover image restoration. To solve these problems, we proposed an improved scheme in the next section.

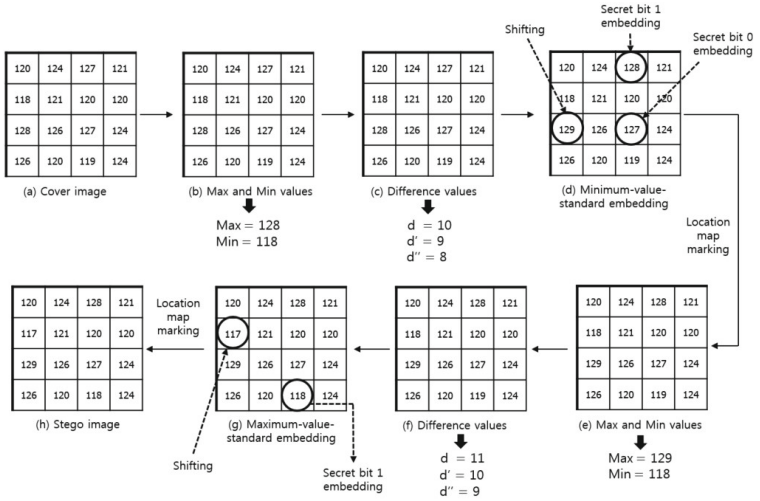


Fig. 2 Embedding example

3 Proposed Scheme

In this section, we propose an improved scheme by using difference values between the maximum and minimum values. The proposed scheme divided into non-overlapping blocks the cover image. And, the proposed scheme utilizes three different values for determining embedding process and extraction process, and the secret message embedded or extracted according to two standards. Two standards are classified as maximum-value-standard embedding or extraction and minimum-value-standard embedding or extraction methods. In the embedding and extraction process, we use the second or third maximum and minimum values for a secret message embedding and extraction. In the case of using the second maximum and minimum values, the first maximum and minimum values are shifting to prevent the failure of the extraction process. In the embedding process, the minimum-value-standard embedding is performed as a priority. On the contrary to this, the maximum-value-standard extraction is carried out as a priority in the extraction process. When the pixel value in the block of cover image is denoted as $C_{i,j}$ and the pixel value in the block of stego image is denoted as $S_{i,j}$, the maximum and minimum values are defined as Max and Min. More detalis are described below.

3.1 Embedding Procedure

The embedding process consists of a difference value calculation step and a secret message embedding step. The embedding conditions are decided by Max, Min and three difference values. The embedding algorithm is discussed below.

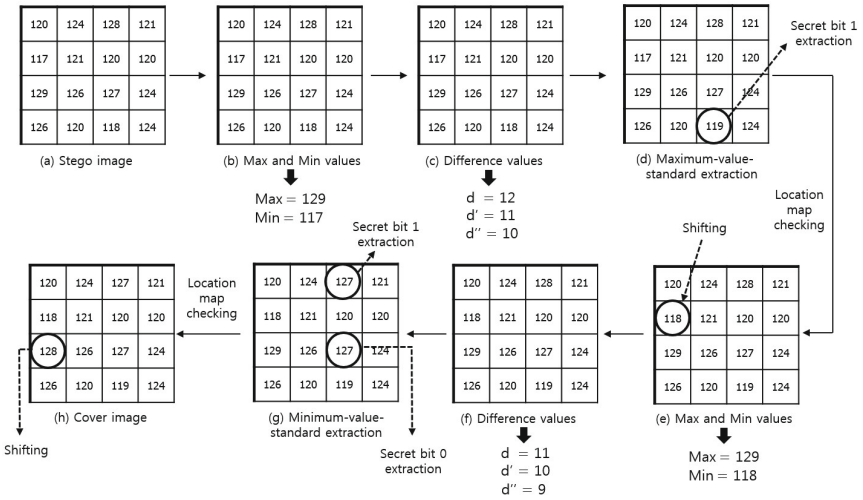


Fig. 3 Extraction example

1) Embedding Algorithm

Input: A cover image C and a secret message s
Output: A stego image S and the location map
Step 1: The cover image is divided into non-overlapped $n \times n$ blocks.
Step 2: Minimum-value-standard embedding for a block $C_{i,j}^{l}$.
(2-1) Find the Max and Min in the block.
(2-2) Calculate difference values d, d', d'' as follows.

$$\begin{cases} d = Max - Min \\ d' = d - 1 \\ d'' = d - 2 \end{cases} \tag{6}$$

(2-3) Embed a secret bit s_k.
if $(Min + d') == C_{i,j}^{l}$
1) $Max = Max + 1$ and marks a location map.
2) $S_{i,j}^{l} = C_{i,j}^{l} + s_k$
else if $(Min + d'') == C_{i,j}^{l}$
1) $S_{i,j}^{l} = C_{i,j}^{l} + s_k$
Step 3: Maximum-value-standard embedding for a block $C_{i,j}^{l}$.
(3-1) Find the Max and Min in the block.
(3-2) Calculate difference values by using Eq. (6).
(3-3) Embed a secret bit s_k.
if $(Max - d') == C_{i,j}^{l}$
1) $Min = Min - 1$ and marks a location map.
2) $S_{i,j}^{l} = C_{i,j}^{l} - s_k$

else if $(Max - d'') == C^l_{i,j}$

1) $S^l_{i,j} = C^l_{i,j} - s_k$

Step 4: Repeat step 2 and step 3 for the next block until all secret bits are embed to the cover image to generate the stego image.

3.2 Extraction and Recovery Procedure

The extraction process consists of a difference value calculation step, a secret message extraction step and a cover image recovery step. The extraction conditions are decided by maximum values, minimum values and three difference values, in common with the embedding condition. By extraction conditions, a secret bit 1 or 0 is extracted from the stego image. The extraction algorithm is discussed below.

2) Extraction Algorithm

Input: A stego image S and the location map
Output: A cover image C and a secret message s
Step 1: The stego image is divided into non-overlapped $n \times n$ blocks.
Step 2: Maximum-value-standard extraction for a block $S^l_{i,j}$.
(2-1) Find the Max and Min in the block.
(2-2) Caculate difference values using Eq. (6).
(2-3) Extract secret message and recover cover image.
if $(Max - d') == S^l_{i,j}$
1) Extract a secret bit 1.
2) The pixel values increase by 1 in the case of equal to $(Max-d')$.
else if $(Max - d'') == S^l_{i,j}$
1) Extract a secret bit 0.
(2-4) Location map checking and Min shifting.
Step 3: Minimum-value-standard extraction for a block $S^l_{i,j}$.
(3-1) Find the Max and Min in the block.
(3-2) Caculate difference values using Eq. (6).
(3-3) Extract secret message and recover cover image
if $(Min + d') == S^l_{i,j}$
1) Extract a secret bit 1.
2) The pixel values decrease by 1 in the case of equal to $(Min+d')$.
else if $(Min + d'') == S^l_{i,j}$
1) Extract a secret bit 0.
(3-4) Location map checking and Min shifting.
Step 4: Repeat step 2 and step 3 for the next block until all secret bits extracts, and cover image is restoration.

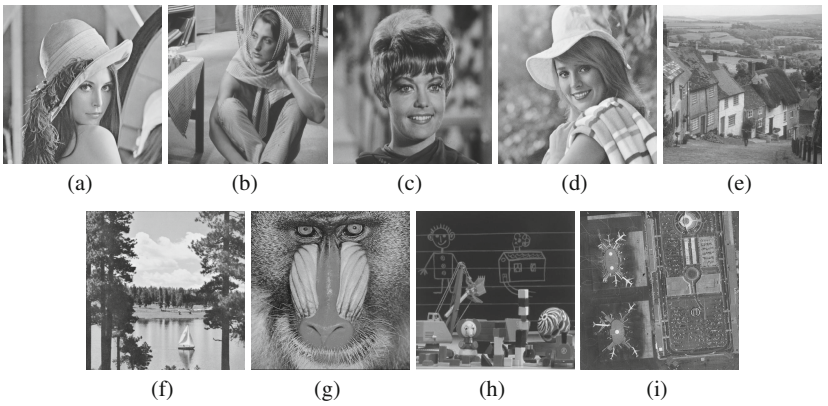


Fig. 4 Nine 512 × 512 size gray-scale images: (a) Lena (b) Babara (c) Zelda (d) Elaine (e) House (f) SailBoat (g) Baboon (h) Toy and (i) Airport

Table 1 The comparison of $PSNR$

	Block size 2 × 2			Block size 4 × 4		
Schemes	PVO	PPVO	The proposed	PVO	PPVO	The proposed
Lena	51.73	52.10	56.93	58.29	55.75	60.14
Babara	51.60	52.59	57.71	58.16	56.50	60.95
Zelda	47.39	52.37	52.38	52.76	55.94	53.25
Elaine	49.67	53.05	58.17	52.71	57.47	61.52
House	52.16	51.62	56.76	58.74	54.59	59.39
SailBoat	51.59	52.92	57.66	58.11	57.16	61.00
Baboon	48.26	53.49	43.66	52.75	58.06	42.11
Toy	52.09	51.88	55.06	58.80	55.92	57.93
Airport	47.34	52.96	57.66	48.84	57.71	61.76
Average	50.20	52.55	55.17	55.46	56.57	57.56

4 Experiment Result

In this section, we have compared the experimental results in that the proposed scheme, Li et al.'s scheme and Qu et al.'s scheme in order to prove the superiority of the proposed scheme. The image quality and embedding capacity are used to evaluate the performance of the proposed algorithm in the data hiding technique. In general, the image quality is measured by using $PSNR(Peak\ to\ Signal\ Noise\ Ratio)$, and embedding capacity is counting on the total amount of secret message in the stego image. $PSNR$ means the degree of image distortion that is defined as follows

$$PSNR = \left(10 \cdot \log_{10}\left(\frac{255^2}{MSE}\right)\right), \tag{7}$$

Table 2 The comparison of capacity

	Block size 2 × 2			Block size 4 × 4		
Schemes	PVO	PPVO	The proposed	PVO	PPVO	The proposed
Lena	27,306	47,162	56,521	8,085	20,731	28,542
Babara	22,630	37,050	46,991	6,657	17,392	12,990
Zelda	27,666	44,542	53,474	7,868	20,173	26,027
Elaine	20,630	28,635	41,446	6,156	13,394	19,397
House	29,043	70,375	68,387	7,622	38,747	43,056
SailBoat	21,858	30,970	47,638	6,622	13,343	23,736
Baboon	13,127	16,338	24,215	6,657	7,201	23,954
Toy	39,969	59,101	90,393	10,542	26,802	54,833
Airport	21,463	29,488	41,067	7,336	18,802	27,972
Average	24,855	40,407	52,237	7,336	18,802	27,972

The above equation is measured by log units and determined on the ratio of MSE. MSE is defined as follows

$$MSE = \left(\sum_{i=1}^{w \times h} \frac{(p_i - p_i')^2}{w \times h} \right), \tag{8}$$

The width and height are defined as w and h in the cover image, and p_i and p_i' represent pixel values of cover and stego image, respectively. In the case that $PSNR$ is 30dB or more, it is impossible to detect the distortion of the image by human eyes. The embedding capacity is determined by $bpp(bit\text{-}per\text{-}pixel)$. The test images are 512×512 gray scale images as shown in Fig. 4. The secret message is randomly generated by random function.

In this section, we compared Li et al.'s PVO scheme, Qu et al.'s PPVO scheme and the proposed scheme. Table 1 and Table 2 show the comparison of $PSNR$ and embedding capacity between the previous schemes and the proposed scheme. In the case of block size 2×2, the average $PSNR$ and capacity are superior in comparison with the other schemes. The $PSNR$ value of the proposed scheme is increased by the 2.62dB compared with PPVO scheme. The embedding capacity of the proposed scheme is increased by 29% compared to the embedding capacity of PPVO scheme. In the case of block size 4×4, the $PSNR$ value of the proposed scheme is increased by the 0.99dB compared with PPVO scheme, and embedding capacity is increased by 48%. As a result, the proposed scheme is outstanding in comparison with other schemes.

5 Conclusion

Reversible data hiding method is used in a variety of ways for securely information transmission. The proposed scheme uses the difference value to obtain a better

performance to compared with previous schemes. In the embedding step, the difference value is calculated by using the maximum value and the minimum value in the block. The secret message is embedded by using the maximum value, the minimum value, and difference values. Due to using the difference value in the extraction step, the proposed scheme fulfills the characteristic of reversible data hiding. In our experimental results, the capacity of the proposed scheme is increased by 29% compared to the capacity of Qu et al.s scheme, and $PSNR$ is increased by 2.62dB. As a result, the embedding capacity and image quality of the proposed scheme are superior in comparison with previous schemes.

In the future work, we plan to group the difference values for increasing the capacity and to adjust the difference values for obtaining the superb image quality.

Acknowledgement This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01060801) and supported by the BK21 Plus project(21A20131600005).

References

1. Johnson, F.N., Jajodia, S.: Exploring Steganography: Seeing the Unseen. Computer **31**, 26–34 (1998)
2. Petitcolas, F.A.P., Anderson, R.J., Kuhn, M.G.: Information hiding-a survey. Proceedings of IEEE Special Issue on Protection of Multimedia Content **87**, 1062–1078 (1999)
3. Johnson, N.F., Duric, Z., Vanstone, S.J.A.: Information hiding: Steganography and Watermarking - Attacks and Countermeasures. Kluwer Academic Publishers (2001)
4. Bai, L., Biswas, S., Blasch, P.E.: An estimation approach to extract multimedia information in distributed steganographyic images. In: Proc. 10th International Conference on Information Fusion. IEEE (2007)
5. Cheddad, A., Condell, J., Curran, K.: Mc Kevit, P.: Digital image steganography: Survey and analysis of current methods. Signal Processing **90**, 727–752 (2010)
6. Awrangjeb, M.: An overview of reversible data hiding. In: Proceedings of the Sixth International Conference on Computer and Information Technology, pp. 75–79 (2003)
7. Tsai, P., Hu, Y.C., Yeh, H.L.: Reversible Image Hiding Scheme Using Predictive Coding and Histogram Shifting. Signal Processing **89**, 1129–1145 (2009)
8. Gui, X., Li, X., Yang, B.: A high capacity reversible data hidiing scheme based on generalized prediction-error expansion and adaptive embedding. Signal Processing **98**, 370–380 (2014)
9. Wu, X., Sun, W.: High-capacity reversible data hiding in encrypted images by prediction error. Signal Processing **104**, 387–400 (2014)
10. Al-Qershi, O.M., Khoo, B.E.: Two-dimensional difference expansion(2D-DE) scheme with a characteristics-based threshold. Signal Processing **93**, 154–162 (2013)
11. Tai, W.L., Yeh, C.M., Chang, C.C.: Reversible Data Hiding Based on Histogram Modification of Pixel Differences. IEEE Transactions on Circuits and Systems for Video Technology **19**, 906–910 (2009)
12. Luo, H., Yu, F.X., Chen, H., Huang, Z.L., Li, H., Wang, P.H.: Reversible data hiding based on block median preservation. Information Sciences **181**, 308–328 (2011)
13. Fu, D.S., Jing, Z.J., Zhao, S.G., Fan, J.: Reversible data hiding based on prediction-error histogram shifting and EMD mechanism. AEU-International Journal of Electronics and Communications **68**, 933–943 (2014)
14. Ni, Z., Shi, Y.U.N.Q., Ansari, N., Su, W.E.I.: Reversible data hiding. IEEE Transactions on Circuits and Systems for Video Technology **16**, 354–362 (2006)

15. Tian, J.: Reversible data embedding using a difference expansion. IEEE Transactions on Circuits and System for Video Technology **13**, 890–896 (2003)
16. Li, X., Li, J., Li, B., Yang, B.: High-delity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion. Signal Process **93**, 198–205 (2013)
17. Peng, F., Li, X., Yang, B.: Improved pvo-based reversible data hiding. Digital Signal Process **25**, 255–265 (2014)
18. Qu, X., Kim, H.: Pixel-based pixel value ordering predictor for high-fidelity reversible data hiding. Signal Process **111**, 249–260 (2015)

The Design, Data Flow Architecture, and Methodologies for a Newly Researched Comprehensive Hybrid Model for the Detection of DDoS Attacks on Cloud Computing Environment

Anteneh Girma, Kobi Abayomi and Moses Garuba

Abstract As cloud computing services are becoming more practical and popular for the sake of its convenience and being more economical, its' security vulnerability has been the continuous threat for both cloud service providers and clients. The more the financial benefits of these services became so attractive and the need for uninterrupted services grows, the distributed denial of services (DDoS) attacks that degrade and down its' service availability has been the major security concern. While researchers try to address this security threat and come up with lasting solutions for early detection of DDoS attacks, the degree of these attacks is getting higher and very sophisticated. The changing and aggressive nature of the attacks make it very severe threat and difficult to easily find remedy. On this research paper, we are presenting the design and dataflow architectures for a solution model that could contribute in resolving one of the major cloud security threat, the early detection of DDoS attacks, using its hybrid approach.

Keywords Cloud security · Cloud design architecture · Detection of DDoS attacks · Conditional entropy

1 Introduction

A Distributed Denial of Service (DDoS) attack is a flooding attack on a certain target network or server system that is launched through many compromised

A. Girma(✉) · M. Garuba
Computer Science Department, Howard University, Howard, USA
e-mail: agirma@howard.edu

K. Abayomi
Mathematics Department, Howard University, Howard, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_34

systems called zombies. DDoS attacks are launched with the intention of service disruption by depleting either the network resources or the host servers. DDoS attacks are caused by sending flood of requests and preventing legitimate users from accessing network services or data center resources. In order for attackers to create large botnets of computers (Zombies) under their control, they have two options: the more common option of using specialized malware to infect the machines of users who are unaware that their machines are compromised, or the relatively newer option of amassing a large number of volunteers willing to use DDoS programs.

Anytime attackers want to launch a DDoS attack, they can send messages to their botnet's C&C servers with instructions to perform an attack on a particular target, and any infected machines communicating with the contacted C&C server will comply by launching a coordinated attack.

2 Related Works

Many researches have been conducted and as many number of different DDoS detection techniques have been proposed. Among these was a simple and efficient hidden markov model scheme for host based anomaly intrusion detection [1]. An entropy based anomaly detection system to prevent DDoS attacks in cloud was reviewed, explored, investigated and proposed as an alternative solution [2]. After investigating the correlativity changes of monitored network features during flood attacks, a covariance-Matrix modelling and detecting various flooding attacks was proposed [3].

An experimented result was also analyzed and presented to support a model that was instrumental to propose a model to detect flood based DoS attack in cloud environment. It provided research results that support how effectively the flood attacks are detected [4]. Researchers also discussed how entropy based collaborative detection of DDoS attacks on community networks could effectively works in theory by applying information theory parameter called entropy rate [5]. Different types of DDoS attacks at different layers of OSI model were discussed and presented, and finally, analyzed the impact of DDoS attacks on cloud environment [6]. The analysis of covariance model for DDoS Detection was discussed and the researchers described how the method can effectively differentiate the traffic between the normal and attack traffic. They also showed how the linear complexity of the method makes its real time detection practical [7].

Another detecting solution framework to predict multi-step attacks before they pose a serious security risk is by using hidden markov model. The study based the real time intrusion prediction on optimized alerts since alerts correlations play a critical role in prediction [8]. The design of two independent architectures for HTTP and FTP which uses an extended hidden semi-markov model to describe the browsing habits of web searchers and detecting DDoS attacks were discussed

and investigated [9]. A survey of different mechanism of DDoS attacks, its detection, and the various approaches to handle them was discussed and explored, to enable the clients review and understand those different parameters having impacts in their decision making process while selecting the right DDoS detecting scheme [10].

The scopes of DDoS flooding attack problems and attempts to combat them have been explored by categorizing the DDoS flooding attacks and classifying existing countermeasures based on different parameters [11]. A comprehensive survey presented DDoS attacks, detection methods, detection tools used in wired networks and internet, and future research direction [12]. The Security problem associated with cloud computing becomes more complex due to entering of new dimensions in problem scope related to its own main attributes. Researchers also proposed a detection scheme based on the information theory based metrics. The proposed scheme has two phases: Behavior monitoring and Detection. Based on the observation, Entropy of requests per session and the trust score for each user is calculated [13]. DDoS attacks could be detected using the application of Dempster Shafer Theory. The theory was applied to detect DDoS threat in cloud environment. It is an approach for combining evidence in attack conditions [14]. The effectiveness of an anomaly based detection and characterization system highly dependent on accuracy of threshold value setting. And this approach described a novel framework that deals with the detection pf variety of DDoS attacks [15]. Cloud specific Intrusion Detection System was proposed and described a defense mechanism against the DDoS attacks. This defense mechanism discusses how to detect the DDoS attack before it succeeds [16]. Effectively detecting the bandwidth limit of a cloud network and the bandwidth currently in use helps to know when a DDoS attacks begin [17]. An approach described based on fundamentals of information theory specifically Kolmogorov complexity to detecting distributed denial of service (DDoS) attacks was proposed. Despite its complexity the scheme enabled early detection [18].

3 Strategy of DDoS Attacks

Even though there are many different types of DDoS attacks, all these attacks have the same attacking strategies and involve four stages of scenarios: ***Selecting agents*** where agents that will perform the attack is chosen, ***Compromising agents*** where the vulnerabilities of the zombie machines will be exploited and the attack codes are transferred, ***Communication*** where attackers will communicate with the handlers to determine which zombie machines are up and running and decide the other parameters, and finally ***Attac*****k** where the hackers initiate the attacks and other attacking parameters could be adjusted.

Fig. 1 Description of DDoS Strategy

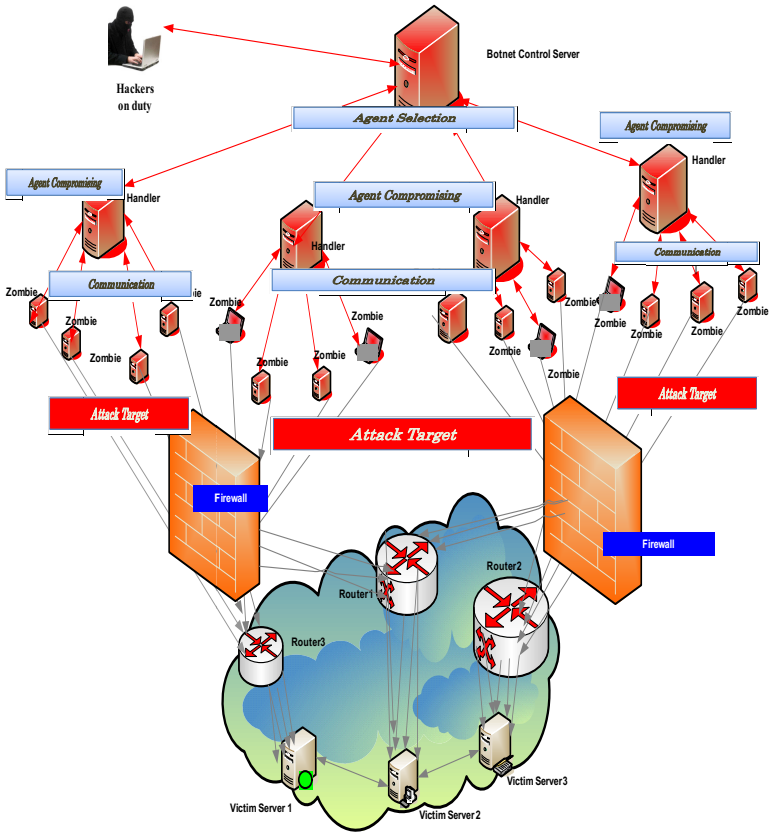


Fig. 2 Detain diagram Sources of an organized flooded DDoS attacking strategies [19]

4 Design Architecture of the New Detection System

The design Architecture includes all entities and structures involved in the overall DDoS attack and detection processes. It clearly shows where the attack packet get launched from, filtered and analyzed, and finally categorized as an attack or not attack.

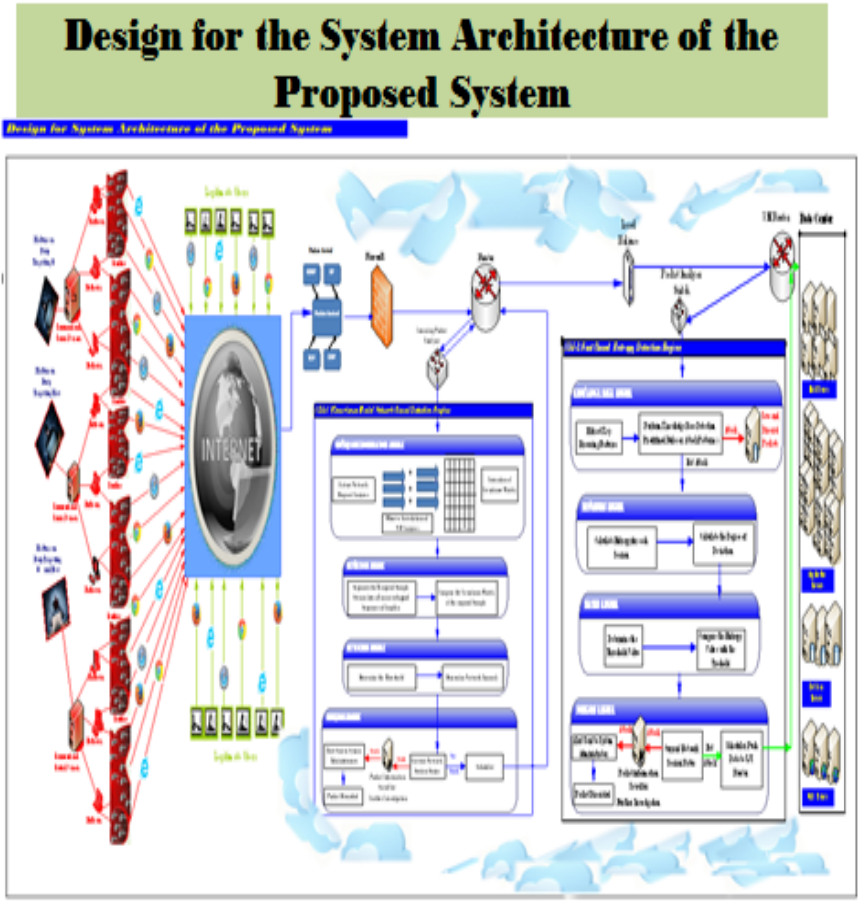


Fig. 3 Design Architecture of the Detection system

5 Data Flow Architecture of the Proposed System

The data flow architecture is designed on the following data flow diagram (figure 4) in more simplified way. The major functional procedures are mentioned and included and give a clear indication for any reviewers regarding how the packets get traversed throughout the system.

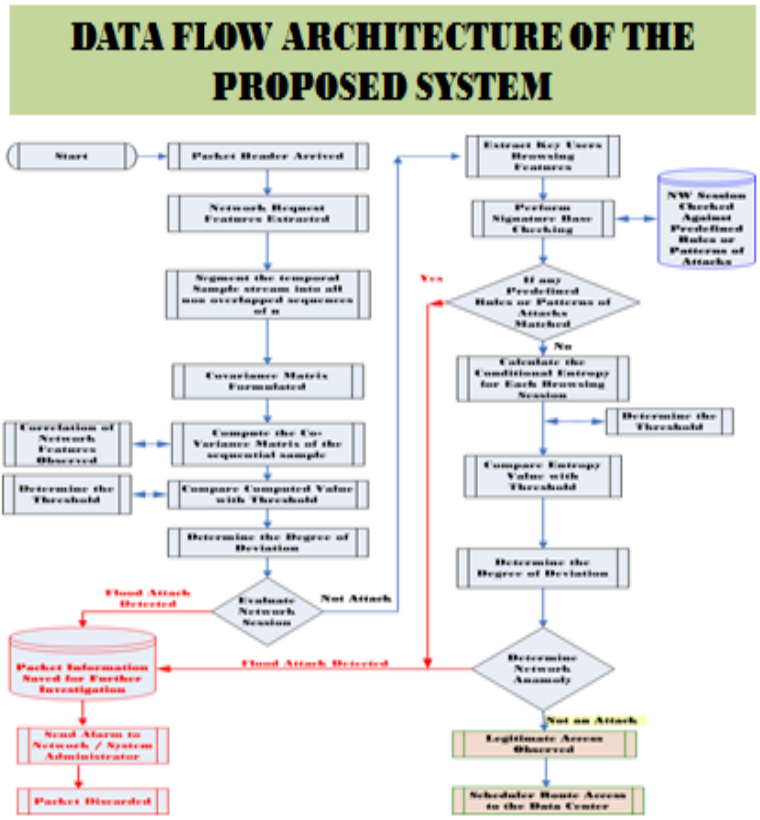


Fig. 4 Data Flow Architecture

6 Algorithm

Our Hybrid Algorithm is consists of two stochastic models. The first one is applied on the selected sample network features and passes the results (attack packet) to the second one that will analyze the same packet but with different selected browsing features that handles mainly the false alarm rate which is one of our main targets. The packet will finally be granted access to the cloud data center if it found not to be an attack after the second packet analysis procedure.

IP Packet Observed

P^* Sample Network features selected
Q^* Sample Browsing features

For Each session of P^* Network features selected

Compute the Kendall tau value using Concordance and Dis-concordance (Using formula 3)

Covariance Matrix C(X) will be modulated (Using formula 4)

The Deviation Function D(.) is computed

Let V = D(.)
Assign V a value equals to 0 or 1

If (V = 0) (Packet not an attack) then

(Perform the next packet analysis procedure)
Browsing features will be considered

For range of interval time observation (t = 1..T)

Compute entropy $H(\mathbf{X})=\sum_{\mathbf{X}} f(\mathbf{X})\ log(f(\mathbf{X})$
Compare the entropy rate $H(\mathbf{X})$ with the threshold

If $H(\mathbf{X})$ is equal or less than the threshold value

Packet discarded
Attack Alarm will be sent to Admin.

End If
End

Else (If V = 1) (Packet an attack)

(False Alarm Analysis is done using conditional entropy)
Compute Conditional Entropy (Using formula 12)

$H(\overrightarrow{\mathbf{x}}|\mathbf{X})=H(\overrightarrow{\mathbf{x}},\mathbf{X})-H(\mathbf{X})$

If $H(\overrightarrow{\mathbf{x}}|\mathbf{X})$ is equal or less than the threshold

Packet discarded
Attack Alarm will be sent to Admin

Else
Grant Access to the Cloud Data Center

End If
End If

End

7 Detection Approaches

Let $\mathbf{X}$ be a $p \times T$ multivariate vector where p is number of network 'features' or variables and T is the number of (discrete time interval) observations. Our objective is to identify the presence of atypical dependence in a sample after establishing a baseline dependence.

In statistical terms we evaluate the dependency among X for an ordinary, non-attack, regime – Let's call it $T_0 = T(\mathrm{X}_0)$, where T is some statistic of the multivariate data and X_0 is the baseline or non-attack, training, data. Then the task is to evaluate the 'distance', via this statistic, between the training data and 'new' data; large values of this distance indicate an 'attack'. In notation

$$D(T_0, T(\mathbf{X})) \tag{1}$$

7.1 Multivariate Correlation

Out of $p^* \leq p$ features the covariance matrix is only taking two features at a time, then testing after imposing a threshold and/or using 0, 1 as the distance. For ordinary $p \times p$ correlation/covariance matrix has entries proportional to:

$$Cov(X_i, X_j) = \boldsymbol{E}[X_i \cdot X_j] \tag{2}$$

there are $\frac{p\,(p+1)}{2}$ unique entries in such a matrix where i and j are now multivariate indices, for each of t: $len\,(i) = \mathrm{len(j)}$. The expectation above, then, should be a scalar and the matrix collecting these will be of dimension $p \times p$.

7.1.1 Kendall's Tau for Bivariate Correlation, Multiple Correlation

The population version of Kendall's Tau is:

$$\tau = P(Concordance) - P(Discordance) \tag{3}$$

These can be collected pairwise into a matrix as well, say, by calculating:

$$\tau_{ij} = P(\{Concordance\} - P(\{Discordance\})[X_i, X_j] \tag{4}$$

or into a matrix of reduced dimension

$$\tau_{ij} = P(\{Concordance\} - P(\{Discordance\})[X_i, X_j] \tag{5}$$

in analogy with the covariance and the same thresholding, can be done. Even to the use of Chebyshev's inequality to impose probability limits on the statistical values.

The sample version of Kendall's Tau is calculated by letting

$$P(\{Concordance\}) = \#\{X_i, X_j > 0 \ \& \ X_i, X_j < 0\} \tag{6}$$

$$P(\{Disconcordance\}) = \#\{X_i > 0, X_j < 0 \ \& \ X_i < 0, X_j > 0\} \tag{7}$$

Where # is the number of instances over all indices i, j.

7.1.2 Entropic Approach

The entropy of a multivariate, discrete, random variable is:

$$H(\mathbf{X}) = \sum_{\mathbf{X}} f(\mathbf{X}) \, log \, (f(\mathbf{X})) \tag{8}$$

This is the entropy across all features of the data, in terms of the model across the entire dimension of the multivariate vector H (.) is a function from $\mathrm{R}^p \rightarrow \mathrm{R}$. In the world of the data which for the moment we assume arrive from a stable process but possible 'attack' regime - we have $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T$ as the multivariate observations. Call these $\overrightarrow{\mathbf{x}}$. So our estimate of the entropy will rely on $f(\cdot)$ - which can be a particular model or an empirical estimator - and be estimated across $t = 1, ..., T$ units of time indexed observations is given by:

$$H(\overrightarrow{\mathbf{x}}) = \sum_{t=1}^{\mathrm{T}} f(\mathbf{x}_t) \, log \, (f(\mathbf{x}_t)) \tag{9}$$

Consider setting

$$H(\mathbf{X}) = T_0(\mathbf{X}) \tag{10}$$

as the *baseline* entropy among the features. Then we can consider

$$D(T_0, T(\overrightarrow{\mathbf{x}})) = H(\mathbf{X}) - H(\overrightarrow{\mathbf{x}}) \tag{11}$$

7.1.3 Conditional Entropy

Alternately we compute:

$$H(\overrightarrow{\mathbf{x}} | \mathbf{X}) = H(\overrightarrow{\mathbf{x}}, \mathbf{X}) - H(\mathbf{X}) \tag{12}$$

as measures of (an increase in) dependency among suspected attack data. In (12), a large distance between the baseline entropy and the entropy of the data (i.e. calculated across some time steps) is the signal for an attack. In (13), we use the entropy itself (i.e. the dependency between the features under the model and those of the data) as the measure of an attack. In (13) this measure is function of the model $f(\mathbf{X})$ on the 'training'

data; perhaps an empirical estimator $\hat{f}(\overrightarrow{\mathbf{x}})$ or the probability of the data given the model $f(\overrightarrow{\mathbf{x}})$. If we eschew an empirical estimator and calculate $f(\overrightarrow{\mathbf{x}})$ in (13) we can think of this as similar to a *likelihood based approach*, but where we access the likelihood via the entropy function. In a sense, this method is more 'complete': the mass from the entire probability distribution (via the model f and estimator $\hat{f}$) is used and not just the expectation.

8 Conclusion

On this paper we have presented and discussed the design architecture, data flow architecture, and algorithm of our newly hybrid approached solution model that could be a better choice in terms of detecting the DDoS flood attacks. The preliminary results showed that our detection technique is very encouraging and promising. We shall present the details of our research results, data analysis, and our recommendation for future works ideas on the next upcoming conference.

References

1. Hu, J., Yu, X., Qiu, D., Chen, H.-H.: A simple and efficient hidden markov model scheme for host-based anomaly intrusion detection. IEEE Network (February 2009)
2. Syed Navaz, A.S., Sangeetha, V., Prabhadevi, C.: Entropy based anomaly detection system to prevent DDoS attacks in cloud. International Journal of Computer Applications (0975-8887) (January 2013)
3. Yeung, D.S., Wang, X.: Covariance-matrix modeling and detecting various flooding attacks. IEEE Transactions on Systems, MAN, Cybernetics- Part A: Systems and Humans **37**(2) (March 2007)
4. Ismail, M.N., Aborujilah, A., Musa, S., Shahzad, A.: Detecting Flooding based DoS attack in cloud computing environment using covariance matrix approach. In: ICUIMC (IMCOM) (2013)
5. Yu, S., Zhou, W.: Entropy-Based Collaborative detection of DDoS attacks on community networks. In: Sixth annual IEEE International Conference on Pervasive Computing and Communications (2008)
6. Sha, J.J., Malik, L.G.: Impact of DDoS attacks on cloud environment. International Journal of Research in Computer and Communication Journal **2**(7) (July 2013)
7. Jin, S., Yeung, D.S.: A covariance analysis model for DDoS attack detection. In: IEEE Communications Society (2004)
8. Sendi, A.S., Dagenais, M., Jabbarifar, M.: Real time Intrusion prediction based on optimized alerts with hidden markov model. Journal of Networks **7**(2) (February 2012)
9. Ankali, S.B., Ashoka, D.V.: Detection architecture of application layer DDoS attack for internet. Advanced Networking and Applications **3**(1), 984–990 (2011)

10. Er. Kakkar, S., Er. Kumar, D.: A survey on distributed denial of services (DDoS). International Journal of Computer Science and Information Technologies **5**(3) (2014)
11. Patcha, A., Park, J.-M.: An overview of anomaly detection techniques: Existing solutions and latest technological trends. Science Direct, Computer Networks **51** (2007)
12. Bhuyan, M.H., Kashyap, H.J., Bhattacharyya, D.K., Kalita, J.K.: Detecting distributed denial of service attacks: methods, tools and future directions (December 2012). http://www.garykessler.net/library/ddos.html
13. Renuka Devi, S., Yogesh, P.: Detection of application layer DDoS attacks using information theory based metrics. CS & IT-CSCP, pp. 217–223 (2012)
14. Lonea, A.M., Popescu, D.E., Tianfield, H.: Detecting DDoS attacks in cloud computing environment. International Journal of Computing and Communication **8**(1), 70–78 (February 2013). ISSN 1841-9836
15. Gupta, B.B., Misra, M., Joshi, R.C.: An ISP level solution to combat DDoS attacks using combined statistical based approach. Journal of Assurance and Security **2**, 102–110 (2008)
16. Goyal, U., Bhatti, G., Mehmt, S.: A dual mechanism for defeating DDoS attacks in cloud computing model **2**(3) (March 2013)
17. Panda, B., Bhargava, B., Pati, S., Paul, D., Lilien, L.T., Meharia, P.: Monitoring and managing cloud computing security using denial of service bandwidth allowance (2012)
18. Kulkarni, A.B., Bush, S.F., Evans, S.C.: Detecting distributed denial-of-service attacks using kolmogorov complexity metrics. In: GE Research & Development Center (February 2002)
19. Girma, A., Abayomi, K., Garuba, M., Li, J., Liu, C.: Analysis of DDoS attacks and an introduction of a hybrid statistical model to detect DDoS attacks on cloud computing environment. In: ITNG-2014 (April 2014)

Educational Gaming: Improved Integration Using Standard Gaming Genres

Ben Brown and Sergiu Dascalu

Abstract A common problem seen in educational games is a lack of tight integration of educational content within the 'fun' portion of the game. We propose and test an educational game using the Shoot'em Up genre as a template for gameplay in which the educational content within the game is constantly presented to the user with the goal of ensuring that this content feels integral to the game. The goal is to enhance the appeal and fun factor of the game in the hope of increasing a learner's motivation to play the game, thus increasing their exposure to the contained educational content.

Keywords Educational gaming · Shoot'em up · Learning · e-Learning · Gaming

1 Introduction

Educational gaming is a topic of much discussion. As a research topic there are a huge number of questions being explored. These range from exploration of frameworks that can be used to improve educational games [1], new methods that can be used to improve aspects of educational games like adding adaptive learning [2,3], to general questions about efficacy of educational games [4,5] , how to better test for said efficacy [6], or what really defines an educational game[7,8]. Given the number of questions and inherent diversity of the educational gaming field [9,10], one may wonder where to focus if we intend to make meaningful improvements on what is considered the state of the art in educational gaming.

To tackle this we will quickly establish our own definition of what we consider to be a worthwhile educational game, discuss why educational games are important, look at the specific problem we will be focusing on, and then of course look at what we propose as a method of improving the state of educational gaming.

B. Brown(✉) · S. Dascalu
University of Nevada, Reno, USA
e-mail: abenyoucantrust@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_35

1.1 What is Educational Gaming?

The topic of what makes an educational game may initially seem trivial, but there appear to be 3 primary schools of thought on the subject whose definitions differ greatly. In one group we see a very broad view being taken, which would define educational games as being any games which result in a positive educational impact on their players. In papers like [11] we see games that were not created as educational games being tested as educational tools. On the other extreme we see a view cited in [10] in which an educational game need not only be designed to be educational, but must present its content in such a way that the user must be able to integrate new knowledge from the game without previous exposure to the subject, and excludes any game which is drill based. The third view is somewhere in the middle, and while not often cited explicitly, seems to allow for games that are educational according to [10], but which also allows for drill based gaming. This is the definition we will be following for the purposes of our research.

1.2 Why Educational Gaming?

Given some definition, the next thing we must establish is why educational games are important, at least for the context of this paper. Looking at the majority of the papers that exist on the topic, two broad reasons are cited as to being major attractive factors to educational games: Motivation, and Automated Evaluation.

Motivation is the most obvious reason that educational games exist, and its explanation can be summed up as follows. Education is not always exciting, thus by putting educational content into a game context we can enhance the motivational factor to partaking in that content. Thus by making education fun, a potential learner is more likely to partake in education.

Automated Evaluation is also simple. The idea is that within an educational computer game we can keep track of how well a student comprehends a subject in real time, reporting this to both the student and educators. Thus via the use of an educational game we can better track educational outcomes. There is also an added benefit. As seen in papers like [2,3] this automated collection of data may be used to tailor educational content to a student. Thus a student showing excellent understanding of a topic can have the difficulty of that topic raised, or may be presented with more complex subjects.

A third reason to look at educational gaming that is not as widely espoused in other research is an ability to engage a learner without requiring the presence of an educator to facilitate that engagement. While much of the research in educational games focuses on their application in the classroom, we see an advantage of the software as being able to encourage educational pursuits outside the classroom space, and thus consider the ability to promote self-driven educational pursuit an important characteristic.

1.3 A Problem

As mentioned above the state of the research into educational gaming is such that there are a huge number of potential research subjects to discuss, thus we must focus in on one subject of interest. For this we will be focusing on a topic discussed in S. Fisch's *Making Educational Games "Educational"* [12]. In this paper one of the author's major points is on a perceived lack of tight integration of educational content into. He suggests among other things that many educational games appear to make use of a method of teaching similar to a something known as the "Seductive Detail". The seductive detail refers to a topic studied within the realm of textbook efficacy and refers to the act of an author taking a boring topic and attempting to make it more appealing by inserting tangentially related entertaining content. An example of such a detail would be a text book looking at the formation of lightning which also discusses a man who survived being hit by lighting on 7 different occasions. The idea is that this entertaining story will hold the attention of a student in an otherwise dry text, but in reality there is a large amount of data to suggest that this doesn't work. What appears to happen is that a student exposed to this method tends to remember the entertaining content over the educational content. Thus not only does the use of the seductive detail fail to enhance educational content, but Fisch suggest it may have the more long term negative effect of making educational content less appealing to a student as it may make it appear as little more than a hindrance to them enjoying the fun content.

So, how does this apply to educational gaming? Well according to Fisch a number of early educational games were little more than simple pen and paper drills put onto a computer, which when completed would reward the user with an animation. These are posited as approximating the seductive detail as the educational content is still being presented as it would outside a game, but the 'reward' for completing the drill was made more entertaining than a grade, even if this reward had no inherent association with the educational content. Thus we get the seductive detail paradigm. The problem is that when we look at modern educational video games a number of them make use of this same basic stratagem, although the entertaining portion of the game has evolved slightly from simply showing off some animation.

1.4 A Modern Educational Game

While this is far from universal there is a specific design that we can see in a number of modern educational games. In these games the content is presented in the following way. First a learner is introduced to some entertaining content. This might be something akin to the animations mentioned by Fisch, showing a popular cartoon character introducing a subject, or in many cases we see gameplay akin to a standard video game. A student will interact with this entertaining content for some amount of time, and then the game will switch to an educational phase. At this point the learner is asked to answer some number of questions based on the

educational topic. After these questions are answered, the game returns to the 'fun' content. This results in a bifurcated game, one in which the educational content and the entertaining content are two distinct portions of the game, and this is a method which very much approximates that of the 'Seductive Detail' discussed before.

1.5 The Solution?

According to Fisch the solution to this dilemma is tightly integrating educational content into the entertainment content of a game. Fisch sites SimCity as example of this more idealized educational game. In that game any educational content is directly related to the core gameplay. This solution seems obvious, thus the question is why this hasn't become a more common standard. The main reason is complexity. Looking at SimCity as an educational game we see the educational content as city planning. How do you set up a city with proper transportation, business opportunities, enough power, and minimal pollution, all within the confines of city budget? The answer here is to create a city planning simulation, hence SimCity. In this case the entertaining content is the educational content. While creating such a game is a non-trivial task, it is easy to see how a subject like city planning can be taught via simulation, and how a simulation can be turned into a game. What we will be looking at is how to tightly integrate educational content into a game that cannot be easily modeled as a simulation. There are a number of subjects like math, spelling, and music, which require some amount of repetitive drill based learning, so how do we integrate those subjects into a gaming context without relying on the seductive detail?

2 Our Proposal

2.1 Altering Standard Video Game Genres for Use in an Educational Context

Looking at the problem above, we have this general question: "How can we tightly integrate educational content into a video game." We have an example of a game in which this was well achieved in the form of SimCity, thus we have two potential new questions we can look at. These would be: "Given the successful integration of educational content into the simulation game genre, how can we integrate educational content not directly tied to simulation into a simulation game?" or "Given the ease in which some educational content can be turned into a simulation game, is there any other genre of game which will lend itself to similar tight integration of educational content". This second question will be the one we will be focusing on.

2.2 Trivial Cases

There are of course some genres that present trivial cases to this answer. Any form of quiz based game can of course be easily adapted to most subjects. These games fail to distinguish themselves from standard teaching methods though, and we don't see them as being a major source of enhanced educational motivation. Instead, what we will be discussing are games that don't have any obvious educational content present.

2.3 Necessities for Drill Based Education

We have implied above that our main goal is to look for ways to teach educational subjects that require drill based learning in a gaming context without relying on the bifurcated gaming paradigm seen in so many commercially available educational games. So our overarching goal is the creation of a game in which educational content is constantly presented to the user, and in which that content is tightly integrated into the core gameplay. For this we will be looking at genres that lend themselves to some key attributes:

- There should be one primary gameplay goal for ease of conversion, as having to alter multiple gameplay mechanics to make each educational is a complex task.
- The main gameplay goal should be repetitive, as our educational drills have to be mapped into this game play goal.
- The gameplay goal should be tied to actions that happen consistently through the game.
- We must be able to map our drills onto this goal without dramatically changing the base gameplay. Looking at existing genres is moot if we alter them to a point where they are unrecognizable.

2.4 The Shoot'em Up

Given the attributes stated above we have chosen the Shoot'em Up (SHMUP) genre as our test case. The SHMUP genre generally refers to games in which a user controls some avatar that navigates through a scrolling world destroying as many enemies as possible without being destroyed. This genre meets each of the goals we've stated. The primary goal of the game is to simply control a user avatar, navigating to line up shots and avoid enemies, and does not go beyond this. While gameplay includes power-ups, these don't tend to change core gameplay mechanics. The gameplay is highly repetitive and consistent, seeing the player fire projectiles seemingly without stop. The last goal though, the mapping of educational content is a separate discussion.

2.5 Music Education as the Base Case

Mapping educational content into the SHMUP in such a way as to maintain the integrity of the SHMUP's core gameplay is a non-trivial task. For this we have chosen to create an initial edition of this game using music as our chosen subject for the following reasons: First, musical skills require frequent repetition to master, and a number of different sub topics within the musical spectrum can be taught via drill based learning. This fact is key to any subject we might consider given our focus on drill based education. The second reason we tested music is that it can be considered a mostly grade level agnostic topic. Skills that can be learned at a primary education level are just as applicable to secondary education and beyond, thus music affords us an unusually large potential pool of users to test the software on. The third reason is that most of the topics we can attempt to teach have answer forms which are easily mapped onto single key inputs. For a SHMUP this is ideal as we can hopefully allow a user to fire quickly as opposed to needing to wait while they type in multi keyed answers. Given this subject, we can now discuss the actual game which has been created.

3 Music SHMUP

3.1 Design

Given our proposed genre and educational topic we can discuss how the game will be designed in order to create an educational SHMUP. The first element that must be tackled is integration of educational content into the SHMUP. Our chosen method consists of the following: A specific topic in musical education was chosen, key signature memorization. This was chosen for a number of reasons. The content is easy to display, as we simply need to show a key signature and a note as a question. The potential answers are easily mapped to 3 keys, representing a sharp, flat, and natural. This was then mapped into the game by displaying the question in the games heads up display (HUD), which refers to data presented in a game that is not necessarily part of the gameplay, but displays pertinent info like score, player health, etc. This is seen in Figure 1. One caveat to take note of in the images seen is that the key chosen is called "Crazy Major". This is a key created for testing purposes which contains both sharps and flats, which is impossible in actual key signatures. Actual gameplay is designed to work as follows. A player controls a user avatar via use of the WASD keys. They defeat enemies by firing projectiles, which is done by answering the questions displayed in the HUD. The questions change every couple seconds, which in this case means the note requested is changing, while the key signature itself remains constant for the duration of a level, although this can be changed to allow for both the notes and key signature to change. With this we have the basis of a game which makes use of standard SHMUP mechanics, but sees gameplay requiring constant educational input. For this version of the game some other alterations were made. First, the projectiles

being fired contain a symbol representing the players answer. This was done so that the educational content appears more tightly integrated, as this allows a user to fire a wrong answer which, when it contacts an enemy, does no damage. Thus we have a negative repercussion to a wrong answer which results in a meaningful effect to gameplay. In this version we are sacrificing some of the speed seen in most SHMUP's (as mentioned above many SHMUPS see a player firing almost constantly) in order to more tightly bind our educational content into the gameplay, thus we have made a second change to the base SHMUP. We have limited the number of enemies that appear in any given enemy wave, and have made it so those enemies must be killed in order to complete a level. This was done to ensure that in order to complete a level a player would have to correctly answer some minimum number of questions (although they can actually answer more as a correct question must result in a projectile being fired that actually hits an enemy in order to progress). After the standard enemies are all defeated a player must face a boss. This is simply an enemy that is more dangerous to the player, and requires multiple projectile hits to defeat. When the boss is present we also present a change to the HUD in which we show the bosses health via a life bar so that a user can track their progress in defeating a boss, as can be seen in Figure 2. This follows directly from standard SHMUP gameplay elements.

Fig. 1 HUD and Gameplay

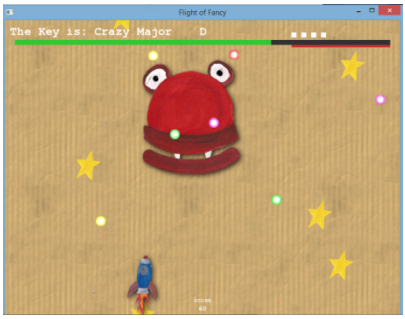


Fig. 2 Boss Enemy and Life Bar

3.2 Result and Expansion in Music Game

As we can see we have accomplished the base task of creating an educational game in which the educational content of the game is tightly integrated into the fun content. This version is not necessarily an idealized example, as there are a number of design changes we could make which might improve or worsen the game, but those changes will be discussed later in the paper. What we can say is that we have accomplished the goal of keeping the educational content of the game in constant focus while the standard gameplay elements of the SHMUP remain intact. At this point though the overall educational scope of the music game is somewhat limited, thus we will briefly mention future potential expansions of the music game into other sub topics, and how those would be achieved.

The first expansion envisioned would be on chord intervals. This topic would see a player given two notes and asked what the interval between them is (so a B flat and B would be an interval of 1, a B flat and C would be 2, etcetera). This would require that the game be programmed with knowledge of each interval, but the HUD would not require major changes, and the method of answer would remain mostly constant (using a number pad each answer can be easily shown), with the main caveat being the ability to look at a 10 interval, which would likely require a single odd mapping of 10 to a key like +.

Future improvements would likely require that a module be made that could show actual music staffs with notes generated on it based on a clef. With this we could do an enhanced interval solver, and a note recognition system, although mapping every potential note to a key intuitively has its own complexities. With note and chord generation via sound, ear training lessons could also be implemented.

4 Future and Ongoing Work

4.1 Changes to Educational SHMUP

Looking at the current game, we can discuss potential methods of improving the core gameplay to enhance fun factor and educational integration, other subjects that this game design could be applied to, the need for new HCI methods, and potential for efficacy testing.

We will begin by looking at improvements to the game design we currently have. Here there are a number of potential areas for improvement. The first has to do with the placement of the educational content. We have reason to question whether the current placement within the HUD is optimal, thus the most obvious potential change would be alteration of the HUD in such a way as to maximize visibility of the content, while attempting to make the content appear better integrated into the game. This is non-trivial as in general HUD's are best suited for displaying information whose size is relatively static. We also have reason to question whether having notation appear on the bullets is useful. While it seemingly adds towards the integration of educational content into the game, it limits the ability to implement some power ups, and its actual effect on promoting education may be minimal.

A larger potential change that has been contemplated is a revision that sees the use of a power meter and auto fire mechanic. This would see the ship constantly firing projectiles (as mentioned, in many SHMUPS releasing the fire key serves no purpose so this would not be too out of line with the games archetype) as long as a players power meter was partially filled. This meter would fill every time they answered a question correctly, but would lower any time an answer was missed (or it would simply not fill up and continue to drain at its standard rate). This would allow for a much larger number of enemies to appear as projectiles could be fired in a much larger number, would allow for a more standard time based enemy wave system, with enemies that need not be defeated to win a level, which overall would help make the game look and feel more like a traditional SHMUP, but could break some of the immersion of the educational content.

4.2 Application to Other Subjects and New HCI

Given that our primary requirement for this game is that the subject needs to be one that is learned from quick drills, we are somewhat limited in other potential subjects. Math is the most obvious subject, and is one that is currently being experimented with. A second version of the game not shown here has been created using math topics instead of music as its educational topic. The game required some changes to the core gameplay, such as needing to add a space bar enabled slow motion mechanic to allow for the use of multi digit answers (while held down a user can type multiple numbers and when released these would be fired as a single answer) as well as changes to the HUD. Monsters in the game also contained part of the arithmetic problems being solved as this was seen as a good way of better integrating the educational content into the game, by making it so that each enemy required a unique answer to be defeated. This seemed potentially useful, but testing still needs to be done, and separating the content from the HUD may end up making the game somewhat confusing. The games are promising, but are not the main focal point of current research. The main issue is that these versions are limited to arithmetic due to the need of having questions that can be answered quickly, and this limits the potential appropriate age ranges of users to an extent where it is harder for our team to attempt any real efficacy testing.

The second subject that has been discussed is spelling. Spelling would be potentially viable, but leads to the question of how you ask a user to spell a word without typing that word on the screen, and how do allow input of a word when part of our keyboard is devoted to movement controls, and knowing that proper typing requires a two handed position. This leads directly to a topic of future research that we consider key to future creation of good educational games: Novel HCI methods. For example, one such method would be the use of text to speech technology, and voice recognition.

One of the major questions we have for each game is how to present questions clearly, and how to map answers to those questions to the limited number of inputs we have available, thus the use of voice control becomes quite tempting. By using text to speech technology questions can be asked without use of the HUD (although this leads to its own questions as to how a user can repeat a question, and ensuring they aren't bombarded by computer speech), and by allowing for vocal input the whole keyboard can be used for ship control, or we could switch control to other traditional input devices like a game controller or touchscreen. This of course may simply be infeasible given the current state of voice recognition, but does lead to one very important conclusion. Looking at these games it is our opinion that one of the biggest questions facing drill based educational gaming (and perhaps educational gaming in general) is natural ways to answer questions while still maintaining control of the game. Without coming up with novel ways to interact with the computer, the need to have one of a player's hands dedicated to controlling in game action for most genres, and this limits our ability to let them answer many questions in an intuitive fashion, thus this is a topic of future research that we believe needs to see a greater research emphasis.

4.3 Efficacy

At this point it should be apparent that we are lacking in efficacy testing. While we have demoed the current games to small groups of individuals, these tests were not done in such a way as to prove the efficacy of our games in a statistically relevant way. Thus one of the largest pieces of future research that needs to be done is a controlled experiment comparing these games with other educational games on the same subject, and looking for differences in learning rate, and motivational factors.

4.4 Conclusions

At this point our main conclusion is that it is indeed possible to create an educational game in which the educational content is tightly bound to the entertaining content within that game. Specifically we have shown that with some modifications the SHMUP genre of games makes a reasonable template for games that require rapid drills to facilitate learning. We have also posited that one of the greatest challenges to use of this genre in an educational context (a challenge which we predict will apply to games in any genre) is capturing a user's answers to questions posed within the game in an intuitive fashion. As most educational subjects see would be learners answering questions that are not easily mapped to a limited number of keys, novel ways of capturing this input may be needed to expand this research beyond the subjects covered here.

A note that must be made is that even within the topics of math and music, the modification of the SHMUP genre cannot be viewed as a universally ideal solution. While we have evidence that educational gaming is especially useful for students who are failing to comprehend a subject taught by other means, [5] one topic seemingly ignored in much of the research done on the use of educational games is user preference. A learner who finds SHMUPs highly engaging may find this particular game to be greatly appealing. A user who has no interest in SHMUPs though may find that this specific game elicits no motivational benefit over a pen and paper drill, or worse may see a negative motivational outcome, finding the game to be more frustrating than traditional educational methods. Thus one important take away that we believe needs to be stressed is that while educational gaming may provide an extremely useful educational tool, no one game will ever be an educational 'silver bullet'.

Acknowledgment This material is based upon work supported in part by the National Science Foundation under grant number IIA-1329469.

References

1. Ahmad, M., Rahim, L.A., Arshad, N.I.: A review of educational games design frameworks: an analysis from software engineering. In: International Conference on Computer and Information Sciences (ICCOINS), pp. 1–6 (2014)
2. Peirce, N., Conlan, O., Wade, V.: Adaptive educational games: providing non-invasive personalised learning experiences. In: Second IEEE International Conference on Digital Games and Intelligent Toys Based Education, pp. 28–35 (2008)
3. Thomas, J.M., Young, R.M.: Annie: automated generation of adaptive learner guidance for fun serious games. IEEE Trans. Learn. Technol. **3**(4), 329–343 (2010)
4. Backlund, P., Hendrix, M.: Educational games - Are they worth the effort? a literature survey of the effectiveness of serious games. In: 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES), pp. 1–8 (2013)
5. Virvou, M., Katsionis, G., Manos, K.: Combining software games with education: evaluation of its educational effectiveness. Educ. Technol. Soc. **8**(2), 54–65 (2005)
6. Wideman, H.H., Owston, R.D., Brown, C., Kushniruk, A., Ho, F., Pitts, K.C.: Unpacking the potential of educational gaming: a new tool for gaming research. Simul. Gaming **38**(1), 10–30 (2007)
7. Frazer, A.: Towards better gameplay in educational computer games: a Ph.D. thesis. PHD Thesis, University of Southampton (2010)
8. Kafai, Y.B.: Playing and making games for learning instructionist and constructionist perspectives for game studies. Games Cult. **1**(1), 36–40 (2006)
9. Hwang, G.-J., Wu, P.-H.: Advancements and trends in digital game-based learning research: a review of publications in selected journals from 2001 to 2010: Colloquium. Br. J. Educ. Technol. **43**(1), E6–E10 (2012)
10. Dondlinger, M.J.: Educational video game design: a review of the literature. J. Appl. Educ. Technol. **4**(1), 21–31 (2007)
11. Kozuki, K., Imachi, M., Ueno, M., Tsubokura, A., Tsushima, K.: Computer game and educational system. In: International Conference on Computers in Education, Proceedings, vol. 2, pp. 1377–1381 (2002)
12. Fisch, S.M.: Making educational computer games educational. In: Proceedings of the 2005 Conference on Interaction Design and Children, New York, NY, USA, pp. 56–61 (2005)

Designing a Web-Based Graphical Interface for Virtual Machine Management

Harinivesh Donepudi, Bindu Bhavineni and Michael Galloway

Abstract In any cloud computing architecture, virtualization is very important. We can accomplish this with the creating of virtual machines, guest operating systems that run on top of a host operating system by means of a hypervisor. There are many hypervisors (Xen, KVM, Virtual Box), each having advantages and disadvantages. The goal of this project is to use the KVM hypervisor, which comes standard with many Linux OS distributions, to host virtual machines that are deployed from a web-based graphical user interface.

1 Introduction

KVM (Kernel-based Virtual Machine) is a virtualization infrastructure for the Linux kernel. It represents the latest generation of open source virtualization. KVM is implemented as a loadable kernel module that converts the Linux kernel in to a bare metal hypervisor. The KVM developers, instead of creating a major portions of an operating system kernel themselves, as other hypervisors have done, devised a method that turns the Linux kernel itself to a hypervisor. It requires a processor with hardware virtualization extensions (Intel VT or AMD-V). It consists of a loadable kernel module, *kvm.ko* that provides the core virtualization infrastructure and a processor specific module, *kvm-intel.ko* or *kvm-amd.ko*. KVM also requires a modified QEMU although work is underway to get the required changes upstream.

Using KVM, we can host multiple virtual machines running unmodified Linux or Windows images. Each virtual machine has its own private virtualized hardware which runs on the host machine. KVM can be managed either through a graphical management tool, similar to VMware products or virtual box or through a command

H. Donepudi · B. Bhavineni · M. Galloway(✉)
Department of Computer Science, Western Kentucky University,
Bowling Green, KY 42101, USA
e-mail: jeffrey.galloway@wku.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_36

line using several methods. The most popular GUI is called virtual machine manager (VMM), which was developed by Redhat. The tool is also known by its generic package name virt-manager. It also supports XEN machines. KVM is primarily a command line tool not as intuitive or friendly as VMware and virtual box products but we can have better productivity and more control in the long run [2, 11].

2 Installation

2.1 Network Configuration

Once the pre-checking is complete we have to connect the virtual machine to an external network. There are few different ways to access the virtual network. For our project we will set an internal bridge to the host. The bridge will work when the physical network device used is a wired device. The VM guests can have their own virtual network interfaces and IP addresses. To create a network bridge on the host we have to install the bridge utility packages by the following command:

```
$ sudo apt-get install bridge-utils
```

To set up the bridge interface we have to edit the */etc/network/interface* file. The file is reconfigured so that the bridge is assigned an IP address using DHCP. Apart from installing the bridge utility packages we can configure the network with the *iproute2 ip* command which is default. Then restart the network:

```
$ sudo /etc/init.d/networking restart
```

If you are remotely connected using SSH, your connection may drop once you restart the network. This happens if /textiteth0 is configured as a static IP address rather than using DHCP IP address. Just SSH back into the machine using the new IP address.

2.2 KVM Configuration

The network is configured and we can install KVM and configure it. KVM can be installed with the following command:

```
$ sudo apt-get install qemu-kvm libvirt-bin Ubuntu-vm-builder
```

We can check whether the KVM is installed or not with the following command:

```
$ virsh −c qemu:/// system list
```

When KVM is installed it creates a default network with a configuration file. If the default configuration is suitable for some situation we can keep it as it is, but if we need to create a new configuration then we must execute the following commands in *virsh* shell:

```
net−destroy default
net−undefine default
```

After the KVM is installed the Ubuntu VM builder is also installed. VM builder is the best tool for creating the virtual machines which run on Ubuntu host. It is a script that automates the process of creating a virtual machines. This Ubuntu VM builder was first introduced as a shell script in Ubuntu 8.04. Later it was used by developers to test their code on virtual machines. This provides an easy way to maintain our web interface as an alternative to command line interactions. When KVM is installed, a network is created with a default configuration. We have configured the bridge network previously which uses DHCP to allow hosts to obtain IP addresses dynamically. To modify this network, first start a *virsh* shell and execute the following commands:

```
net−destroy default
net−undefine default
net−list all
```

The *list all* command will show that the default network has been removed. So, to configure the bridge network, exit *virsh* and enter the command:

```
$ nano configuration file
```

and make the necessary changes to the file and to configure the network. In the configuration file enter the following:

```
<network>
    <name>br0</name>
    <forward mode='bridge'/>
    <bridge name='br0'/>
</network>
```

Save and exit the file and return to the *virsh* shell in order the define the network:

```
net−define br0.xml
net−autostart br0
net−start br0
```

Now reboot the host to ensure everything is up and running. To check this run the following command:

```
$ virsh net-list -all
```

2.3 Installing a Virtual Machine(VM)

Once KVM installed and configured, a machine reboot is required before creating virtual machines. Without a system reboot, an error may be reported to the virtual machine logs located in the */var/log/libvirt/qemu* directory. Next, we have to create an image for the virtual machines. For that we have to create a new directory as we cannot create it in root directory. We can create a new directory for each VM we are about to create and each VM will have a subdirectory called *ubuntu-kvm*. If we try to install a second VM, it will alert us with an error that *Ubuntu-kvm* already exists unless we run *vmbuilder* with the *dest=DESTDIR* argument. Now we will use the *vmbuilder* tool to create the VM, which uses a template to create virtual machines in the */etc/vmbuilder/libvirt/* directory. We create a copy of that directory in the user defined destination address. After that we have to partitioning for the VM. We create a file called *vmbuilder.partition* and define the desired partitions. If we have to get a unique openssh key for each VM we have to install *openssh-server*. We cannot install this while we create VM. We have to create a script which will be executed when the VM is booted for the first time. The following command is used to install the openssh server:

```
$ sudo apt-get install -qqy -force-yes openssh-server
```

It will install the *openssh-server* and force the user to change the password when they log in for the first time. To create the virtual machine, change the directory to the VM directory and run *vmbuilder* [12].

```
$ sudo vmbuilder kvm ubuntu
    --suite=precise flavour=virtual arch=amd64
    --mirror=http://us.archive.ubuntu.com/ubuntu -o
    --libvirt=qemu:///system --ip=192.168.1.122
    --gw=192.168.1.1 --part=vmbuilder.partition
    --templates=mytemplates --user=kvmuser --name=kvmuser
    --pass=ubuntu --addpkg=vim-nox --addpkg=openssh-server
    --addpkg=unattended-upgrades
    --addpkg=acpid-firstboot=/var/lib/libvirt/images/vm1archer/boot.sh
    --mem=256 --hostname=vm1archer --bridge=br0
```

The option *-part* specifies the file about the partitioning details which is related to the present working directory. *-template* holds the template file. *-firstboot* specifies the file which has to be executed first when the VM is rebooted. *-mirror* specifies the default Ubuntu repository. If the IP address is specified by the *-ip* line, we have to

specify the gateway IP using the *-gw* line in the file. The gateway IP address is same as that we have used in our */etc/network/interfaces* file. After the VM is built we will find an XML configuration file for the virtual machine in the */etc/libvirt/qemu/* destination. Now, the user has the ability to connect to the virtual machine. We use *virsh* for connecting to the VM.

```
$ virsh −connect qemu:///system
$ virsh list −all
$ virsh define /etc/libvirt/qemu/VMname.xml
```

Once we run the above the command we will enter into the virtual shell prompt. To see the all VMs that are installed in the host use the below command:

Starting a VM for the first time, we have to define it in the XML file located in the qemu directory:

Now start the VM and connect using the following commands:

```
$ virsh start VMname
$ ssh username@ipaddress
```

2.4 Installing a Guest

So far we have installed KVM and VM builder. Next, the user should install a guest OS on the virtual machine. For that we need to execute the following command:

```
$ sudo ubuntu−vm−builder kvm trusty
```

The VM builder will install the trusty, a flavor of Ubuntu, on your virtual machine as a guest operating system. We can manage our installed virtual machines from the *virsh* shell. The virtual machines must be defined before management can take place. The *list* command in the *virsh* shell will list all the virtual machines which are currently on the host system. The *list* command with *–all* option shows all the active and inactive domains currently on the host system. When the virtual machine is created, it is in the shutoff state. Open the configuration file for the new guest and note the MAC address. Make an entry in the DHCP server of the Mac address and start the guest.

We can undefine the virtual machines with *undefine* command in the virsh shell. The virtual machine must be shut down in order to be undefined. The forcible way to shut down a virtual machine is to use *destroy* command. To start a virtual machine. Open the configuration file, set the MAC address, and then start the machine with the *start* command. Before starting a new VM for the first time, you must define it from its XML file located in */etc/libvirt/qemu/* directory. Whenever you modify the VM XML file in */etc/libvirt/qemu* you must run the *define* command and start the VM [6].

3 Experiment Design

In our design we are considering two cloud servers named *janeway* and *archer* in which the Ubuntu OS is installed. The two servers are connected to client machine using a NAT router and switch over gigabit Ethernet. A switch is a multiport network bridge that is used to forward data at the datalink layer and uses packet routing to transfer packets. The client is connected to the server through a NAT router. The user is able to deploy and terminate virtual machines in the web browser. The web interface also gives information about the running virtual machine instances and available virtual machine images.

To create the web interface, we have to install apache2 web server on both the cloud servers *janeway* and *archer*. Next, we installed PHP5 on the cloud servers for allowing the client to invoke server commands using the web interface. We used Ajax and HTML on the client side to access the PHP scripts which are executing on the servers. As PHP is installed on the server side to execute the commands we use shell_exec for shell scripting in server. The shell_exec command execute through the shell and returns the output. Thus function is identical to the back tick operator. This function is disabled when PHP is running in safe mode. Ajax is a technique for creating better, faster and more interactive web applications with the help of HTML and javascript. Conventional web applications transmit information to and from the server using synchronous requests. With Ajax when submit is pressed, JavaScript will make a request to the server, which then interprets the results and updates the current screen. Ajax offers a special technique for client-side javascript to make background server calls and retrieve additional data as needed. With Ajax we can get a better balance between client functionality and server functionality when executing the action requested by the user. Using Ajax and JavaScript we can access PHP script using these calls. JavaScript is a scripting language whose code is written in plain text and embedded into HTML. When a client request an HTML page, the HTML page can contain JavaScript. HTML is built of markup that web browsers understand, parse and display. It is a language that describes document formatting and content, which is basically composed of static text. Client side technologies enable the web client to do more interesting things than displaying static documents. Server side technologies are those that enable the server to store logic to build the web pages very easily.

3.1 Installation of DHCP

The Dynamic Host Configuration Protocol (DHCP) is network service that enables host computers to be automatically assigned settings from a server as opposed to manually configuring each host. Computers configured as DHCP clients have no control over the settings they receive from the DHCP server and the configuration is transparent to the user. The most common settings provided by a DHCP server are

IP address, subnet mask, DNS, and default gateway router IP address. At a terminal prompt enter the following command to install *dhcpd*,

```
$ sudo apt−get install isc−dhcp−server
```

We have to change the default configuration by editing the */etc/dhcp3/dhcpd.conf* to a desired configuration. Next, we edit the DHCP server file to specify the interfaces *dhcpd* should listen to and assign a static IP address to the interface that we use for DHCP. If you use *eth0* for providing address in the 192.168.1.x subnet, then you should assign for an instance IP 192.168.1.1 to the *eth0* interface using the network manager. Errors are possible during the installation, so to mitigate this, we assign the IP address randomly in the *dhcpd.conf* file. We have to specify the IP address range, for example, between 192.168.1.10-192.168.1.100. We can have any range but it as to be in between 192.168.1.x-192.168.1.n. It will lease an IP address for 600 seconds if no lease time is specified. The time frame can have a maximum value of 7200 seconds. Next restart the DHCP service to apply correct IP addresses to virtual machines. Occasionally we had some permission issues with the DCHP server, so we have to check the permissions and open the *apparmor* profile for *dhcpd*.

```
$ sudo apparmor_status
```

The above command shows the list of profiles, if it has */usr/sbin/dhcpd* profile then stop the *apparmor* daemon and edit the file with the root permissions and then restart the daemon. After installing the DHCP, we connect to the virtual machine using the vmbuilder with the DHCP inline command which is given below:

```
$ sudo vmbuilder kvm ubuntu–suite=precise −−flavour=virtual
    −−arch=amd64 –mirror=http://us.archive.ubuntu.com/ubuntu
    −o −−libvirt=qemu:///system −−part=vmbuilder.partition
    −−templates=mytemplates −−user=kvmuser −−name=kvmuser
    −−pass=ubuntu −−addpkg=vim−nox −−addpkg=openssh−server
    −−addpkg=unattended−upgrades −−addpkg=acpid
    −−firstboot=/var/lib/libvirt/images/vm1archer/boot.sh
    −−mem=256 −−hostname=vm1archer −−bridge=br0
```

3.2 Address Resolution Protocol (ARP)

ARP is used for mapping an internet protocol addresses to physical machine addresses. The ARP cache table is used to maintain the relation between mac address and its corresponding IP address. It provides the protocol rules for making this correlation and provides address conversions in both directions. A host machine is getting an incoming packet with the destination address on a local area network arrives at the gateway, it asks for the ARP program to find the physical host or MAC address for the incoming IP address. The ARP program looks in the cache table, and if it finds the address, it routes the packet to the required destination address. If there is

no entry in the cache table, ARP will send the packet to the all the clients and if any destination founds out that it belongs to that client it routes that packet to that destination. This protocol updates the ARP cache table for future reference and then sends the packet to the MAC address. We can find the ARP address of the available hardware with the following command.

The ARP [10] cache table is stored in the system routing table which dynamically create the host routes. ARP entries can be added, deleted or changed with the ARP utility. The entries added manually are permanent or temporary and published in which case the system will respond to the ARP requests to the hosts. A host which responds to an ARP mapping request for the local host address. If the local host responds to the requests for addresses other than itself with its own address is called proxy APR. The route to a directly attached Ethernet network is prescribed as cloning. Normal timeouts for these routes will be 20 minutes after validation. Entries are not validated when not in use [13].

3.3 Problems with DHCP

There are some cases where we only know the domain name of a remote host. For this purpose we use domain name service protocol to get the corresponding IP address. When the sender only knows the receivers IP address, the interface uses Address Resolution Protocol(ARP) for finding the receiver MAC address. For finding the receiver MAC address, the sender will transmit an ARP broadcast over Ethernet which is received by every other device on the Ethernet. The device with the matching IP address will reply back with an ARP reply along with the MAC address of the device. Otherwise, when the PC boots, it only has knowledge of its own MAC address. In this case, we reverse ARP and DHCP to get the IP address. RARP is similar to the ARP message format. When the booting system sends the ARP broadcast request, it sends its own hardware address in both sender and receiver in the encapsulated ARP packet. The RARP server will fill in the correct sending and receiving IP addresses in its response to the message. Similar to RARP, DHCP provides IP address, default gateway IP address, sometimes DNS server IP address to the DHCP client. However, there are two big advantages to using a DHCP IP pool, which is a range of IP addresses dynamically assigned to requesting PCs. The second advantage is there is no need for a static MAC-to-IP mapping using the DHCP server [13].

4 Results

The apache 2 web server is installed in the servers' *janeway* and *archer*. PHP is installed on the servers through Shell_exec we can execute the shell scripts. By using JavaScript, Ajax and HTML on the client we have successfully created the user interface for maintaining virtual machines on the two servers. Using the client we

are able to deploy the web browser used to manage virtual machines. Management activities include launching the virtual machine, stop the VM and listing all the virtual machines that can be or already deployed. We are able to launch independent GUI interfaces for each server through which we can maintain the virtual machines located on each server. The architecture of the entire cloud system with browsers accessing the cloud servers from the client machine is shown below. We have used the address resolution protocol to communicate with the router and the physical machine address. We have installed the DHCP service on the cloud servers with which we can automatically configure the IP addresses according the number of virtual machines instantiated. When we open the browser for managing the virtual machines, it will prompt for the login credentials to access the VM management dashboard. Refer to Fig. 5.

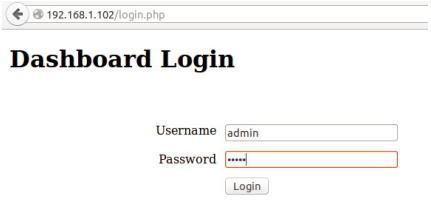


Fig. 1 Login Page

Fig. 2 Home Page

It will route us to the next page which shows all the managing functionality buttons for managing the virtual machines in the cloud servers. We can manage the virtual machines using the buttons in the interface. The *show All VMs* button results the available VM in the cloud servers.

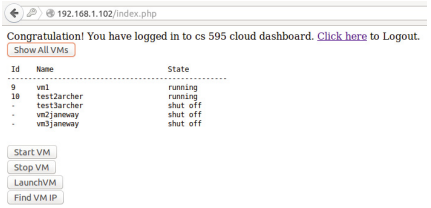


Fig. 3 Viewing available VMs

Fig. 4 Start a VM

In the available list we can choose which VM to start or to stop. Select the VM and perform the start or stop functionality at the buttons below the list. Refer to the Fig.8. When we view the available VM again we can see the selected VM is up and running and we can shutdown running VMs as well. Refer to the Fig.9. and Fig.10.

We can launch a new VM [9] with the *Launch VM* button in the dashboard. We can see the script while the VM is being launched. We can find the IP address of the VM as well from the dashboard by the *find Vm IP* button.

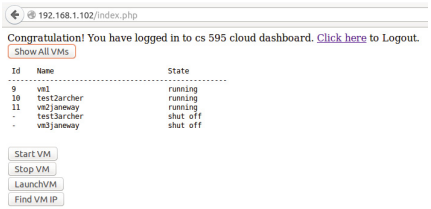


Fig. 5 VM is running

Fig. 6 Launch a VM

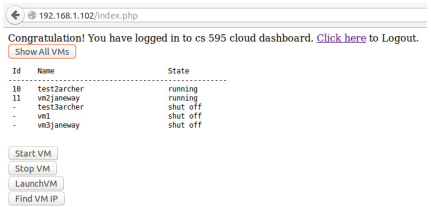


Fig. 7 Launch a VM

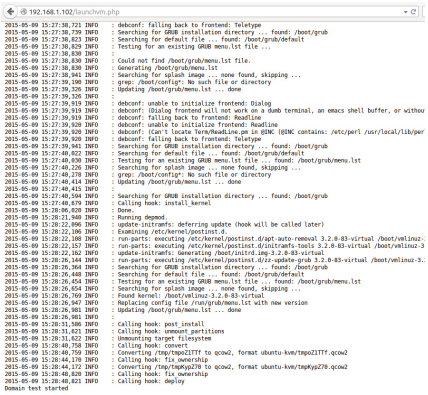


Fig. 8 Output of launching a VM

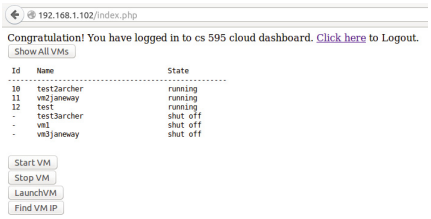


Fig. 9 New VM is running

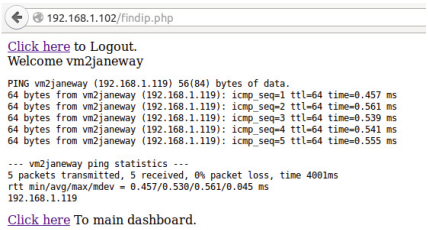


Fig. 10 Ping statistics of the VM

5 Conclusion and Future Work

This work involves the creation of virtual machines on Ubuntu servers using VM builder a command line interface. An end user may have difficulties in using the command line interface to create and manage virtual machines, so an idea was developed to create a web interface which eases the use of virtual machine creation and management. The web interface is developed using PHP, Ajax, JavaScript and PHP embedded shell script. The user can create, start, stop and find the IP address of any virtual machine hosted by the cloud architecture. The web interface is secured by the login feature which provides the user authentication as there is no RSA key

exchange between the guests and hosts in the current approach. The future work includes creating custom virtual machine images. There is a scope to design a web interface which can handle multiple servers in a cluster. Designing a load balancing system is much needed to handle these servers in a cluster which includes the limiting the number of virtual machines per server to optimize performance. There is a plan to include creation of VM images with predefined set of software packages. For example, an option would be given to the user to build a VM with default installation of apache hadoop framework [12].

References

1. Smith, G.: Wobblycogs -. Installing and Running KVM on Ubuntu 14.04 Part 1. Wobblycogs, October 27, 2014, Web, May 11, 2015
2. Ubuntu Documentation: KVM. N.p., n.d. Web, May 11, 2015
3. Digitalocean.com: How To Install Linux, Apache, MySQL, PHP (LAMP) stack on Ubuntu | DigitalOcean (2015). https://www.digitalocean.com/community/tutorials/how-to-install-linux-apache-mysql-php-lamp-stack-on-ubuntu (accessed May 12, 2015)
4. Ellingwood, J.: How To Configure the Apache Web Server on an Ubuntu or Debian VPS | DigitalOcean. Digitalocean.com (2015). https://www.digitalocean.com/community/tutorials/how-to-configure-the-apache-web-server-on-an-ubuntu-or-debian-vps (accessed May 11, 2015)
5. Help.ubuntu.com: PHP5 - Scripting Language (2015). https://help.ubuntu.com/lts/serverguide/php5.html (accessed May 11, 2015)
6. Hertzen, N.: php.js - PHP VM with JavaScript. Phpjs.hertzen.com (2015). http://phpjs.hertzen.com/ (accessed May 11, 2015)
7. Linuxjournal.com: Virtualization with KVM | Linux Journal (2015). http://www.linuxjournal.com/article/9764?page=0,1 (accessed May 11, 2015)
8. Folgar, F., Loureiro, A., Pena, T., Zablah, J., Seoane, N.: Implementation of the KVM hypervisor on several cloud platforms: tuning the apache cloudstack agent. In: 2014 IEEE Intl. Conf. on High Performance Computing and Communications, 2014 IEEE 6th Intl. Symp. on Cyberspace Safety and Security, 2014 IEEE 11th Intl. Conf. on Embedded Software and Syst (HPCC,CSS,ICESS) (2014)
9. Ferril, P.: Scripting KVM with Python, Part 1: libvirt. Ibm.com (2015). http://www.ibm.com/developerworks/library/os-python-kvm-scripting1/$#$list5 (accessed May 11, 2015)
10. Molenaar, R.: ARP (Address Resolution Protocol) explained - Networklessons.com. Networklessons.com (2013). https://networklessons.com/cisco/arp-address-resolution-protocol-explained/ (accessed May 11, 2015)
11. qumranet: KVM: Kernel-based Virtualization Driver (2006)
12. Analysis of virtual machine creation characteristics on virtualized computing environment. In: 2011 7th International Conference on Networked Computing and Advanced Information Management (NCM), Gyeongju (2011)
13. Taddia, C., Mazzini, G.: DNS reverse lookup statistics. In: 2005 International Conference on Wireless Networks, Communications and Mobile Computing (2005)
14. Khosla, R., Fahmy, S., Hu, Y.: Content retrieval using cloud-based DNS. In: 2012 Proceedings IEEE INFOCOM Workshops (2012)

Towards Providing Resource Management in a Local IaaS Cloud Architecture

Travis Brummett and Michael Galloway

Abstract Cloud Computing is a rapidly growing branch of distributed computing. A vertical implementation of a cloud architecture could be used to replace a traditional computer lab within an educational setting. However, to do this the architecture requires a middleware that can communicate across nodes. This paper discusses a middleware developed in Python that uses sockets to communicate between compute nodes and the head node within such an architecture. Specifically, the middleware uses a socket connection between a client program installed on the head node and server programs installed on each compute node to poll the compute nodes for information. It then uses that information to carry out a load balancing algorithm that checks the available resources on each compute node and starts a virtual machine(VM) on the node with the most available resources. This paper will discuss in detail how these functions are accomplished.

1 Introduction

Before we can discuss load balancing we first must understand the basics of cloud computing. There is not one single definition for cloud computing. In fact it is defined in different ways by a multitude of different people. However, the National Institute of Standards and Technology (NIST) defines cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable resources that can be rapidly provisioned and released with minimal management effort or service provider interaction." [1] The NIST also identifies five core characteristics of cloud computing. The first is on-demand self service. Users must be able

T. Brummett · M. Galloway(✉) ·
Department of Computer Science, Western Kentucky University,
Bowling Green, KY 42101, USA
e-mail: jeffrey.galloway@wku.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_37

to acquire access to resources without interaction with the service provider. A cloud must also have broad network access. Resource pooling is also highly important in cloud computing. Resources must be available to multiple users at the same time. Rapid elasticity is also a key characteristic. A cloud needs to be able to grow and shrink in the number of resources in can provision based on demands. Typically, cloud computing is provided in a pay-per-use model like a utility, meaning that there must be some form of service management for a cloud.

There are three commonly used service models. They are Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). There are also a few common deployment models. The first, a public cloud is available to a large group. The second is referred to as a private or local cloud. A cloud using this deployment model is provided by either an internal or external source for a specific group. A combination of the previous two deployment models is referred to as a hybrid cloud. Finally, a community cloud can be made up of both public and private clouds and provides services to multiple organizations supporting the same community [1], [2].

The purpose of this research is to develop a IaaS vertical cloud architecture that can replace traditional computer labs through the use of virtual machines(VMs) and thin clients. However, before we could do that it was necessary to develop a middleware to manage and perform load balancing for our architecture. Currently, we have an architecture that can support up to 50 active VMs at a time. In this paper we will discuss our architecture and the middleware that was developed to manage it.

In section 2, this paper will discuss the purpose of load balancing and the previous research conducted on the topic. Section 3 will discuss the importance of scalability and its execution within the proposed middleware. In section 4, focus will shift to the middleware itself. The topics will include a brief overview and the setup of the hardware and related software that was required to build the middleware. The server side program which was installed on each each compute node will then be discussed. Following that, the primary functions within the head node's client program will be discussed in detail. Section 5 will outline the future plans for this research and section 6 will be the conclusion.

2 Related Work

The authors of [3] state that the two main tasks of load balancing are resource allocation and scheduling in a distributed environment. The authors then compare several load balancing algorithms. They conclude that a static load balancing scheme is the easiest to simulate and monitor but due to it requiring prior knowledge about the resources, all the resources must be homogeneous. This means that such a structure couldn't scale to the heterogeneous model of the cloud. Thus, the authors of [3] conclude that dynamic load balancing techniques in distributed or hierarchical environment provide better performance. They also state that performance of the cloud

computing environment can be further maximized if dependencies between tasks are modeled using workflows.

[4] states that there are five metrics of load balancing. The authors say that most current load balancing techniques adequately focus on performance, response time, scalability, throughput, resource utilization, fault tolerance, migration time and associated overhead. However, the authors of [4] and the authors of [5] both feel that energy consumption and carbon emission should also be considered. These authors claim that this is necessary in a continued pursuit to develop green load balancing methods.

[6] discusses various load balancing algorithms before suggesting one of their own. The authors discuss using a greedy algorithm for load balancing. They say that while the greedy algorithm is simple and requires little processing, it leads to poor resource utilization. This is due to the fact that it uses the first available resources it comes across which means that all the resources are used on one node at a time. The authors then discussed the Round Robin. According to the authors the advantages to this method are that it utilizes all the resources in a balanced order and that virtual machines(VMs) are evenly distributed across nodes. The authors propose the use of a weighted Round Robin algorithm. This proposed algorithm uses a weighted table to help select a node. The idea behind this method seems to be similar yet slightly different to the load balancing method discussed in section 4.

3 Scalability

Scalability is an important characteristic within cloud computing architectures. In fact, one of the main ideas of cloud computing is having resources to provision on-demand. Thus, it becomes necessary to be able to expand or reduce the number of machines within a particular architecture. This means that a middleware also has to be able to adjust to a change in the number of nodes within an architecture.

While there are a multitude of ways to accomplish this, the use of a database seemed to be the best solution and is the approach used within this middleware. If the need arises for more machines, an administrator using this architecture and middleware could simply hook the machine into the router and get an IP through DHCP. It does not matter what size memory the machine has or how many cores it runs as long as it supports virtualization. The administrator just needs to decide how many virtual machines(VMs) that the machine will allow to be launched per each core. The number that the administrator selects cannot be greater than 8 VMs per core due to a limitation within KVM itself. This number along with the new node's name and IP are then simply added to a table in a database. This allows the middleware to access the information it requires. A more detailed discussion of the database will follow.

4 Middleware

4.1 Overview

The middleware discussed in this chapter was developed in Python and uses sockets to communicate. Python was chosen for the middleware's development due to how easily sockets can be implemented. Python also offers the ability to use a list of lists. This is important because the middleware uses a list to hold lists of information regarding each node. The ability to easily parse strings that are returned from commands passed to the compute nodes also aides in the load balancing process. Comparisons are also easy to do using these Python lists. Python has yet another benefit. It allows more than one value of varying types to be returned from methods which made it the ideal language for developing the middleware.

The middleware is comprised of two parts. The first is the *viceclient.py* which accesses a database to get a list of nodes. It uses the list of nodes to poll each one for information by passing a series of commands through the socket. It then uses a load balancing algorithm to select a virtual machine to launch and a node to launch it on. If there is a free VM that is started, *viceclent.py* prints out the IP associated with the VM. If all nodes are full, a message is printed instead. The *viceclient.py* is executed on the head node.

The second half of the middleware is called *viceserver.py*. It runs on the compute nodes and listens for commands from *viceclient.py*. The compute nodes are always running the *viceserver.py* program. Thus, they are always listening for the client-side's commands. When the *viceserver.py* receives a command it passes back the results to the *viceclient.py*. These two programs will be explained in more depth later.

4.2 Architecture

Before we can discuss the load balancing software we must first describe the architecture upon which it was deployed. Our architecture is made up of a head node and five heterogeneous compute nodes. The head node has an Intel Core i3 dual core processor and 8GBs of RAM. The two larger compute nodes have Intel Core i5 quad core processors and 16GBs of memory. The last three compute nodes have an Intel Celeron dual core processor and varying amounts of memory. Ubuntu 14.04 LTS Server was installed on each machine.

4.3 Compute Nodes

The compute nodes serve as the machines that launch and host the virtual machines. It is the resources found on these nodes that the virtual machines utilize. Therefore, they need to have virtualization extensions enabled. Compute nodes also need to be

able to receive commands from the head node. To do this the middleware requires that the python script entitled *viceserver.py* which causes each compute node to listen for commands from the *viceclient.py* program executed from the head node. However, to adequately use the commands that are required by the middleware some software had to be set up first.

Setup. The setup process was the same across all compute nodes. After the hardware itself was hooked into the network, Python3, vm-builder [7], libvirt and kvm [8] were installed. Yet another tool, arp-scan, was downloaded using the command *sudo apt-get install arp-scan* [9]. This tool will later be used by the viceclient.py program to get the IP addresses assigned to VMs. However, to be useful the sudoers on each compute node had to be given passwordless privileges. To do so I placed the line *ALL ALL=(ALL) NOPASSWD:ALL* in each node's sudoers file.

Networking. At this point any virtual machines created on these nodes will only be accessible on the local node. The virtual machines need to be able to access the router and be assigned an IP through DHCP. Thus it is necessary to create a virtual bridge and define a new network setting for the virtual machines. The guides provided by [10] provided instructions to add a network bridge to the */etc/network/interfaces* file. Next, the resources of [11] where used to create a XML file for a new default network. The contents of the file can be seen in figure 1. Then the commands to create a new network for virtual machines were issued. They were *virsh net-destroy default*, *virsh net-undefine default*,*virsh net-define network.xml*, *virsh net-start default*, and *virsh net-autostart default*. The first two commands stopped the network that would be utilized the virtual machines and undefined the previous network settings. The next command defined new network settings for the VMs using the *network.xml* file. The final two commands started the network defined by the previous command and switched on its auto-start option.

```
<network>
    <name>default</name>
    <forward mode="bridge"/>
    <bridge name="br0"/>
</network>
```

Fig. 1 *network.xml* File.

Virtual Machine Creation. After the network has been setup and arp-scan has been installed, Virtual Machines are defined on each compute node using a variation of the commands found at [7] and [12]. This means the middle-ware does not have to wait to define the VM and that it can simply start it. At this point the process changes slightly between each node. This is due to the fact that the limit of virtual machines per core on each node was decided before hand and the amount of cores

vary between machine. However, each node had the vmbuilder command ran on it the maximum number of times. Each VM had a unique name starting at vm1 and ending at vm50 with a directory of the same name. This naming scheme allowed each VM and directory to be uniquely named across the entirety of the cloud. The command was slightly modified each time. Once all the VMs had been defined the *viceserver.py* file was moved onto each node and started.

viceserver.py. The purpose of the *viceserver.py* program is to open a Python socket up to listen for commands from the *viceclient.py* program. When commands are received by the *viceserver.py* program the results are passed back to the *viceclient.py* socket and the server simply resumes waiting. After the program is launched on each compute node, they are simply left in their listening state indefinitely. Figure 2 shows the Python code which was used.

```
import socket
import time
from subprocess import PIPE, Popen

def main():
 host = ''    # Symbolic name meaning all available interfaces
 port = 12345  # Arbitrary non-privileged port

 s = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
 s.bind((host, port))
 s.listen(5)

 while True:
   conn, addr = s.accept()
   print('Connected by', addr)
   data =cmdline(conn.recv(1024).decode())
  #if not data: break
   conn.send(data)
   conn.close()

def cmdline(c):
 process=Popen(
         args=c,
         stdout=PIPE,
         shell=True)
 return process.communicate()[0]

main()
```

Fig. 2 *viceserver.py* Contents.

4.4 Head Node

The head node serves to manage compute nodes. It issues commands to the compute nodes and launches virtual machines on the compute node that is best suited to host it. However, it does not host any VMs itself. The head node also store the database

containing the information related to the compute nodes and hosts the web server that will be used to create the cloud architecture's user interface. After the *viceserver.py* program is waiting for commands on all the compute nodes, the head node can issue commands from the *viceclient.py*. However, to maintain scaliability, the head node also requires some setup.

Setup. The head node also requires that Python3 be installed. In addition the head node needs a way to manage a database through Python code. Thus, Once this Psycopg was installed using the *sudo apt-get install python3-psycopg2*. Psycopg is a PostgreSQL adapter for the Python programming language [13]. After this installation was complete, Psycopg could be imported into our Python program allowing it to access the vice database and run queries on it.

Database. In order to keep the load balancing algorithm dynamic and scalable, the *viceclient.py* program was designed to read information from a database. Thus, PostgreSQL was installed onto the head node. Using the SQL commands found at [14], a database called *vice* was created. Within the *vice* database a table called *nodes* was created. The nodes *table* consisted of three attributes, the node name, the node IP, and the number of VMs allotted per each core(npc). The last of which cannot exceed 8 due to KVM limitations. Each compute node was entered into this table with its own related information. Whenever a new node is added all an administrator would have to do is set it up as discussed previously and add its information into the *nodes* table.

In addition to the *nodes* table, the *vice* database also has a table named after each individual compute node. These tables consist of the names and MAC addresses associated with the VMs on that compute node. After each VM was created, the command *sudo virsh dumpxml vmname | grep 'mac address'* where vmname was actually the VM's name was ran on the host compute node. This command retrieved the MAC address from the XML file. These were the values used to fill the database tables. This allows for resource expansion within previously utilized hardware and since the MAC addresses assigned to the VMs are static they can be used with arp-scan to find the machine's IP address.

viceclient.py. Now that the database is created, *viceclient.py* can be ran from the head node. The first thing the program does is call the *loadBalance* method which can be seen in figure 3. The first thing it does is call the *poll* method to get information about each node. The purpose of *poll* is to build a list of lists which contains information regarding each node. It can be seen in figure 4. The *poll* method first calls *sqlops*. It passes the string "nodes" to the function call. This allows the *sqlops* method access to the *nodes* table where it builds a list of tuples containing each entry regarding the information for each node entered into the *nodes* table. The *(*sqlops) returns a list containing the node name, IP address, and number of VMs per core(npc). The *(*poll) method then uses the IPs in this returned list to expand upon the information regarding each node.

```
#Description: A function that uses data returned from other functions to decide which node to launch a vm on and then decides which vm to launch.
#Input: None
#Output:None
def loadBalance():#will be called through a web interface.
 glist=poll()#calls poll() and stores into into glist.
 for i in range(0,len(glist),1):#An embedded for loop is used to compare the usage percentages of nodes.
  for j in range(0,len(glist),1):
   if (glist[i][9]<glist[j][9]):#The node with the lowest usage percentage is moved to the front of the list.
    glist[i],glist[j]=glist[j],glist[i]
   elif(glist[i][9]==glist[j][9] and glist[i][5]>glist[j][5]):#If usage percentages are equal the amount of free memory is used to dictate which node goes first. The node with the most free memory goes to the first position.
    glist[i],glist[j]=glist[j],glist[i]
 if(glist[0][9]==1):#After the nodes have been sorted the best availiable node will be in position 0 of the list. If it's percentage used is 1 or 100% we know all the nodes are full. Thus, a statement is printed.
  print("Sorry, all resources are in use.")
 else:#Otherwise, we have our node and begin to determine which vm to launch.
  vmList=sqlops(glist[0][0])#Passing in the node name to the sqlops function, we store a list of the nodes possible vms and their MAC addresses.
  rL=virshList(glist[0])#A virsh list command creates a list of VMs that are already running.
  vm=vmList[0][0].strip()#the vm variable is used to hold the name of the vm to be started.
  mac=vmList[0][1].strip()#The mac does the same only for the MAC.
  for i in range(0,len(vmList),1):#Embedded forloops are used to compare a two latest lists against each other.
   flag=False#A flag is set to signify if a vm has been found or not.
   for j in range(2,len(rL)-1,1):#The vmList is compared to the running vm list.
    if(vmList[i][0].strip()==rL[j][1]):#If we find the current vm we are looking at in the running list we cannot launch it.
     flag=True#Thus, our flag becomes True.
     break#breaks out of the inner loop and moves on to the next possible vm.
   if(flag==False):#If the flag is False.
    vm=vmList[i][0].strip()#The vm variable recieves the name of the vm that wasn't running.
    mac=vmList[i][1].strip()# Same with mac.
    break#break out of outter loop.
  virshStart(glist[0][1],vm)#Calls virshStart with the nodes' IP and the vm to start.
  print("vm started: ",vm)#Prints
  flag2=True#A flag to keep track if we have found the IP of the started vm.
  while(flag2):#While the flag2 is true
   time.sleep(2)#We have a 2 second delay because it takes a moment for the vm to start and get an ip.
   arpL=arpScan(glist[0])#calls arp--scan on the node to get vm macs and ips.
   for i in range(2,len(arpL)-3,1):#Searches the list.
    if(mac==arpL[i][1]):#When our MAC is found we print the associated IP.
     print("SSH at: ",arpL[i][0])#Prints the ip
     flag2=False#changes flag to exit while
     break#breaks for
```

Fig. 3 Method: *loadBalance*.

The *poll* method next calls *free*, *lscpu*, and *virshCount*. It uses these methods to add on to the inner lists that represent each individual node within the bigger list. These three functions are extremely similar except for the location of the information added to the list. The *free* method returns the amount of total memory, used memory,

```
#Description: This function is used to cycle through the nodes to get the
    ↪need information from the database and combining it with the
    ↪information gathered by the system calls to make a master list.
#Input:None
#Output:List
def poll():
 lst=sqlops("nodes")#lst is filled with the info placed into the db table "nodes
     ↪"
for i in range(0, len(lst),1):#for loop goes through each element in the list which
     ↪represents each row of a db.
  port = 12345#abritary port number for socket use
  total,used,fm=free(port,lst[i])#info from free command placed in variables
  cores,vms=lscpu(port,lst[i])#same as above only from lscpu function
  VMCount=virshCount(port,lst[i])
  node=list(lst[i])
  node.append(int(total))#the info gathered from the previous two lines is added one
      ↪ by one to the node element in lst
  node.append(int(used))
  node.append(int(fm))
  node.append(int(cores))
  node.append(int(vms))
  node.append(int(VMCount))
  node.append(VMCount/vms)
   lst[i]=node
 #lst is returned
return lst
```

Fig. 4 Method: *poll*

and free memory. The *lscpu* method returns the number of cores and the max number of virtual machines. The latter is the number of cores multiplied by the npc entered into the *nodes* table. Finally the *virshCount* method returns the number of virtual machines listed as running. The *poll* method adds these items to the list at each index along with a usage percentage in decimal form determined by the number of running VMs divided by the maximum VM total. *Poll* then returns a list in which each index is a list of information regarding one compute node. For example, index zero of the returned list would refer to a list containing the first compute node's name, IP, npc, total memory, used memory, free memory, maximum number of VMs, number of VMs running, and usage percentage in decimal form.

Now that the *loadBalance* method has all of this information, it can start comparing the nodes based on their usage percentage. The method uses a simple sorting pattern and sorts the lists on their usage percentage found at the 9th index. If the usage percentage is the same for two nodes the method looks at the 5th index where free memory is located and gives priority to the node with the most free memory. After the nodes are sorted, the node that is best suited to launch a VM is located in position 0 of the main list. This node's usage percentage is then compared against 1. If the percentage is equal to 1 that means all the virtual machines are in use and a message is printed to the user.

However, if the node is not full, the *loadBalance* calls the *sqlops* method with the name of the selected node. This creates a list of virtual machines and MAC addresses from the database table with the same name as the node. The *loadBalance* method then calls *virshList* with the list related to the selected node. This method creates a list of currently running VMs. A for loop then uses a flag to compare the two new

lists. Whenever a VM is found in the list from the database but not in the list from the *virshList* that VM is selected to be started. The loops are broken and the *virshStart* method is called with the node IP and the VM name as arguments. The *virshStart* method simply passes the command to start the VM it is given.

After the start command has been sent, the *loadBalance* sets a new flag to the boolean value True. It then enters a while loop for as long as this flag remains True. Inside the while loop the *loadBalance* method waits two seconds before calling the *arpScan* method. The delay is to give the virsh start command time to run. The *arpScan* method is passed the list containing the node information and uses it to send the arp-scan command to that node. A list containing all the IPs that are in use and the MACs that are using them is returned. This list of MACs is then compared against the MAC associated with the VM that was launched. If the MAC is not found in the list, the VM does not have an IP yet and the while loop iterates once more. If the MAC is found the while loop is broken and the IP is printed. The user can now SSH into the VM using the given IP.

5 Future Work

Currently, the middleware is executed manually. In the future a web interface will be build around it that will allow the methods within to be called. There will be two types of user an administrator and a student. The student will only be able to start a VM through the load balancing method. An administrator will be able to launch any VM he desires and poll any node for its information. Work could also be done to make the load balancing process more efficient and green. The eventual goal of this research is to replace a traditional computer lab with a cloud architecture and thin clients.

6 Conclusion

The middleware discussed in this paper can be used to adequately manage a cloud. It is highly scalable due to the use of a database. It can also adequately load balance by polling each node in the database for information and choosing the best one by comparing the usage percentage and free memory of each. Once a user interface is built, this architecture can be used along with thin clients to create a computer lab for use in an educational setting.

References

1. Madisetti, V., Bahga, A. (ed.): Cloud Computing A Hands-On Approach (2014)
2. Smith, K., Galloway, J., Vrbsky, S.: A survey of cloud computing architectures. Master's thesis, The University of Alabama

3. Katyal, M., Mishra, A.: A comparative study of load balancing algorithms in cloud computing environment (2014). http://arxiv.org/ftp/arxiv/papers/1403/1403.6918.pdf
4. Chana, I., Kansal, N.J.: Cloud load balancing techniques: A step towards green computing. IJCSI International Journal of Computer Science Issues **9**(1) (2012)
5. Mohan, K.G., Megharaj, G.C.: Two level hierarchical model of load balancing in cloud. International Journal of Emerging Technology and Advanced Engineering **3**(10), October 2013
6. Biradar, S., Supreeth, S.: Scheduling virtual machines for load balancing in cloud computing platform. International Journal of Science and Research (IJSR) **2**(6), June 2013
7. Creating kvm machines with boxgrinder and vmbuilder. http://www.admin-magazine.com/Articles/Building-Virtual-Images-with-BoxGrinder-and-VMBuilder
8. Kvm/virsh. https://help.ubuntu.com/community/KVM/Virsh
9. Arp-scan user guide
10. How to configure a linux bridge interface. http://xmodulo.com/how-to-configure-linux-bridge-interface.html
11. libvirt.org. https://libvirt.org/
12. Virtualization on ubuntu 14.04 with virsh and vmbuilder. http://blog.viktorpetersson.com/post/108451140634/virtualization-on-ubuntu-1404-with-virsh-and
13. Psycopg postgresql database adapter for python. http://initd.org/psycopg/docs/install.html
14. Postgresql tutorial. http://www.tutorialspoint.com/postgresql/index.htm

Integration of Assistive Technologies into 3D Simulations: An Exploratory Study

Angela T. Chan, Alexander Gamino, Frederick C. Harris Jr. and Sergiu Dascalu

Abstract Currently, there are limited, commercially available video games for people with disabilities. Sim-Assist is a software system that aims to allow people with disabilities to interface with a three-dimensional simulation game of Air Hockey. This is accomplished through various integrated assistive technologies, such as brain-computer interfacing, voice commands, and speech-to-text capabilities. With Sim-Assist, users are able to play Air Hockey without depending on sight, and in a hands-free manner. We conducted an exploratory study to provide the foundation for integrating assistive technologies in 3D simulations, including scientific simulations and serious games. In this paper, we introduce the research and development of assistive technologies, focusing more on the BCI software component; outline our system design and implementation; provide a short walkthrough of the interface; and briefly discuss our preliminary results.

Keywords Assistive technologies · Brain-computer interfaces · Speech-to-text · Air hockey · 3D simulation

1 Introduction

According to the 2014 Disability Statistics Annual Report, the average percentage of disability in the United States population in 2013 was 12.7%; this translates to approximately 41 million people in the United States with some disability [1]. Recently, assistance for people who are care-dependent has gained importance as a research field because it could potentially improve the quality of their lives as well as their relatives'. Blindness and vision impairment also increases rapidly as people grow older, particularly after the age of 75 [2].

A.T. Chan(✉) · A. Gamino · F.C. Harris Jr. · S. Dascalu
Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA
e-mail: {achan,agamino}@nevada.unr.edu, {fredh,dascalus}@cse.unr.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_38

Software-based assistive technology (SBAT) research is the main motivation for this paper. Our system, Sim-Assist, aims to integrate SBATs into a 3D simulation. The main goal of Sim-Assist is to allow people with disabilities to play the 3D virtual game Air Hockey. In the near future, we plan to use Sim-Assist as a foundation to integrate assistive technologies (AT) into 3D scientific simulations.

This paper aims to provide an exploratory study of the possibilities of integrating assistive technologies into a 3D environment. In Section 2, we investigate the surveyed research and development in SBATs. In Section 3, we outline the system design and implementation of our system. In Section 4, we provide a brief walkthrough of Sim-Assist's interface. In Section 5, we discuss our preliminary results. Finally, we provide our concluding statements and discuss the directions of our future work in Section 6.

2 SBAT Background

Software-based assistive technologies are products that offer people with disabilities better accessibility to computers. For those who are unable to see, hear, speak, or lack motor skills, assistive technologies can alleviate their lives a little more. Some of their distinguishing categories include brain-computer interfacing, gesture recognition, and speech recognition. Brain-computer interfacing mainly targets users who lack motor skills, but can also be used by users who are unable to see, hear, and speak; gesture recognition targets users who are unable to see, hear, or speak; and speech recognition targets users who are unable to see and possibly those who lack motor skills.

One of the main challenges pertaining to developing software for assistive technologies (ATs) is the unsuccessful collaboration between developers and disability specialists. The disjoint between the functionality requirements of disabled users and the efforts of system developers creates poor usability of the end product for the users. Another challenge is that the complex solutions are not necessarily the best. The more tools that are embedded in a program, the more complex the system becomes, making it more difficult for users to learn. There also seems to be a large spectrum in the research and development of assistive technologies because there are various types of disabilities. This ranges from aiding students learn in the academic environment (applies to those with learning disabilities, blindness, and deafness) to assisting the elderly.

Research in software-based assistive technologies is important as society grows increasingly more technologically advanced. As computers are meant to improve the quality of our lives, they are used daily. Therefore, it makes sense to use computers as one of the tools, if not the main tool, for assisting people with disabilities. In order to keep up with the advancements in everyday technology, research in the software for such computers for assistive technologies becomes important. The most distinguishing characteristics of software-based assistive technologies are the abilities to convert text from the computer to auditory speech and vice versa, and respond to: speech commands, gesture commands, and certain brain

signals. These substantial advancements within this area would benefit people who have a disability such as: auditory, hearing, speech, or motor impairments.

2.1 Related Research

Assistance for people who are care-dependent has gained importance as a research field because it could potentially improve the quality of their lives as well as their relatives. Such people may be dependent due to motor, sensory, or cognitive impairments. As a result, performing daily life activities such as feeding, hygiene, personal care, etc. may be difficult. Gomez et al. presented and evaluated a system that shows adaptive manuals for daily life activities for people with disabilities [3]. This proposed system is based on mobile devices and Quick Response (QR) codes, and will help with the rehabilitation process of patients with acquired brain injury. Additionally, Donoghue et al also aimed to create a system for people with motor disabilities [4]. Specifically, they developed a human application of neural interface systems (NICs) for people with paralysis. However, the NIC is in its early stages of development. The pilot trial results provided proof that a neuron-based control system is feasible. Thus, in this system, signals can be detected, decoded, and used for real time operation of computer software, ATs, and other devices.

According to the 2012 National Health Interview Survey, 20.6 million American adults reported experiencing vision loss [2]. Individuals who experience vision loss can be referred to those who claim that they have difficulty seeing, even while wearing glasses or contact lenses, and people who are blind or unable to see at all. Blindness and vision impairment also increases rapidly as people grow older, particularly after the age of 75 [5]. At this age, people will struggle to learn Braille and become accustomed to visual impairment without dependence on others. With this issue as motivation, Narasimhan et al developed system to provide the visually impaired an independent shopping experience [5]. This system will allow the visually impaired shopper to find the correct aisles and shelves that contain the desired product. The system can also distinguish between different products of the same type (e.g., Marinara sauce vs. Alfred sauce) once the general shelf location of the desired product has been identified. This system can also be used for accomplishing other daily tasks such as cooking.

Although there are different levels of disability, assistive technologies should not solely target severe disabilities. According to Edyburn, little attention has recently been dedicated to the assistive technology needs of students with mild disabilities [6]. Mild disabilities may consist of learning disabilities, behavioral disorders, and mental retardation. Some tools such as iPing, Co:Writer, and WebMath have been developed for these targeted users, but there is little evidence suggesting students with mild disabilities have access to such technologies [6]. In addition, due to the lack of performance measurement tools for assessing users' progress in the academic field, it is difficult to justify claims that the technology will enhance performance.

There have also been efforts to develop assistive technology for the elderly. So far, assistive technology targeted towards those who are older has yet to improve the quality of their lives. To target this issue, Bright and Coventry suggested potential design strategies to maximize a product's usability and usefulness for older adults [7]. By doing so, their main goal was to show the importance of considering both psychological and socio-emotional design requirements when creating such technology.

2.2 Development Efforts

It appears that research and development efforts have not yet advanced to a stage that allows people with disabilities to interact with 3D environments. However, there are existing user interfaces (UIs) that allow people with disabilities to interact with 2D environments in such a way that is more user-friendly and accessible for them. Through the developments of these existing works, we may see user interfaces for 3D environments in the future. We explore these mentioned works and the methods that were used in such interfaces.

Visual Impairments

There are various degrees of visual impairment, also described as "low vision" [8]. Variations of low vision may include: a diminished acuity; the loss of one's field of vision; one's sensitivity to light; and distorted vision or loss of contrast. Currently, a large selection of assistive technologies is made commercially available to the general public such as screen readers and magnifiers, braille keyboards, and text-to-speech software. However, these generic products only solve he most basic problems that one with low vision must encounter. Researchers are now exploring alternative methods that can further expand the limited capabilities of these people.

Auditory Interfaces

People who are blind depend heavily on their sense of hearing. Previous research in spatial audio revealed that it has potential as a new medium for creating nonvisual interfaces [9]. Researchers Winberg and Hellstrom used spatial audio to create an auditory interface that utilizes direct manipulation for graphical user interfaces (GUIs) to become more accessible for the blind [10]. Although direct manipulation improves the average user's usability and learnability of a system, it is not available in modern screen readers that most blind users depend on [10].

Previous research in this type of interaction for blind users revealed systems that include using tactile devices with audio [11, 12] and a system that uses 3D audio and a data glove [13]. However, in these systems, only a subset of the interface objects are presented to the user, and the user has to browse the auditory space in order to get an overview of the object. Therefore, the user is interacting directly with interface objects instead of using direct manipulation, according the definition mentioned earlier. Another system that uses auditory interfaces is a system developed by Pitt and Edwards [14]. In this system, the cursor becomes a "virtual microphone" that is used for selecting items from a menu [14].

In order to determine if it is possible for a UI to utilize auditory direct manipulation, and if this method of interaction is worth pursuing, Winberg and Hellstrom implemented an auditory form of the game "Towers of Hanoi" [10]. Each disc had a corresponding sound that was different in pitch and timbre. The smaller the disc, the higher the pitch. The sounds also had slightly different tunes with respect to each other to further distinguish the size of the disc. The tower the disc is located on would be determined by the various levels of stereo panning and amplitude envelopes. Furthermore, the vertical position of the disc is distinguished by the length of the sound. In this system, the mouse is similar to the one used by Pitts and Edwards. Users can move the cursor to each tower based on differences in volume. Results from testing this system revealed that auditory direct manipulation is possible and possesses potential of being significant for blind computer users [10].

Custom Interfaces

Most graphical user interfaces are designed to only cater towards the general public. However, there are some cases where the interface can be easily rendered to suit one's unique capabilities. For users with restricted vision, enlarging the fonts and visual cues appropriately may suffice. Similarly, for users who partially lack motor control, it may be easier to interact with interfaces that incorporate widgets with large targets so that dragging is unnecessary. In cases where the use of ATs are required, a specific GUI design may be desired. For instance, it might be more user-friendly to design a hierarchical structured interface for navigation through UIs that utilize screen readers.

When designing a GUI, it is important to focus on the users' needs. However, it is unreasonable to design a specific GUI that handles each individuals' unique preferences. A paper written by Gajos et al proposes the idea of generating a custom user interface which would better suit users, especially for users with visual and motor impairments [15]. The authors present a system called SUPPLE++ that automatically generates UIs that caters specifically towards either an individual's motor or vision capabilities, or both. Their preliminary results indicated that their system allowed one user to complete tasks, which she could not perform using a standard interface. Compared to other users, the SUPPLE++ allowed the user to save time by 20% [15].

Motor Impairments

Most of the interactions between the user and the GUI rely on selection operations [16]. Actions such as inputting data can also be achieved by selecting various letters through on-screen keyboards. This type of selection operation is executed by pointing the cursor at a desired area of the screen, and clicking to indicate the selection of the item the cursor is pointing to.

For people with motor disabilities there are several devices that can be used as an alternative for the mouse. For example, the "Tonguepoint" [17], as the name suggests, is a device in which the user can operate the cursor using his/her tongue.

Results for users who have tested this system revealed that they were able to achieve a performance level that was only 5-50% slower than users who used a standard pointing device.

Another alternative device is the "Headmouse", which is a pointing device that moves the cursor based on head movements [18]. This is similar to Emotiv's Neuro Mousecontrol application that allows the user to control the cursor using head movements while he/she is wearing the Epoc headset [19]. However, whereas the Neuro Mousecontrol application depends on cognitive action to trigger left and right mouse button operations, the Headmouse device achieves these actions when the user dwells over a particular key on an on-screen keyboard for a set period of time, or by using a remote switch. There are also more complex approaches for computer interface functionality such as eye-tracking systems. These systems only depend on the user's ability to control the gaze of his/her eye. The cursor follows the eye's line of gaze, and the clicking operation is executed whenever the cursor remains in a location for a certain amount of time, which is similar to the Headmouse device.

Another complex approach is the Brain-Computer Interface that uses electroencephalogram (EEG) signals. However, Andrew Junker developed a device called Cyberlink that allows the user to control the cursor using both electromyography (EMG) (produced by muscle movements from the head and neck) and EEG signals (produced by thoughts) [20]. Unfortunately, this method of interaction hinders the overall control process of the cursor because intensive concentration is required, which may be difficult users.

Barreto et al. proposed a system that only utilizes EMG signals for cursor control [16]. In their approach, cursor movement depended on the natural and voluntary movements of the user's face such as clenching different sides of the jaw (left, right, and both) and eyebrow movement (up and down). These types of movements are usually controllable for people who are impaired from the neck down. The system was tested by six subjects that participated in 20 trials. In each trial, subjects were asked to click on the Stop button by moving the cursor from one corner on the computer screen to the center, where the button is located. Results have shown that the subjected required an average of 16 seconds for completing the task [16]. Though this time is notably greater compared to the amount of time it would take for an unimpaired subject to accomplish the same task using a mouse, the response time would still be usable with most standard GUIs. In addition, this proposed interface may potentially be more affordable and portable than other complex interfaces like the eye-tracking system.

Issues/Limitations

In the case of auditory interfaces, one issue is its scalability as the complexity increases. In Winberg and Hellstrom's Towers of Hanoi game, only five discs were used and users seemed to perform relatively well. However, Towers of Hanoi had rule restriction that were able to offer some guidance for the users. In a desktop interface, in which there will inevitably be a larger amount of objects, will blind users benefit from auditory direct manipulation?

Another issue in auditory interfaces is representing the selection state in a GUI. The selection state for when users execute selection operations such as highlighting text or selecting objects to drag and drop needs to be represented. In the game Towers of Hanoi, users were not able to recognize whether or not they were actually moving the disc to the next tower instead of just hovering the cursor over each tower. Furthermore, if large numbers of interaction objects are present, the user may be limited to his/her memory. To forget the location of an object would be time consuming for the user to constantly check each item present one by one for the correct object. Another issue would be using various sounds could cause the user to forget the meaning of each sound during the beginning stages of using the system. This means that auditory interfaces may be more difficult to learn for users.

For UIs that accommodate specifically to people with motor impairments, there may be some possible disadvantages depending on the category of motor disability. For instance, a user with cerebral palsy may not be able to operate the Tonguepoint device due to lack of motor abilities in the tongue. Similarly, a user with spinal fusion may not be able to move his/her head, thus unable to fully operate any of the mentioned BCI systems. However, for those who are capable of operating a BCI system, one limitation is the reaction speed of the cursor. In addition, it takes a considerable amount of concentrate to train and execute commands. If the user becomes distracted while training the system, it will be difficult to accurately perform a desired task. Furthermore, it is difficult for one to measure the effectiveness of a BCI system against a mouse when executing the same tasks. This is because Fitts' law is unable to describe the performance of people with disabilities [21].

Two major disadvantages of eye tracking systems are that they are significantly expensive and require a considerable amount of focus to control. If the user loses focus and unintentionally stares into one location on the UI, he/she may accidentally click the cursor. Another limitation is that the calibration of the device can be easily lost if the user changes position with respect to the screen [16]. Furthermore, gazing at a location that is out of the camera's field of vision will hinder the system's operation.

3 System Design and Implementation

Our system uses a 3D version of the Air Hockey game, implemented by us. Sim-Assist is innovative because it allows people to play an engaging game such as Air Hockey with reduced effort. In addition, there are no existing commercially available video games that incorporate brain-computer interface (BCI) technology.

The main functionalities of Sim-Assist include: the ability to move the Air Hockey paddle using BCI capabilities; an audio notification for each new score; and the ability to input player information, restart, and pause and resume the game using speech. The implemented requirements of the Sim-Assist system are shown in Table 1. The intended users of this system are people with visual and motor

impairments, or both. However, people without disabilities may also use Sim-Assist to experience a new form of gaming (BCI-based).

Sim-Assist was developed in C++ on a Microsoft Windows platform, using Microsoft Visual Studio 2012. To play the Air Hockey game, Sim-Assist requires the Emotiv Epoc headset for all BCI-related user inputs [19]. To record the user's voice for speech commands and speech-to-text capabilities, the Microsoft Kinect peripheral was used, with a dependency on the Microsoft Speech Platform.

The Sim-Assist system, from a high-level perspective, consists of different layers in its architecture, as shown in Fig. 1.

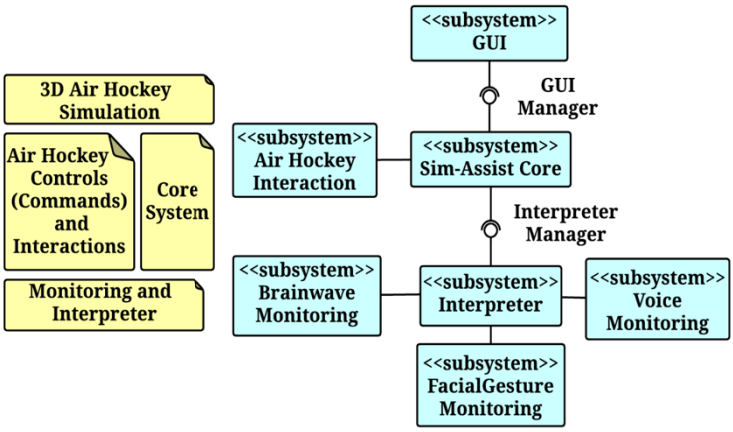


Fig. 1 System-level diagram of the Sim-Assist system. Layers are represented in yellow, and components are represented in blue.

Table 1 Functional Requriements For Sim-Assist

Functional Requirements
Sim-Assist shall: ● Display a 3D simulation game of Air Hockey. ● Allow the user to control the mouse using head movements. ● Grant the user the ability to control the mouse using various thoughts. ● Permit the user to use voice commands to pause/un-pause the game using speech. ● Notify when user scores a goal. ● Enable the user to toggle voice detection. ● Allow the user to input player information using speech. ● Present the option to select number of players. ● Provide the capability to quit the game using speech. ● Enable the user to enter/exit full screen mode using speech. ● Authorize the user to assign facial expressions specific to mouse control. ● Permit the user to assign thoughts to specific mouse control. ● Accept multiple users to play at one time.

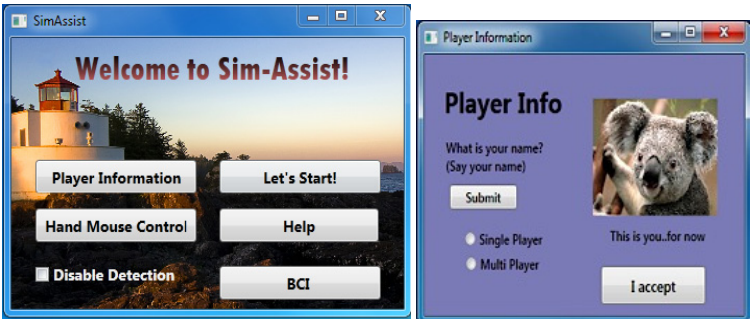
The lowest layer is the monitoring and interpreter, where Sim-Assist detects speech, facial and head gestures, and brainwaves. This is accomplished by the brainwave, facial gesture, and voice monitoring subsystems. The interpreter subsystem then takes these possible commands and determines if they are actual commands. If so, they are sent to the Sim-Assist core system. This system is responsible for all the major decisions within the system, including holding registered commands, as well as managing GUI information. The top layer of Sim-Assist is the Air Hockey game, which consists of the GUI subsystem that communicates with the Sim-Assist core system.

4 System Walkthrough

Upon startup, Sim-Assist will greet the user with a welcome window, as shown in Fig. 2a. The user must enter his or her information prior to playing Air Hockey, as shown in Fig. 2b. The system will prompt the user to enter his or her name verbally. A database of approximately 200 popular names (both genders) in the U.S. is used to help the system recognize the user's name. The user also has the option to select how many players will participate in the game (one or two).

Once player information is submitted, the user must click on the "BCI" button from the welcome window. The training session for the headset will launch, together with Emotiv's "Neuro Mousecontrol" application. Once the training session is completed, the user can map trained commands to the mouse control. This will allow the user to move the mouse, along with playing the game, using his or her thoughts.

Finally, the user can launch the game by clicking on the "Let's Start!" button. The user will be presented with an Air Hockey table, along with a scoreboard, two paddles, and a puck, as shown in Fig. 3.



(a) (b)

Fig. 2 a) Sim-Assist welcome window; b) Player information window.

It is important to note that our settings were that Player 1 could control the paddle using only his or her thoughts and facial expressions, while Player 2 can control the paddle only with keyboard presses. Either player could also: enter or exit full screen mode, as well as pause, resume, or restart the game using speech commands.

Fig. 3 Snapshot of 3D Air Hockey game

5 Preliminary Results

One limitation of our system was that Player 1 could use only BCI controls, while Player 2 could use only keyboard controls. Upon system testing, we realized that the Air Hockey game is too difficult to play with BCI controls alone. While playing the game, Player 1 was under too much stress and had difficulties executing certain thoughts accurately. Since Player 2 could control his or her paddle using only keyboard presses, this served as an unfair advantage. Consequently, in our future versions, we plan to fix the current player settings of the game.

One significant constraint in our system was that in order for Player 1 to use the Epoc headset, he or she needed to have little or no hair, as shown in Fig. 4. Otherwise, the sensors on the headset were unable to detect any signals.

Fig. 4 Player using an EEG headset to interact with the 3D Air Hockey game

Because the Epoc was used, an advantage of Sim-Assist is that there are many different types of thoughts the user can use to map to a specific mouse control. To compare, systems that use NeuroSky's "MindWave" headset will operate only if the user is in a focused or relaxed state [8]. Using this headset will be too restrictive for developers to create a robust system.

6 Conclusion and Future Work

Sim-Assist allows people with disabilities to interface with a three-dimensional simulation game of Air Hockey. Inspired by the HeadMouse and Barreto's system, Sim-Assist users can also interact with the interface using facial expressions and head movements. Our system utilizes brain-computer interfacing, voice commands, and speech-to-text capabilities, which target people with visual and/or motor impairments. We will leave the addition of text-to-speech capabilities for future work.

Sim-Assist serves as a foundation to integrate assistive technologies into a larger variety of 3D simulations. Once further improvements are made, we will be able to apply its concepts to 3D scientific simulations. We plan to use the concepts of AT-integration from Sim-Assist for the virtual watershed simulation environment that is part of the NSF-funded Western Consortium for Watershed Analysis, Visualization, and Exploration (WC-WAVE) project, as shown in Fig. 5 and Fig. 6 [22]. This will allow scientists and researchers with disabilities to interface with and navigate through a 3D environment.

We also strongly consider using a different EEG head-set. We hope that with the newest Emotiv Epoc model, Epoc+, we will have more accurate EEG readings and be able to write our own mouse control application.

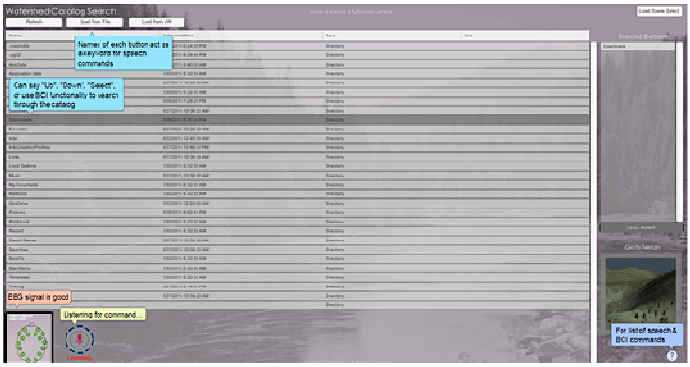


Fig. 5 Mock-up interface of Catalog Search in the Virtual Watershed Platform with AT-integration

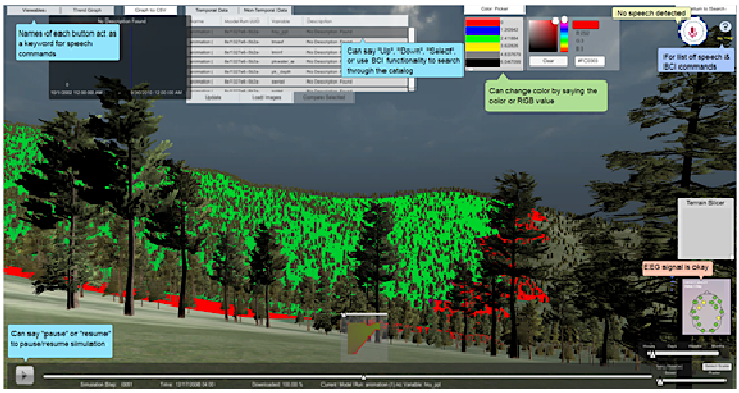


Fig. 6 Mock-up interface of the Virtual Watershed Platform with AT-integration

Acknowledgment This material is based upon work supported by the National Science Foundation under grant number IIA-1329469.

References

1. Annual Disability Statistics Compendium. http://disabilitycompendium.org/
2. Facts and Figures on Adults with Vision Loss. http://www.afb.org/info/blindnessstatistics/adults/facts-and-figures/235
3. Gomez, J., Montoro, G., Haya, P.A., Alaman, X., Alves, S., Martinez, M.: Adaptive manuals as assistive technology to support and train people with acquired brain injury in their daily life activities. Personal and Ubiquitous Computing **17**, 1117–1126 (2013)
4. Donoghue, J.P., Nurmikko, A., Black, M., Hochberg, L.R.: Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia. The Journal of Physiology **579**, 603–611 (2007)
5. Narasimhan, P., Gandhi, R., Rossi, D.: Smartphone-based assistive technologies for the blind. In: Proceedings of the 2009 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems, pp. 223–232. ACM (2009)
6. Edyburn, D.L.: Assistive technology and mild disabilities. Special Education Technology Practice **8**, 18–28 (2006)
7. Bright, A.K., Coventry, L.: Assistive technology for older adults: psychological and socio-emotional design requirements. In: Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA-2013), pp. 9–12. ACM (May 2013)
8. Low Vision Aids and Technology. Macular Degeneration Foundation. Accessed July 30, 2015. http://www.mdfoundation.com.au/resources/1/MDFLowVisionAids.pdf
9. Burgess, D.: Techniques for low cost spatial audio. In: Proceedings of the 5th annual ACM Symposium on User Interface Software and Technology (UIST 1992), pp. 53–59 (1992)
10. Weinberg, F., Hellstrom, S.O.: The quest for auditory direct manipulation: the sonified towers of hanoi. In: Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT 2000), pp. 75–81 (2000)

11. Mynatt, E.D., Weber, G.: Nonvisual presentation of graphical user interfaces: contrasting two approaches. In: Proceedings of the Human Factors in Computing Systems Conference, CHI 1994, pp. 166–172. ACM Press (1994)
12. Petrie, H., Morley, S., Weber, G.: Tactile-based direct manipulation in GUIs for blind users. In: Conference Companion on Human Factors in Computing Systems, CHI 1995, pp. 428–429. ACM, New York (1995)
13. Savidis, A., Stephanidis, C., Korte, A., Crispien, K., Fellbaum, K.: A generic direct-manipulation 3D auditory environment for hierarchical navigation in non-visual interaction. In: Proceedings of Assets 1996, pp. 117–123. ACM (1996)
14. Pitt, I.J., Edwards, A.D.N.: Navigating the interface by sound for blind users. In: Proceedings of the HCI 1991 Conference on People and Computers VI, pp. 373–383 (1991)
15. Gajos, K., Wobbrock, J., Weld, D.: Automatically generating user interfaces adapted to users' motor and vision capabilities. In: Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology, pp. 231–240. ACM (2007)
16. Barreto, A., Scargle, S., Adjouadi, M.: A practical EMG-based human-computer interface for users with motor disabilities. Journal of Rehabilitation Research and Development **70**, 53–64 (2000)
17. Salem, C., Zhai, S.: An isometric tongue pointing device. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 1997), pp. 538–539 (1997)
18. HeadMouse Extreme. Origin Instruments. http://www.orin.com/access/headmouse/ (accessed March 23, 2015)
19. Emotiv. http://emotiv.com (accessed May 10, 2015)
20. Junker, A., Sudkamp, T., Eachus, T., Mikov, T., Wegner, J., Edmister, E., Livick. S., Heiman–Patterson, T., Goren, M.: Hands-free computer access for the severely disabled (unpublished)
21. Pino, A., Kalogeros, E., Salemis, E., Kouroupetroglou, G.: Brain computer interface cursor measures for motion-impaired and able-bodied users. In: Proceedings of the 10th International Conference on Human-Computer Interaction, vol. 4, pp. 1462–1466 (2003)
22. Dascalu, S.: Scientific collaboration in virtual environments: the western consortium for watershed analysis, visualization, and exploration project. Invited talk. In: Proc. CTS 2014, pp. 560–561 (2014)

Data Profiling Technology of Data Governance Regarding Big Data: Review and Rethinking

Wei Dai, Isaac Wardlaw, Yu Cui, Kashif Mehdi, Yanyan Li and Jun Long

Abstract Data profiling technology is very valuable for data governance and data quality control because people need it to verify and review the quality of structured, semi-structured, and unstructured data. In this paper, we first review relevant works and discuss their definitions of data profiling. Second, we offer a new definition and propose new classifications for data profiling tasks. Third, the paper presents several free and commercial profiling tools. Fourth, authors offer a new data quality metrics and data quality score calculation. Finally, authors discuss a data profiling tool framework for big data.

Keywords Data profiling tools · Data governance · Big data · Data quality control · Data management

W. Dai(✉)
Information Science, University of Arkansas at Little Rock, Little Rock, AR, USA
e-mail: wxdai@ualr.edu

I. Wardlaw · Y. Li
Computer Science, University of Arkansas at Little Rock, Little Rock, AR, USA
e-mail: {iawardlaw,yxli5}@ualr.edu

Y. Cui
College of Information Engineering, Guangdong Mechanical and Electrical Polytechnic, Guangzhou, Guangdong, China
e-mail: cuiyu@gdmec.edu.cn

K. Mehdi
Software Development Group, Collibra Inc., New York, NY, USA
e-mail: kashif.mehdi@collibra.com

J. Long
Information Science and Engineering, Central South University, Changsha, Hunan, China
e-mail: jlong@csu.edu.cn

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_39

1 Introduction

Data is ubiquitous. People use digital maps to navigate in major cities, send emails to communicate, and buy e-tickets online every day. The more computers assist people, the larger the volume data that will be stored. Big data means a new era in data utilization and exploration; and its key characteristics are volume, variety, and velocity [1].

Data can be structured, as is scientific and statistical data; semi-structured, such as PDF and XML; or unstructured, like raw video or audio [2]. Structured data can be accessed by tools like SQL, while XQuery is usually used to access semi-structured data. Because of the complexity of unstructured data, data profiling tools usually focus on structured data and semi-structured data.

Data quality plays an important role in data governance. According to [3] and [4], data governance takes a unique role in companies because of the Sarbanes-Oxley (SOX) Act and Basel II. Data is a valuable asset for customers, companies, and governments. The target of Total Data Quality Management (TDQM) is to offer high-quality information products (IP) to users, so definition, measurement, analysis, and improvement of data quality are included in TDQM cycle [5] [6].

Data profiling technology can improve data quality. In [7], [8], and [9], data profiling tools discover the pattern of datasets and offer a scores of projects regarding TDQM, including data cleaning, data integration, and data analysis. Data scientists, end-users, and IT engineers utilize these tools to improve data quality because data profiling tools easily show the frequency patterns of addresses information, credit cards, and phone numbers.

The remainder of this paper is organized as follows: Section 2 reviews relevant works and discusses a new definition of data profiling. Section 3 presents new classifications of data profiling task; Sections 4 and 5 introduce several data profiling tools, including both free and commercial software (Fig. 1). Section 4 discovers data quality metrics and an exponentially weighted data quality score in detail. Section 7 discusses a data profiling tool framework for big data. Section 8 concludes main ideas, and section 8 addresses future work.

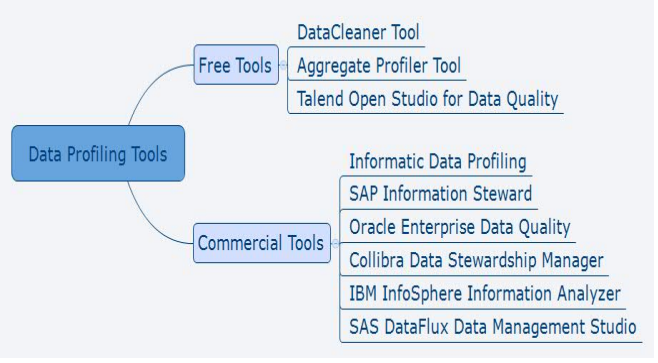


Fig. 1 Selected Data Profiling Tools

2 Data Profiling Definition

Different authors give different definitions of data profiling. [7] describes it as "the set of activities and processes to determine the metadata about a given dataset." According to [10], "data profiling is a process whereby one examines the data available in an existing database or flat file and collects statistics and information about that data. These statistics become the real or true metadata." [11] defines it as "referring to the activity of creating small but informative summaries of a database." These definitions of data profiling seem convincing initially, but [7] only believes that data profiling is the metadata regarding datasets, and [10] does not mention unstructured data at all. Definition of [11] is vague because data has different types, structures, and utilization.

Data profiling can be utilized at different stages of data governance. Thus, in our opinion, profiling is the process of verifying users' structured data, semi-structured data, and unstructured data, gathering data structure, data pattern, statistical information, distribution messages, and reviewing data attributes for data governance, data management, data migration, and data quality control.

3 Data Profiling Tasks

The data profiling includes in multiple tasks for data governance. According to [8] and [10], people need to profile data only when cleaning, managing, analyzing, or integrating it; evaluating its quality; performing data archaeology; optimizing queries; or engaging in Extract, Transform and Load (ETL) projects. However, data profiling could also be used for compressing, verifying, masking, auditing, migrating, archiving, and recovering data as well as for generating testing data and data health reports.[7] and [8] list classifications of data profiling tasks; however, these classifications are not good enough. For example, [8] differentiates data quality jobs based upon whether they use a single data source or multiple sources to separate different data quality jobs. In our view, data profiling tools consists of five primary jobs:

1. Metadata Profiling: discovering metadata information, such as data structures, creators, times of creation, primary keys, and foreign keys.
2. Presentation Profiling: finding data patterns, including text patterns, time patterns, and number patterns. such as address pattern, date patterns, and telephone patterns.
3. Content Profiling: reviewing data basis information, including accuracy, precision, timeliness, null or non-null.
4. Set Profiling: analyzing data from collections or groups; for example statistics, distribution, cardinality, frequency, uniqueness, row count, maximum or minimum values, mean, and redundancy.
5. Logical Rule Profiling: reviewing data based on business logical rules or business glossary domain, such data logical meanings, business rules, and functional dependency.

Different data profiling missions could be used for different projects, such as Table 1.

Table 1 Data profiling for user scenario

Profiling Missions	Profiling Scenario
metadata profiling	Data management, data integration, ETL, data migration
Presentation Profiling	Data compression, data audit
Content Profiling	Data compression, Data management,
Set Profiling	data audit, Data compression, data management
Logical Rule Profiling	data audit, data management

Some sorts of data profiling (such as content profiling and logical-rule profiling) are strongly linked to business domain, but others are not. Fig 2 shows this relationship.

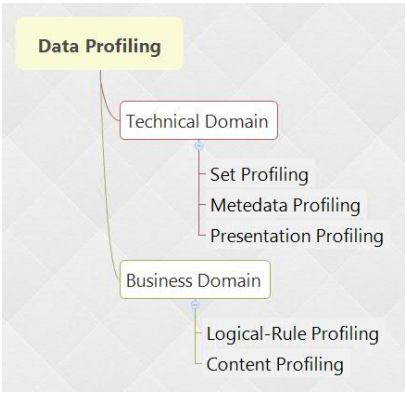


Fig. 2 Data Profiling in different technical or business domains.

4 Free Data Profiling Tools

Some tools are free software and open source; however, many, but not all free data profiling tools are open source projects. In general, their functionality is more limited than that of commercial products, and they may not offer free telephone or online support. Furthermore, their documentation is not always thorough. However, some small companies still use these free tools instead of expensive commercial software, considering the benefits that free tools provide.

1. Aggregate Profiler Tool

Aggregate Profiler (AP) is an open source project developed in Java [7]. AP supports both traditional database and big data, such as Hadoop or Hive, and it offers statistical analysis, pattern matching, distribution chat, basket analysis, etc.

AP also supports data generation, data preparation, data masking features, and address correction for data quality projects. Moreover, this tool offers data validation (metadata profiling, analytical profiling, and structural profiling), and data quality (removing duplicate data, null values, and dirty data).

2. Talend Open Studio for Data Quality

Talend Open Studio for Data Quality (TOSDQ) [8] is also based on Java and is a mature open source tool. TOSDQ offers navigator interface to access databases and data files. This tool supports catalog analysis, time correlation analysis, column analysis, table analysis, column correlation analysis, and schema analysis; it also supports column functional dependency, redundancy analysis, numerical correlation analysis, nominal correlation analysis, connection analysis, column set analysis, and match analysis. Furthermore, TOSDQ reports several different types of statistics indicators, including simple statistics, text statistics, summary statistics, pattern frequency statistics, Soundex frequency statistics, phone number statistics, and Fraud detection (Benford's law frequency).

3. DataCleaner

DataCleaner [9] is a commercial tool for data profiling and data cleaning, but it has a free version which offers multiple data profiling functions, including pattern matching, boolean analysis, weekday distribution, completeness analysis, value matcher, character set distribution, value distribution, date gap analysis, unique key check, date/time analysis, string analysis, number analysis, referential integrity, and reference data matching.

5 Commercial Data Profiling Products

Commercial data profiling products usually come packaged in data governance suites. These products have multiple functions, high performance, and strong capabilities; they can connect to other suites to provide comprehensive solutions for customers. Moreover, these software is not only powerful, but end-users also can find online services and telephone support.

1. IBM InfoSphere Information Analyzer

IBM InfoSphere Information Analyzer (IIA) [10] is part of IBM's data governance suite that includes InfoSphere Blueprint Director, Metadata Workbench, DataStage, QualityStage, Data Click, Business Glossary, and Information Services Director. IIA supports column analysis (statistics, distribution, cardinality, and value analysis.), identifying keys and relationships, discovering redundant data, comparing data and structures through history baselines, analyzing data via data rules, and importing and exporting data rules.

2. Informatic Data Profiling

Informatic Data Profiling is a key component of PowerCenter [11]. This profiling software supports aggregate functions (count null values, calculate averages, get maximum or minimum values, and get lengths of strings), candidate key evaluation (unique or non-unique), distinct value count, domain inference, functional dependency analysis, redundancy evaluation, and row count. In addition, users can add business rules (verbose mode) or configure profile functions in this tool.

3. Oracle Enterprise Data Quality

Oracle Enterprise Data Quality (EDQ) [12] permits address verification, profiling data (files, databases, and spreadsheets), standardization, audit reviews (incorrect values, missing data, inconsistencies, duplicate records, and key quality metrics), matching and merging columns (duplicate prevention, de-duplication, consolidation, and integration), and case management (data reviewing). Furthermore, the tool can utilize pre-built templates or user-defined rules to profile data. EQD also can connect to other Oracle data governance products, including Oracle Data Integrator and Oracle Master Data Management.

4. SAP Information Steward

SAP Information Steward can improve information quality and governance [13] via the Data Insight module (data profiling and data quality monitoring), Metadata Management module (metadata analysis), Metapedia Module (business term taxonomy), and cleansing package builder (cleansing rules). The data insight module can define validation rules, determine profiling (column, address, uniqueness, dependency, and redundancy), import and export metadata, and create views [13].

5. SAS DataFlux Data Management Studio

SAS DataFlux Data Management Studio (DDMS) [14] is a data governance suite that consists data profiling, master data management, and data integration. This data profiling tool covers key analysis (primary and foreign keys), pattern frequency distribution analysis, redundant data analysis, and data profiling reports.

6. Collibra Data Stewardship Manager

Collibra Data Stewardship Manager (DSM) [15] module is part of Collibra's Data Governance Center that also includes Business Semantic Glossary (BSG) and Reference Data Accelerator (RDA) module. DSM also provides historical data quality reports around trend analysis and reports to understand the impact of resolved data issues. In addition, DSM provides fully configurable data quality reporting dashboard (see figure below) by bringing data quality rules and metrics calculated in one or multiple sources (data quality tools, databases, and big data).

6 Data Quality Metrics

Several academic papers offer data quality measurement methods. In [16], the authors mention information quality dimensions including accessibility, completeness, and security; however, these dimensions only focus on information quality for data governance. [17] separates data quality problems according to how a data administrator might view them-for example, single-source, multi-source, instance level, schema level. The authors state that these problems could be solved by ETL, but these procedures only improve the quality of data after it has been collected and stored in a data warehouse making the solution inflexible. [18] presents algorithms for calculating data quality measures, such as free-of-error, completeness, and appropriate-amount-of-data.

Some papers describe how to build metrics for data quality. In [19], the authors offer many metrics (25 candidate metrics and 18 subsets of metrics) regarding data quality. However, they do not mention how to implement them. [20] offers a blueprint for improving data quality, but it only focuses on the business value of such improvement. In [21], the authors discuss how to build the data metrics of data warehouse from such components as table metrics, start metrics, and schema metrics, but they do not mention how to profile data quality, enhance data quality of data structures, or improve unstructured data quality. In [22], the authors mention dimensions of data quality (for instance, uniqueness, accuracy, and consistency), and they also detail how to build a model for quantifying data quality performance.

1. Data Quality Indicators

Qualitative indicators are the best way to measure data quality, but sometime we utilize dimensions to define data quality. In [24], authors enumerate six indicators of data quality (uniqueness, accuracy, consistency, and etc.). Table 2 shows these qualitative indicators and definitions

Table 2 Six Indicators of data quality [24]

Indicator	Definition
Accuracy	The degree to which data correctly describes the 'real world' object or event being described.
Completeness	The proportion of stored data against the potential of '100% complete'
Consistency	Similarity when comparing two or more representations of something against its definition.
Timeliness	The degree to which data represents reality at the required point in time.
Validity	Conformity of data's syntax (format, type, range) to its definition.
Uniqueness	Nothing will be recorded more than once based upon how it is identified.

2. Data Quality Metrics

Qualitative indicators do not offer precise indexes of data quality, but quantitative metrics deliver accurate methods to monitor, control, and improve data quality, which is valuable for data profiling. Metrics consist of directly measured regulations. [24] discusses twenty quantitative metrics of data governance for health organizations; however, these metrics only target at health field. In this paper, we offer comprehensive metrics for data quality, shown in Table 3.

Table 3 Quantitative Data Quality Metrics

Indicator	Metric
Accuracy	Percent of data is correct (correct data / total data) (e.g. ZIP code, SSN)
Completeness	Percent of data is completeness data (e.g. phone number, address)
Consistency	Percent of data is correct consistency (Such as business rules and logical rules of Consistency)
Timeliness	Percent of data is correct timeliness (For example, ages, educational degree at a special time or date.)
Validity	Percent of data is validity (Such as first name, last name, suffix, and etc.)
Uniqueness	Percent of data is uniqueness (e.g. primary keys, foreign keys)

Moreover, we can get the same results if we use simple data instead of the full volume data.

Data Quality Metric (DQM) formula usually need to build data quality indexes, or weights, which depend on specific business scenarios. For example, banks usually send bills via mail or email, so home and email addresses are important to them. However, they may not care much about your cell phone number because they do not need it. Customers could define their own weights for data quality metrics. Table 4 shows an example of metric weights for the address column.

Table 4 METRIC WEIGHTS OF Address Column

Metric	Weight
Accuracy : % of data is correct (W_1)	30
Completeness: % of data is completeness data (W_2)	15
Consistency: % of data is correct consistency (W_3)	10
Timeliness : % of data is correct timeliness (W_4)	05
Validity: % of data is validity (W_5)	20
Uniqueness: % of data is null value or non-null values (W_6)	20
Total Weight Value:	100

DQM formula is

$$DQM(t) = \frac{1}{w_1+w_2+w_2+\cdots+w_n}\sum_{i=1}^{n}\left(Percent_{Correct_{of_{Data_i}}} \times W_i\right) \quad n = 6; \quad (1)$$

$$w_1+w_2 + w_2 + \cdots +w_n = 1\,,\ n = 6\,; \quad (2)$$

3. Exponentially Weighted Data Quality Score

Exponentially Weighted Data Quality Score (EWDQS) is strongly related to time and DQM. If we want to create a data quality score, we should build exponentially weighted moving average formula.

The DQM is a series EWDQS, which may be calculated recursively:

$$EWDQS(1) = DQM(1) \quad (3)$$

$$EWDQS(t) = \lambda * DQM(t) + (1-\lambda) * EWDQS(t-1) \quad t > 1 \quad (4)$$

Where:

1) A higher λ increases previous DQM quicker and 1 >λ>0.
2) DQM (T) is the value of the EWDQS at any time period t.
3) EWDQS (T) is the value at any time period t.

4. Case Study

Imagine Bank A has a lot of address data, and the DBA wants to check its data quality. After profiling the data via data quality tool in January, he notices some problems: 5% of the data is incorrect, 7.8% of data contains null values, and 12.5% of data is inconsistent. The engineer utilize our formulas to measure the data quality score.

According to Equations 1 and 3:

DQM (1) = (1-0.05) × 30 + 15 + 5+ (1-0.078) × 20 + (1-0.12.5) × 10 +20 = 95.69

EWDQS (1) = 95.69

In February, the engineer check it again, it also has some data quality problems: 4.5% of the data is incorrect, 7.0% of data contains null values, and 10.5% of data is inconsistent.

According to Equations 1 and 4, and if λ =0.75:

DQM (2) = (1-0.045) × 30 + 15 + 5+ (1-0.070) × 20 + (1-0.105) × 10 +20 = 96.2

EWDQS (2) = 0.75× DQM (2) + (1-0.75) × EWDQS (1)

= 0.75×96.2 +0.25 × 95.69 = 96.07

Therefore, the exponentially weighted data quality score of Bank A shifts from 95.69 and 96.07 in the month of January.

7 Data Profiling Tool Framework for Big Data

There is limited document mention data profiling tools framework because commercial companies consider this to be a trade secret [19] [20] [21] [22] [23] [24]. However, as big data increases, data governance, data quality control, and data profiling technology grow more important [1] [25] [26]. People need data profiling tools to perform the big data analysis necessary to improve data quality, but they are currently challenged by limited performance and unexpected robustness. Customers want these tool to profile real-time data, static data profiling, simple data, and full volume data, and we should strive to satisfy their requirements. The data profiling framework has six layers: the hardware layer, data layer, parallelism layer, algorithms layer, function layer, and Web UI layer (Fig. 3).

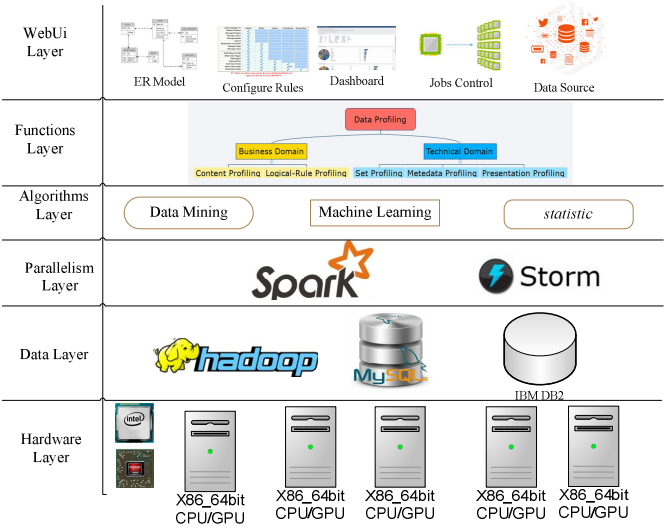


Fig. 3 Data Profiling Architecture

Each layer has different application (Table 4). “At the Algorithm Layer, for instance, machine learning algorithms could be employed that automatically analyze the patterns and rules of data as well as data structures in order to enhance data quality and facilitate its governance. These algorithms could run on x86 machines or graphics processing units (GPU), such as NVidia CUDA [27].

Table 5 Data profiling framework

Layer	Application
Web-UI Layer	User interface for ER model, business rules, KPI dashboard, batch or real-time job maintenance, input/output data source configuration.
Function Layer	Scores of data profiling missions (Section 3).
Algorithm Layer	Utilize data mining, machine learning, and statistics, or other algorithms.
Parallelism Layer	Apache Spark [28] for static data and full volume data; Apache Storm [29] for real-time data.
Data Layer	Business rules, configuration data, metadata store in Hadoop, traditional databases.
Hardware Layer	Integrate CPU and GPU clusters for improving performance, especially real-time or machine learning tasks.

8 Conclusion

As data profiling enriches data quality, data management, and data governance, it is important for customers, data scientists, and DBA to use data profiling tools. Data is an asset for all users, so data quality should be controlled by procedures, rules, people, and software. In this paper, after reviewing existing relevant works, a new data profiling definition and data profiling tasks were presented. Authors discuss several free or paid data profiling tools. Moreover, the paper introduced a method for building data quality metrics and showed how to calculate data quality scores. Finally, a data profiling tool framework was given.

9 Future Work

Data profiling only grows more important in this new era of big data. We will continually read more academic papers, technical documents, and develop codes about algorithms for data profiling on big data and will find new ways to extend the functions of data profiling in the future.

References

1. Zikopoulos, P., Eaton, C.: Understanding big data: analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media (2011)
2. Buneman, P.: Semistructured data. In: Proceedings of the Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, pp. 117–121. ACM (1997)
3. Buneman, P., Davidson, S., Fernandez, M., Suciu, D.: Adding structure to unstructured data. In: Database Theory, ICDT 1997, pp. 336–350. Springer, Heidelberg (1997)
4. Khatri, V.: Brown, C.V: Designing data governance. Communications of the ACM **53**(1), 148–152 (2010)
5. Pipino, L.L., Lee, Y.W., Wang, R.Y.: Data quality assessment. Communications of the ACM **45**(4), 211–218 (2002)

6. Wang, R.Y.: A product perspective on total data quality management. Communications of the ACM **41**(2), 58–65 (1998)
7. Kumar, R., Yadav, A.: Aggregate Profiler – Data Quality. http://sourceforge.net/projects/dataquality/
8. Talend Company. Talend Open Studio for Data Quality. http://www.talend.com/products/data-quality
9. DataCleaner Company. DataCleaner Manual. http://datacleaner.org/resources/docs/4.0.10/pdf/datacleaner-reference.pdf
10. IBM Company. InfoSphere Information Server: Information Center. http://www-01.ibm.com/support/knowledgecenter/SSZJPZ_9.1.0/
11. Informatica Company. Data Profiling Solutions. https://www.informatica.com/data-profiling.html
12. Oracle Company. Oracle Enterprise Data Quality. http://www.oracle.com/us/products/middleware/data-integration/enterprise-data-quality/overview/index.html
13. SAP Company. SAP Information Steward. http://scn.sap.com/docs/DOC-8751
14. SAS Company. SAS Products: DataFlux Data Management Studio. http://support.sas.com/software/products/dfdmstudioserver/
15. A Data Governance Solution Tailored for Your Role. Collibra Solution Comments. https://www.collibra.com/solution/
16. Pipino, L.L., Lee, Y.W., Wang, R.Y.: Data quality assessment. Communications of the ACM **45**(4), 211–218 (2002)
17. Rahm, E., Do, H.H.: Data cleaning: Problems and current approaches. IEEE Data Eng. Bull **23**(4), 3–13 (2000)
18. Lee, Y.W., Pipino, L.L., Funk, J.D., Wang, R.Y.: Journey to data quality. The MIT Press (2009)
19. Moody, D.L.: Metrics for evaluating the quality of entity relationship models. In: Conceptual Modeling, ER 1998, pp. 211–225. Springer, Heidelberg (1998)
20. Ballou, D.P., Tayi, G.K.: Enhancing data quality in data warehouse environments. Communications of the ACM **42**(1), 73–78 (1999)
21. Calero, C., Piattini, M., Pascual, C., Serrano, M.A.: Towards Data Warehouse Quality Metrics. In: DMDW, p. 2 (2001)
22. Loshin, D.: Monitoring Data Quality Performance Using Data Quality Metrics: A White Paper. Informatica, November 2006
23. The Six Primary Dimensions for Data Quality Assessment. The Six Primary Dimensions for Data Quality Assessment. http://www.enterprisemanagement360.com/white_paper/six-primary-dimensions-data-quality-assessment/
24. The Ultimate Guide to Data Governance Metrics: Healthcare Edition:40 Ways for Payers and Providers to Measure Information Quality Success. The Ultimate Guide to Data Governance Metrics: Healthcare Edition (2012). http://www.ajilitee.com/wp-content/uploads/2013/09/Ultimate-Guide-to-Data-Governance-Metrics-for-Healthcare-Ajilitee-June-2012.pdf
25. Zikopoulos, P., Eaton, C.: Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media (2011)
26. LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Big data, analytics and the path from insights to value. MIT Sloan Management Review **21** (2013)
27. CUDA GPUs. NVIDIA Developer. June 4, 2012. https://developer.nvidia.com/cuda-gpus
28. Apache Spark™ - Lightning-Fast Cluster Computing. Apache Spark™ - Lightning-Fast Cluster Computing. http://spark.apache.org/
29. Apache Storm. http://storm.apache.org/

Understanding the Service Model of Mobile Social Network in Live 4G Network: A Case Study of WeChat Moments

Weijian Sun, Dandan Miao, Xiaowei Qin and Guo Wei

Abstract Driven by the rapid advancement of both cellular network technologies and mobile devices, mobile social network (MSN) services significantly facilitate personal and business communications and therefore inevitably consume substantial network resources and potentially affect the network stability. Despite of their huge user base and popularity, little work has been done to characterize the service model of MSN services. Thus, in light of many works on MSN services focusing on their various different aspects, this paper aims to provide a deep understanding of the service model of MSN services from different levels. In order to reach credible conclusions, our research is based on a large-scale data set collected inside a province wide 4G cellular data network in China. Specially, we choose WeChat Moments (WM) as a case study since it is one of the most popular MSN services in China with 600 millions of active users per month.

Keywords 4G Cellular network · Big data · Mobile social network · Service model

1 Introduction

With the rapid development of cellular network technologies, powerful mobile devices and various requirements of social contacts, the number of mobile social network (MSN) consumers has been growing continuously. What is worth to mention is that the emerging technologies and researches of big mobile traffic data analysis clearly yield enormous potentials when it comes to building and understanding large-scale service model of MSN services in live cellular network. However,

W. Sun(✉) · D. Miao · X. Qin · G. Wei
Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei, China
e-mail: {sunwj,ddmiao}@mail.ustc.edu.cn,
e-mail: {qinxw,wei}@ustc.edu.cn

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_40

despite of the popularity [1],[2],[3],[4],[5],[6], few studies focus on its service model. Yang et al. [7] mainly focus on the social behavior of MSN which can improve our understanding of the nature of MSN services, but few information can be provided to network operators. In this paper, we take a first look at WM traffic on a large-scale commercial 4G network. The reason we choose WeChat Moments (WM) as a case study of is that it is one of the most popular MSN services in China with 600 millions of active users per month. Our goal is to understand the service model of WM from individual and aggregated level. To the best of our knowledge, our study is the first to investigate the characteristics of traffic generated by WM in live cellular network. We summarize our key contributions as follows.

- **Large-scale measurement:** We conduct the first large-scale measurement study of cellular WM traffic. Our dataset collected from the **S1-MME**, **S1-U**, **S11** and **S6a** interfaces, covering about 10 thousands eNodeBs in the northwest province of China within a region of 0.45 million km^2.
- **Accurate recognition of WM traffic and session:** We recognize the traffic of WM by the $(host, url, useragent)$ pair of HTTP get/post request. 293 actions of WeChat can be recognized and 9 of them are corresponded with WM through the structure of url. One WM session is defined as continuous HTTP get/post requests that belong to one usage of one user in each time. These 9 WM actions are integrated into 4 class and each class can be treated as the beginning of each WM session since each of them are repetitively tested and verified at terminal-side by our OTTCAP [4].
- **Individual model of WM:** From the perspective of individual traffic, the bimodal nature of WM session inter-arrival time is confirmed. Two kinds of distributions of session length are obtained and we find it is determined by the social contents that users visit in each session. Furthermore, we compare the user activeness with corresponding model's power exponent and find the monotonic relation between them. We also compare the session length generated by different terminal types and summarize distinction of different terminals.
- **Aggregate model of WM:** From the perspective of aggregated traffic, temporal and spatial dynamic patterns of WM traffic, which represent users' preference, are analyzed with the help of cells' GPS information. We also examine the fitting results of aggregated traffic within one cell to candidate distributions and find that power-law models are also suitable to characterize the aggregated traffic of WM. Furthermore, the model's parameters can be fitted to Gauss distribution which means the WM model can be formulized in cell level.

The remainder of the paper is organized as follows. In Section II, we describe our methods and dataset. In Section III, we provide detailed results. Finally, we conclude this paper with a summary and future research direction in Section IV.

2 Background

2.1 Data Set

In this paper, we use one anonymous data set from a tier-1 4G cellular network carrier in China. User-plane and control-plane data, which are collected from **S1-MME**, **S1-U**, **S11**, **S6a**, are combined together to generate the service records of all users. The data contains the following information for each user's HTTP get/post request(record): *cell id*, *IMEI*, *begin time*, *end time*, *packet length*, *host*, *url*, *useragent* and so on. All user identifiers are anonymized to protect privacy without affecting the usefulness of our analysis. In general, our data set contains about 1.6 million 4G users and their over 8 billion HTTP requests during two weeks of April in 2015 in one northwest province (14 cities) of China. Furthermore, all cells' GPS information are provied which can be utilized to show the spatial distribution of traffic. The volume of our data is 6 Terabyte (1TB = 1024GB) which means some big data analyzed tools should be utilized and this will be presented in next subsections.

2.2 Recognition of WM Traffic and Session

In order to separate WM from other serivces, we first test the url of all WeChat actions such as open a dialog, open a picture or visit a friend's page and so on in our OTTCAP, and analyze the corresponding url information. We find the consistency of WeChat urls, for example,**weixin.qq.com/cgi-bin/micromsg-bin/XXXX** is the basic pattern of WeChat url. We name **XXXX** action, and when action begins with **mmsns**(perhaps means micro message social networking services), the corresponding url belongs to WM. We find 273 actions of WeChat and 9 of them begin with

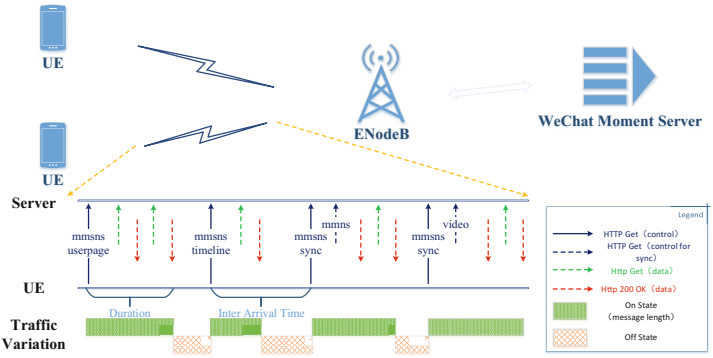


Fig. 1 An illustration of four kinds of WeChat Moments activities.

mmsns. Futhermore, we also need to distinguish the beginning of one WM session. After the large amounts of test of WM in our OTTCAP, we identify 4 combinations of these 9 actions which are shown in Fig.1. **mmsnsuserage** and **mmsnstimeline** are intuitive ones, however, when detect **mmsnssync**, other following actions such as the **mmsns** and **video** should be combined together to distinguish the beginning of one WM session. The definition of session length and session inter-arrival time, which are the key parameters of service model, are illustrated in Fig.1.

2.3 Analysis Procedure

Since the network operators not only care about user behavior of WM but also concern the aggregated behavior in the cell level, we analyze service model of WM from two levels - individual level and aggregated level. Raw records, which are integrated from **S1-MME**, **S1-U**, **S11**, **S6a**, are stored in our HDFS. Map-reduce jobs are written to parse such big volume of traffic data. *Service Parese Job* is used to transmit the records, which belong to WM, into desired format shown in Table. 1. There are 16 items in the records of the desired format, for example, item 2 represents the total elements that user have read in one session. *Individual traffic job* and *Aggregate traffic job* are utilized to generate corresponded data set which will be separately analyzed through our matlab scripts. The analysis results will be illustrated in Section III.

Table 1 Session format of WM after *Service Parese Job*

index	item	index	item	index	item	index	item
1	user_id	5	clicked videos	9	session duration	13	upload packet number
2	total element number	6	comment number	10	session active time	14	download packet number
3	not clicked pictures	7	session begin time	11	upload traffic	15	user terminal type
4	clicked pictures	8	session end time	12	download traffic	16	cell_id

3 Service Model of WM from Two Levels

3.1 Individual Model of WM Session

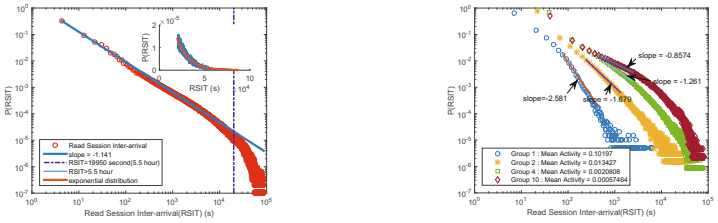
Distribution of Individual Session Inter-Arrival Time. We first analyze the overall behavior of individual session inter-arrival time. Analysis on burstiness and memory [8] is helpful to quantitatively understand WM's individual model. Burstiness of a user is defined as $B = \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau}$, where m_τ and σ_τ are the mean value and standard deviation of the series of inter-arrival time, respectively. With the definition, B is a real value in the bounded range from -1 to 1, where $B > 0$ means that the series has burstiness effects.

Memory is used to measure the inter-arrival time correlations, i.e., whether a short (long) inter-arrival time is followed by a short (long) one. The memory coefficient of a user is defined as $M = \frac{1}{n_\tau - 1} \sum_{i=1}^{n_\tau - 1} \frac{(\tau_i - m_1)(\tau_{i+1} - m_2)}{\sigma_1 \sigma_2}$, where n_τ is the total number of the inter-arrival time, m_1 (m_2) and σ_1 (σ_2) are sample mean and sample standard deviation of the $\tau_i{}'s$ and $\tau_{i+1}{}'s$. According to the definition, M ranges from -1 to 1, and obviously $M \approx 0$ denotes the short (long) inter-arrival time inclines to be followed by a long (short) one. The mean values of burstiness and memory are $B = 0.5351$ and $M = 0.0604$. That is to say, the WM system has bursty effects but almost no memory effects, which is different from many well-studied human-activated systems [8]. In Fig.2(a), we report the inter-arrival time distribution based on the aggregated data of all users, specifically, 9497926 inter-arrival time of 481538 WM users. The distribution is shown below and t_s denotes Read Session Inter-arrival(RSIT).

$$f(t_s) = \begin{cases} 1.77 \cdot t_s{}^{-1.141} & t_s \leq 5.5hour \\ \frac{1}{6.66 \times 10^5} \cdot \exp(-\frac{1}{7.749 \times 10^5} \cdot t_s) & t_s > 5.5hour \end{cases} \quad (1)$$

As depicted in Fig.2(a), we find the bimodal distribution of session inter-arrival time since when $t_s > 5.5hour$, the distribution can be fitted well with exponential and when $t_s \leq 5.5hour$, it can be fitted with power-law. The power-law exponent is obtained by maximum likelihood estimation with $RMSE = 7.916 \times 10^{-5}$. It is worth mentioning that unless further noted, all the fitted distributions are obtained by this method in this paper and RMSE is exploited as the uniform criterion for gauging the fitting preciseness. As analyzed in [9], the bimodal behavior can be explained by the interaction among individuals since the difference between WM and other social media is that WM has message alert mechanism.

In order to understand the inter-arrival time distribution of our data deeply, we want to explain this distribution by identical user behavior, namely, user activeness. We define user activeness as $A_i = \frac{n_i}{T_i}$, where A_i means the frequency of sessions of an individual. n_i is the total sessions of user i in one day and T_i is the time between the first and the last session. To investigate the role of user activeness, we sort the user by activeness in descending order and divide them into 10 groups. Each groups are



(a) (b)

Fig. 2 (a) Power-law distribution of session inter-arrival time. (b) Power exponent in each group of user activeness.

with equal user number. We report 4 typical power-law distribution of inter-arrival time at a group level, and from Fig.4(b), we find that they have different exponent values. Note that the group with lower activeness has smaller power-law exponent, giving a longer average session inter-arrival time. Finally, we diagram the exponent as a function of user activeness and find a non-trivial, monotonous increase of the exponent with the activeness. This relation indicates the significant role of activeness for the power-law behavior of session inter-arrival time.

Distribution of Individual Session Length. When it comes to distribution of individual session length, there are two basic questions to be answered. First, which distribution is the fittest one of session length and second, just as the role of user activeness to power-law exponent of session inter-arrival time, which parameter has the decisive influence on the distribution of session length. We characterize the statistical feature of WM session first. We analyze 1660498 sessions generated by 182354 users and find that users do not prefer to click pictures when reading WM since only 29% of all sessions have pictures clicked. We judge if a picture is clicked or not by deep parsing the *url* of each picture. Users just like to read WM but do not prefer to comment since only 8% of sessions have user comment inside them. Nine kinds of terminals are extracted from the *user-agent* of WM and the statistical feature of each terminals are shown in Fig.3(c) and Fig.3(d). In Fig.3(c), we compare each terminal group from two perspective, namely, number of terminals and average session length. From number of terminals, we find that most of users that using WM are iphone users which partly reflects that WM users are of high-income group. From average session length, we find that lenovo users tend to read more contents in each session since the average traffic is 190KB, while that of Coolpad users are 120KB. Generally, the size of one picture is 10KB, that is to say, lenovo users tend to read 7 more pictures than coolpad users. We also compare cumulative distribution functions (CDFs)of the session length generated from four terminals, namely, coolpad, huawei, iphone and lenovo in Fig.3(d) and find the same result. Since 20% of WM sessions generated by lenovo terminals are larger than 350KB and when it comes to others such as coolpad terminals, only 10% of their sessions are larger than 350KB.

Next, we calculate the probability density functions (PDF) of session length. At first, we do not concern the content type such as video, picture or comment, in each session and then fit them to common heavytailed distributions, namely, power-law, weibull and lognormal. As shown in Fig.3(a), there is a pdf peak around session length eaquals to 454KB. In order to deepen understand this characteristic, we separately analyze the session length distribution of different kinds of contents since it is intuitively seeing that the session with length equals to 454KB may not just contain pictures. Therefore, we divide the WM sessions into two categories, namely, with video and without video, and plot their corresponding pdf in Fig.3(b). We are please to find that the results are consistent with our conjecture. On the one hand, when a session does not contain video content, its distribution can be well fitted by power-law with exponent equals to -1.756. To see clearly, we exhibit the power-law result in the subfigure on the top of Fig.3(b). On the other hand, when a session contains video contents,

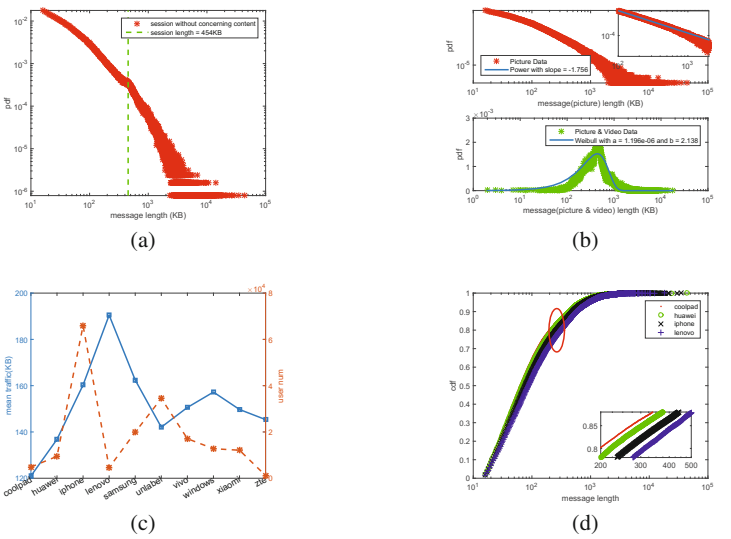


Fig. 3 (a) Session distribution without concerning session concents. (b) Two kinds of distribution of session length. (c) Terminal types of WM users and its corresponded average session length. (d) CDF of session length generated by different kinds of terminals.

which is the new function in the newest version of WeChat, the pdf of session length can be well fitted with weibull distribution with $k = 2.138$ and $\lambda = 1.196 \times 10^6$. Finally, the overall distribution of individual WM session length is

$$f(l_s) = \begin{cases} 28.89 \cdot {l_s}^{-1.756} \& flag = 0 \\ \frac{k}{\lambda} \cdot (\frac{l_s}{\lambda})^{k-1} \cdot \exp(-\frac{l_s}{\lambda})^k \& flag = 1 \end{cases} \tag{2}$$

where $flag = 0$ means there are no video contents in those sessions and vice versa.

3.2 *Aggregated Model of WM Session*

Temporal and Spatial Dynamic of WM Traffic. The temporal feature of WM traffic in 2015-04-03(workday) of one city is depicted in Fig.4(a). The overall traffic trend is in consistent with human schedules since traffic volume during the day is higher than that of mid-night. We find there are 3 local peaks of WM traffic in one day, namely, 11 am, 16 and 23 pm and respectively represents before lunch, before off duty and before sleep. There are also 3 local lowest points, namely, 4 am, 12 am and 21 pm and respectively represents sleep hour, lunch hour and family hour. This temporal features of WM means that it is for leisure and entertainment.

The spatial feature of WM traffic is shown in Fig.4(b). The green star in Fig.4(b) represents cdf of cell and red star represents eNodeB. The two dashed lines mean that whether it is cell or eNodeB, 80% of total traffic is generated from the total 35% cells or eNodeBs. The non uniform nature of spatial traffic provides potentials for the planning of CDN servers of WM. We select a area to visualize its spatial distribution in different hour of day. The longitude of this area is from 103.8 to 103.9 and latitude is from 36.03 to 36.09. We divide this area into 10000 $0.01 * 0.01$ squares and calculate each square's WM traffic by cumulating those cells's traffic that belong to this square. The visualized results of 4 selected hours in the selected area are shown in Fig.4(d). In general, the temporal pattern of this area is the same to the overall temporal analysis, for example, there is almost no traffic generated in 4 am of this area. Furthermore, we also try to analyze the WM's hotspot around [103.835,36.04] since, as shown in Fig.4(c) and Fig.4(d), no matter in which hour or the whole day, it is always the hottest area of WM. We correspond this location and its places of interest(POI) data in Google Map, and find that a second biggest hospital of the analyzed city together with two residential subdistricts are located around there.

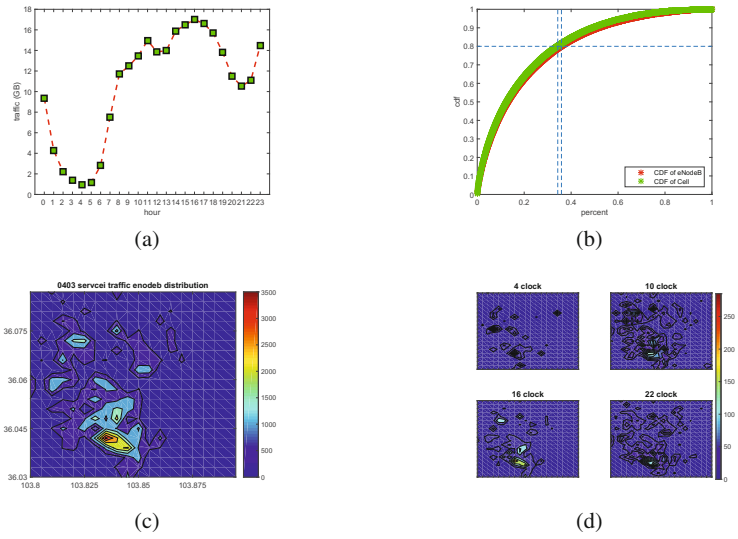
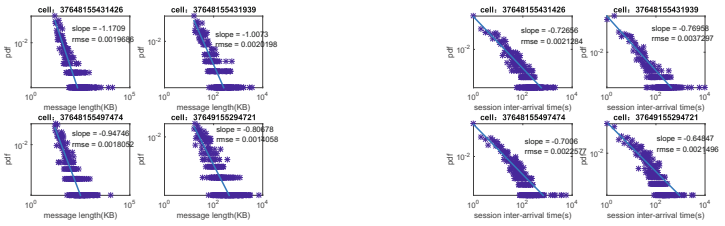


Fig. 4 (a) Temporal distribution of WM traffic in a city. (b) Unbalance of traffic among cells and eNodeBs. (c) Spatial distribution of WM traffic in a city. (d) Spatial distribution of WM traffic in four distinct hours

Distribution of Aggregated Traffic. We try to answer the following two questions, the first is: Are the distribution of session length and session inter-arrival time at cell level same to those of individual ones and the second: Can the cell level's models be formulized. Based on the *cell ID* of each session record, we gather each cell's

WM session. After sorting all the cells by its session number from big to small, we choose the top 400 ones from the whole 4574 cells since we should make sure that each cell has statistical significance. Fig.5(a) and Fig.5(b) verify the fitting preciseness of session length and session inter-arrival time to power-law models in four randomly selected cells. Furthermore, the fitting RMSE and parameters of pawer-law distribution of each cell are also fitted, parameters are shown in Table.2. We achieve the rather good results that each parameters of cell's WM model can be fitted to Gauss distribution, in other words, we can formulize cell level's WM model.



(a) (b)

Fig. 5 (a) Fitting examples of four random selected cells (session length).(b) Fitting examples of four random selected cells (session inter-arrival time).

Table 2 Gauss parameters of aggregated session inter-arrival time fitted RMSE and Power-law paratmeters

Gauss Para	a	b	c
RMSE	0.31	0.0023	0.00053
Power-law ratio	0.28	-0.67	0.077
Power-law slope	0.32	0.074	0.023

4 Conclusion

In this paper, we investigated the traffic characteristics of WM service from two different viewpoints. For individual traffic, we form the model of session length and session inter-arrival time. Bimodal feature can be found in the distribution of session inter-arrival while power-law and weibull can be fitted to the distribution of session length based on the different contents in one session. We further mine the relationship between user activeness and the power-law exponent and find the non-trivial, monotonous increase of the exponent with user activeness. Terminal types are also concerned towards average session length and distribution of session length. For aggregated traffic, we first analyze the temporal and spatial dynamics of WM traffic and deep understand this uniform feature through a selected area. When it comes to the model of WM from cell level, we are surprised to perceive that both session length and session inter-arrival time can be fitted to power-law distribution

and the corresponded parameters can be fitted to Gauss distribution which means that the WM model of cell level can be formulized. Our future research will focus on the implication of WM model in wireless resource allocation mechanism.

Acknowledgements This research is supported by the National Basic Research Program of China (973 Program: 2013CB329004) and National High Technology Research and Development Program of China (863 Program: 2015AA01A706).

References

1. Shafiq, M.Z. Ji, L., Liu, A.X., Pang, J., Wang, J.: Characterizing geospatial dynamics of application usage in a 3G cellular data network. In: 2012 Proceedings IEEE INFOCOM, pp. 1341–1349. IEEE (2012)
2. Sun, W., Qin, X., Wei, G.: Modeling and mining the temporal patterns of service in cellular network. Communications, China **12**(9), 11–21 (2015)
3. Shafiq, M.Z., Ji, L., Liu, A.X., Pang, J., Wang, J.: Large-scale measurement and characterization of cellular machine-to-machine traffic. IEEE/ACM Transactions on Networking **21**(6), 1960–1973 (2013)
4. Sun, W., Qin, X.: End-to-end delay analysis of wechat video call service in live dc-hspa+ network. In: 2014 Sixth International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–5. IEEE (2014)
5. Zhou, X., Zhao, Z., Li, R., Zhou, Y., Palicot, J., Zhang, H.: Understanding the nature of social mobile instant messaging in cellular networks. IEEE Communications Letters **18**(3), 389–392 (2014)
6. Li, R., Zhao, Z., Qi, C., Zhou, X., Zhou, Y., Zhang, H.: Understanding the traffic nature of mobile instantaneous messaging in cellular networks: A revisiting to $alpha$-stable models. IEEE Access **3**, 1416–1422 (2015)
7. Yang, K., Cheng, X., Hu, L., Zhang, J.: Mobile social networks: state-of-the-art and a new vision. International Journal of Communication Systems **25**(10), 1245–1259 (2012)
8. Zhi-Dan, Z., Hu, X., Ming-Sheng, S., Tao, Z.: Empirical analysis on the human dynamics of a large-scale short message communication system. Chinese Physics Letters **28**(6), 068901 (2011)
9. Wu, Y., Zhou, C., Xiao, J., Kurths, J., Schellnhuber, H.J.: Evidence for a bimodal distribution in human communication. Proceedings of the National Academy of Sciences **107**(44), 18803–18808 (2010)

Applying Scrum in an Interdisciplinary Project for Fraud Detection in Credit Card Transactions

Mayara Valeria Morais dos Santos, Paulo Diego Barbosa da Silva, Andre Gomes Lamas Otero, Ramiro Tadeu Wisnieski, Gildarcio Sousa Goncalves, Rene Esteves Maria, Luiz Alberto Vieira Dias and Adilson Marques da Cunha

Abstract This paper aims to describe the use of Scrum agile method for an academic prototype system development using Big Data, Internet of Things, and Mobile Devices, as a solution for detecting and reducing fraud on credit card transactions. It highlights the involvement of graduate students in the Brazilian Aeronautics Institute of Technology (Instituto Tecnologico de Aeronautica - ITA), during the first Semester of 2015. The major contribution is a Proof of Concept (PoC) development, by using an Interdisciplinary Problem Based Learning (IPBL), where students geographically dispersed, from three different courses, worked into a real problem and had to develop a solution within seventeen weeks. At the end of the project, it was possible to deliver an academic prototype system, where mobile devices were able to carry out transactions with performance, using Internet of Things (IoT), and enabling data mass analyses within a Big Data environment in a transparent way.

Keywords Scrum agile method · Fraud detection · Interdisciplinary problem based learning

1 Introduction

According to Gartner [1], in 2015, 4.9 billion things are interconnected like smartphones, computers, cars, and other devices. An increase on connections

M.V.M. dos Santos(✉) · P.D.B. da Silva · A.G.L. Otero · R.T. Wisnieski · G.S. Goncalves · R.E. Maria · L.A.V. Dias · A.M. da Cunha
Computer Science Division, Brazilian Aeronautics Institute of Technology – ITA, Sao Jose Dos Campos, Brazil
e-mail: mayara.vms@gmail.com, paulodiego1@gmail.com, lamas3000@gmail.com, ramirotadeu@gmail.com, gildarciosousa@gmail.com, rene.stvs@gmail.com, {vdias,cunha}@ita.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_41

of 30% since 2014 will reach 25 billion by 2020. Craig Spiezle [2] said that the implementation of security practices must coincide with the speed at which the Internet of Things (IoT) trend is growing.

Once the mobile devices usage has grown gradually, banking services were impacted by people's use. According to a study by EasySolutions [3] between May and June 2015 in Latin America, 52% of respondents use banking services through mobile devices and in Brazil by 97%.

In a report released by Criteo [4] on e-commerce's global panorama, in the first quarter of 2015, 34% of transactions were made through smartphones and tablets. By the end of the year, this figure will reach 40%. In Japan and South Korea, the number of transactions via mobile devices is over 50% [4]. This information appointed to the gradual growth of IoT, although benefited from its convenience and agility, also brings high-impact problems.

The prime problem is linked to fraudulent transactions. In a study issued by ClearSale [5], until the end of 2015, one in six online fraud attempts in Brazil will be through mobile devices.

Comparing all online fraud attempts, 18% will originate on a mobile device, an increase over 2014, when the rate was 7%.

During the first Semester of 2015, within seventeen weeks, it was developed an academic prototype system to prevent and reduce online fraud transactions, named as BDIC-DM (Big Data, Internet das Coisas e Dispositivos Moveis, in Portuguese). In this project, it was possible to perform a Proof of Concept (PoC), using the Interdisciplinary Problem Based Learning (IPBL) approach taken from around fifty graduate students from three different courses at the Brazilian Aeronautics Institute of Technology (Instituto Tecnologico de Aeronautica - ITA).

2 Background

2.1 Big Data

Big Data are large pools of data structured and/or unstructured able to be captured, stored, organized, and analyzed through tools to generate results. According to Marr [6], Big Data consists of volume (amount of data), velocity (speed in data analysis), variety (variety in data types), truthfulness (data consistency), and value (result of data analyses).

2.2 Internet of Things

The conventional concept of the Internet as an infrastructure network reaching out to end-users terminals will fade, leaving space to a notion of interconnected 'smart' objects forming a pervasive computing environment [7]. The next wave in the era of computing will be outside the realm of the traditional desktop. In the IoT paradigm, many of the objects that surround us will be on the network in one

form or another [8]. The basic idea of this concept is the pervasive presence of a variety of things or objects – such as Radio- Frequency IDentification (RFID) tags, sensors, actuators, mobile phones, among others – which through unique addressing schemes, are able to interact with each other and cooperate with their neighbors to reach common goals [9].

2.3 Mobile Devices

Mobile devices are portable computing appliances such as smartphones and tablets. They have been used as personal computers, often replacing them. The term 'mobility' refers to the higher degree of independence from space and time achieved in information and communications technology processes by the employment of mobile devices [10].

2.4 Agile Method Scrum and Its Practices

Scrum is a framework that uses processes and techniques to address and solve complex adaptive problems in an iteratively and incrementally way. Under the pillars of transparency, inspection, and adaptation, Scrum is one of the most flexible and adaptable agile methods. Scrum involves the following roles, ceremonies, and artifacts: (i) roles: Product Owner, Scrum Master, and Team Development; (ii) ceremonies: Sprint Planning Meeting, Daily Meeting, Sprint Review, and Sprint Retrospective; and (iii) artifacts: Product Backlog, Sprint Backlog, and Burndown Chart [11].

3 The Project Overview

The proposed IPBL for the three courses (CE-240 - Database Systems Project; CE-245 - Information Technology, and CE-229 - Software Testing) taught at the graduate program in Electronics and Computer Engineering at the ITA, the BDIC-DM project goal was to develop an academic prototype system for prevention of fraudulent transactions by credit cards via mobile devices, using the Scrum agile method.

To perform real-time fraud detection, via e-commerce, demands immediately answers (approval or denial), to protect online users. This first requirement demanded high performance and high availability architecture. Moreover, the initial project architecture was designed involving two segments: OLAP (Online Analytical Processing) and OLTP (Online Transaction Processing).

The architecture was based on REST services, which can easily be integrated with various platforms in open source tools (e.g. Apache Cassandra, Apache ActiveMQ, Node.js, R, Hadoop, Hive, MySQL, and Linux) lowering the cost of ownership. It allows customizations, besides being tested thoroughly by the community, and supported by companies such as IBM, RedHat, and Microsoft.

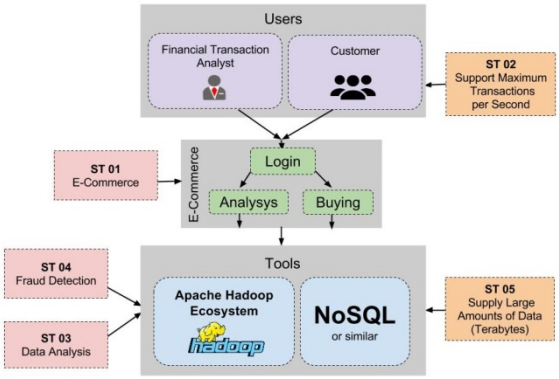


Fig. 1 The project initial architecture.

Three R language models for fraud prevention were developed, in order to improve the user protection. They contributed to the system reliability and coverage, because they checked various parameters at transactions' time.

From the OLAP and OLTP segments were defined five Scrum Teams (STs) responsible for five macro-features: (i) ST01 - e-commerce; (ii) ST 02 - supporting maximum transactions per second; (iii) ST 03 - batch data analysis; (iv) ST 04 - online fraud detection; and (v) ST 05 - supplying large amounts of data (Terabytes).

Each ST had around eleven students from three different courses, featuring interdisciplinarity. Furthermore, a Product Owner (PO), a Scrum Master (SM), and the other students as Team Members (TMs) comprised each ST.

For managing the STs, it was adopted the Scrum of Scrums (SoS) technique. It consisted in managing a large number of people divided into STs. For academic purposes were used two new roles: General Product Owner (GPO) and General Scrum Master (GSM). These roles would have the function of managing all project teams, in order to remove impediments and helping during its activities and ceremonies.

At the end of the BDIC-DM project, it was supplied: (i) three R models for fraud prevention; (ii) Application Programming Interface (API) of fraud prevention engine, where clients connect to transact; (iii) transaction validator application used as the authentication method in two steps for users to transact more securely; (iv) e-commerce to test the integration with the API; (v) control panel, where the transaction history can be analyzed, in order to refine the analysis models; and (vi) API stress test suite.

4 The Project Development

4.1 Sprint 0

The Sprint 0 objective was empowering students from three different courses for new technologies, methods, and artifacts that would be used during the project

development. In the first few weeks, it was performed a dynamic exercise known as Lego4Scrum [12], in order to introduce Scrum concepts, and agile development's benefits and challenges. Furthermore, the professors have provided supporting materials such as books, online courses [13], exercise lists and lectures, which led students to understand the main project concepts of Big Data, IoT, and credit card fraud.

Also in the Sprint 0, students have proposed some increases in project architecture. They choose the NoSQL database Cassandra [14], to handle millions of operations per second, and be friendlier to developers accustomed to relational database. It was also defined a Rest API [15] to enable the link between e-commerce, database, and mobile devices.

4.2 Sprint 1

At the beginning of the Sprint 1, through a video conference, occurred the first part of the Sprint Planning. It is an important Scrum ceremony for planning activities and tasks during the Sprint. The five STs choose the User Stories (USs) responsible for first integration level of the BDIC-DM project. To prioritize the business value, the STs used the Planning Poker technique for estimating the effort of each US. In the second part of the Sprint Planning, the STs created tasks for each US and managed tasks development through the Kanban method [16], using a free online tool named Trello. The tasks had three status (i.e. to do, doing, and done). This allowed the real-time control and management by each ST and its members geographically dispersed.

All the STs had few common characteristics: working remotely, dispersedly, collaboratively, and/or cooperatively. In addition, through the BDIC-DM project portal, hosted on Google Sites, students have been used spreadsheets as project artifacts (i.e. Product Backlogs, Sprint Backlogs, and Burndown Charts) along with all content developed by the ST as it relates to the project documentation (i.e. tutorials, videos, and test cases). With the purpose to storage source-codes, scripts, data definition language, and data manipulation language, it was created an account on Github to the BDIC-DM project (https://github.com/projbdicdm/bdicdm). The Github use has allowed to centralize the project development and provided access to the code for all involved ST members.

To control all artifacts and the project progress, the five STs made online weekly meetings. The ST reported each task's status being developed and planned strategies to optimize deliverables.

In order to centralize all project features, a portal began to be designed, which the main functions were to sell items for the e-commerce's user, and to provide information analysis for the Financial Transaction Analyst (FTA).

In parallel, the establishment of the data model was initiated. It was chosen the MySQL database to support structured data made from adding and/or excluding items to the shopping cart, and the unstructured database Cassandra to support the

transaction data. These unstructured data would be consumed by the Hadoop ecosystem in the next Sprint.

Another Sprint objective was the beginning of the Rest API creation using the Node.js framework [17]. The developed API was based on JavaScript framework called Node.js. It enables to make calls to the database through a search query, via Java Database Connectivity (JDBC). This search query has accessed Hadoop through the Hive and returning data to the API, which were consumed by the view layer.

The creation of login in the Rest API was another Sprint objective, in order to enable the buying method with its confirmation functions. All of these functions should support the data flow coming from the e-commerce. The main contribution of this objective was to send an email to the costumer with the purchase confirmation code.

One of the first Sprint major gains was the creation of the test architecture. The architecture followed the agile testing quadrants and used the Test Driven Development (TDD) supported by testing tools (i.e. JUnit, Selenium, and JMeter).

During the Sprint Review, each ST was able to show the main results obtained from its USs. Sixteen USs composed the Sprint Backlog. Thirteen USs were delivered and accepted by POs, the other USs do not meet the minimal quality requirements described at the Sprint Planning Meeting.

Students also performed the Sprint Retrospective. They described the positive and negative points during the project development and made recommendations for the next Sprint. The main recommendations were: (i) face meetings between ST members; (ii) effective participation of all project members; and (iii) endorse the ST integrations.

4.3 Sprint 2

During the second Sprint Planning, it was selected and discussed the Product Backlog USs for the second Sprint Backlog by the five STs. Each selected US will aggregate to the project new features and integrations. Each ST began creating test plans based on the selected test tools. As a result, test cases were created for each US before the development starts. In addition, the TDD was used to improve quality assurance.

In this Sprint, three different profiles were created to access the portal: (i) the administrator; (ii) the customer; and (iii) the Financial Transaction Analyst (FTA). Within the customer profile, it was possible to visualize items in the catalog, and add them directly to the shopping cart. However, the customer was not able to finalize purchase orders, because the integration with the developed API would be taken on the next Sprint. Although not integrated with the API buying method, it was possible to view charts and generate reports on the FTA profile, using mock data provided by the ST 05. On the next Sprint, it will be consumed data coming from the API. In addition, the API code was refactored to provide information about customer types.

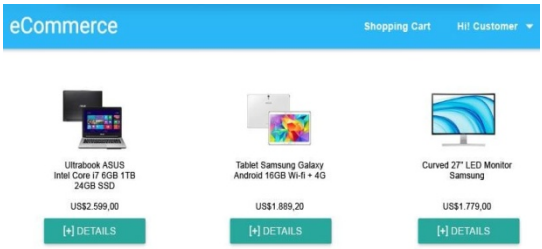


Fig. 2 The e-commerce's catalog items.

It was necessary the installation and configuration of a tool for data transfer between MySQL and Hadoop, in order to allow the flow between data from the relational database for Big Data environment. The Apache Sqoop [18] was chosen mainly because of its robustness and flexibility. The Sqoop allows creating automated scripts that make the synchronization between the data available in MySQL to Hive. The available dataset was successfully imported and scripts were tested. If new data were included in MySQL, synchronization with Hadoop will be automatic.

Some searched models for fraud prevention, in order to identify that a minimal learning curve were found, according to [19] [20] [21] [22] [23]. The adopted model for fraud prevention on the BDIC-DM project was the Hidden Markov Model (HMM) [24]. To use the HMM, it was chosen the R HMM library. The R integration with Cassandra worked using RJDBC library and JDBC drivers provided by Cassandra. In addition, it was used BoxPlot to detect points out of the spending curve. Fraud detection used the following chosen items as parameters: (i) the credit limit; (ii) the geographical distance between transactions; and (iii) the customer buying behavior. The code used these fixed parameters, remaining to the next Sprint for construction of the e-commerce parameterization.

Three tools were created to identify the stress level in the NoSQL database, and match their data generated to the results generated by the implemented model for fraud detection: (i) the Benchmark Tool, which made a metric of transaction number versus request time; (ii) the Data Shuffle, responsible for shuffle the latitude and longitude data generated by transactions; and (iii) the Fraud Report, responsible for showing a percentage of filter transactions that will be analyzed by the FTA.

During the second Sprint Review, each ST presented the developed values and integrations. Twenty-one USs composed the Sprint Backlog. Seventeen USs had been delivered and accepted by POs, the other USs do not meet the minimal quality requirements described at the Sprint Planning Meeting.

During the second Sprint Retrospective the main recommendations made were: (i) an online server was necessary to centralize the deployment to improve STs integration; (ii) to keep the test plans creation at the Sprint starts; and (iii) the communication problem persists between STs.

4.4 Sprint 3

During the third Sprint Planning, the STs aimed to integrate all features developed, and to create an application for mobile devices. This application would serve to confirm the purchase by the customer through a code, because the email service, held at the first Sprint, took too long to finalize purchase orders.

The Sprint challenge was the integration of full features developed following an assigned mission created by the professors. The assigned mission consisted of the following steps: (i) in the customer profile add items directly to the shopping cart; (ii) the purchase confirmation, via mobile device; (iii) provide a fraudulent transaction to verify the implemented model for fraud detection; (iv) insertion of concurrent transactions; and (v) analyzes a data mass in the Big Data environment in a transparent way to the FTA.

One of the Sprint gains was a server availability. It has allowed a better performance of the developed features, since it needed an environment with a more robust hardware than personal computers of project members. Moreover, it was possible to process and analyze, online and batch, in less time.

The developed FTA control panel was part of the last integration level, within the portal. This panel would serve to improve fraud prevention for continuous transactions monitoring. All this information in .pdf format was exported to reports.

The Apache JMeter [25] was used for load testing. It brought a gain reuse of the previous tests, serving also as an API's keeper, since tests could break if API had changes. In addition, to the JMeter, Selenium was also used as an automated Graphical User Interface (GUI) test tool in the e- commerce. These tools usage has implied in time testing optimization and improved USs quality.

In order to perform purchase authentication by the customer via mobile device, it was created an application on the Android platform supported by messages service Apache ActiveMQ [26] (https://goo.gl/Q7N47a). After the purchase made in the e-commerce, the customer received a purchase code notification in your mobile device. For the purchase confirmation, the customer should enter the received code. It was possible to perform fraud detection making twenty transactions per second, using an only one server.

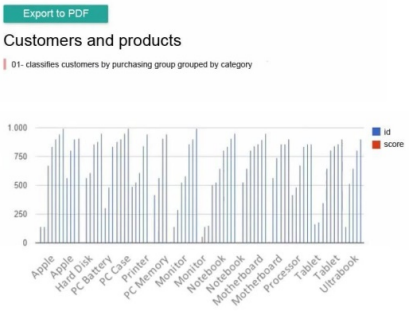


Fig. 3 The FTA control panel.

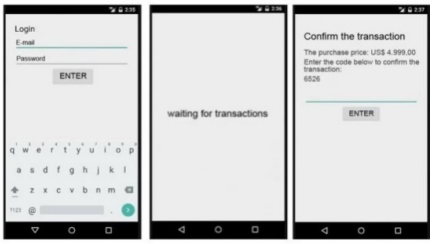


Fig. 4. Android Application to authorize transactions.

It was consolidated the final project architecture, as shown in Fig. 5. In this architecture, it was possible to identify the flow of REST API, which was the link between the e-commerce, structure/unstructured database, an mobile devices.

During the third Sprint Review, the STs presented the final integration work. Fourteen USs composed the Sprint Backlog. All of them were delivered and accepted by POs. At the last Sprint Retrospective, students reported that the communication problem was finally mitigated, because the project members were able to improve their maturity in this project development, understanding the agile method.

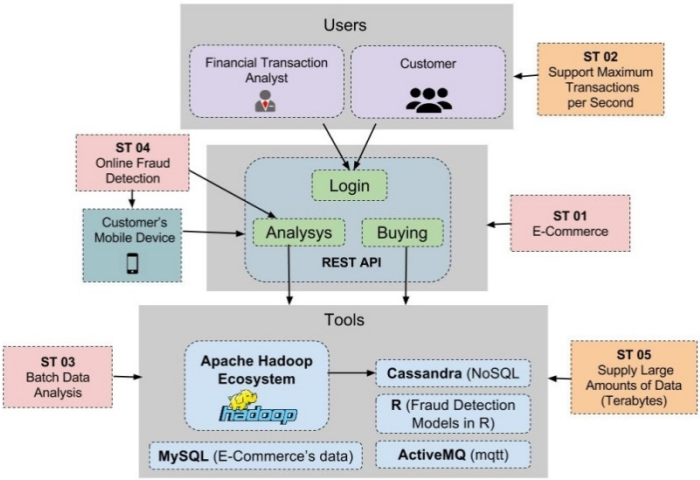


Fig. 5 The project final architecture.

5 Conclusion

This paper aimed to describe the academic prototype development using the agile method Scrum and the IPBL approach, in order to develop a fraud prevention system based on a PoC. It was possible to deliver a portal where mobile devices using IoT with performance in transactions. It also enabled the data mass analysis within a Big Data environment in a transparent way to the FTA.

The result was obtained from a set of open source tools, and the agile method support. Moreover, the main contributions made during the project developed were synthesized in the final architecture.

Arising from students' collaboration, critical, and experimentation, this academic prototype proved to be able to accomplish the proposed assigned mission. Consequently, it has provided for students a real life experience about how high-performance and high availability systems are developed, and how to work with geographically dispersed STs.

Acknowledgements The authors would like to thank the supporting of Brazilian Aeronautics Institute of Technology (ITA) and the involved graduate students from the graduate program in Electronics and Computer Engineering for making possible the BDIC-DM project development.

References

1. Gartner: Gartner Says 4.9 billion connected "things" will be in use in 2015. http://www.gartner.com/newsroom/id/2905717
2. CRN: IoT boom needs to be matched by security – report. http://www.channelweb.co.uk/crn-uk/news/2421745/iot-boom-needs-to-be-matched-by-security-report
3. EasySolutions: Visão dos consumidores latino-americanos sobre a fraude eletrônica – 2015. http://www.easysol.net/pt/research-report-por
4. Criteo: State of mobile commerce Q2 2015. http://www.criteo.com/br/resources/mobile-commerce-q2-2015
5. ClearSale: Tentativas de fraude via aparelhos móveis crescem; saiba se proteger. http://portal.clearsale.com.br/novidades/Tentativas-de-fraude-via-aparelhos-moveis-crescem-saiba-se-proteger
6. Marr, B.: Big Data: The 5 Vs everyone must know. https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know
7. Weiser, M.: The computer for the 21st century. Scientific American, 94–100 (1991)
8. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 1645–1660 (2013)
9. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Computer Networks, 2787–2805 (2010)
10. Fenu, G., Pau, P. L.: An analysis of features and tendencies in mobile banking apps. The 10th International Conference on Future Networks and Communications (FNC 2015) / The 12th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2015), vol. 56, pp. 26–33. Procedia Computer Science, Belfort (2015)
11. Rico, D.F., Sayani, H.H., Sone, S.: The business value of agile software methods, p. 70 (2009)
12. Lego4Scrum. https://lego4Scrum.com
13. Big Data University. https://bigdatauniversity.com
14. Cassandra. https://cassandra.apache.org
15. InfoQ: A Brief Introduction to REST. http://www.infoq.com/articles/rest-introduction
16. Atlassian: A brief introduction to Kanban. https://www.atlassian.com/agile/kanban
17. Node.js. https://nodejs.org
18. Apache Sqoop. https://sqoop.apache.org
19. Dheepa, V., Dhanapal, R.: Analysis of credit card fraud detection methods. International Journal of Recent Trends in Engineering **2**(3), 126–128 (2009)
20. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. ACM Computing Surveys (CSUR) **41**(3), 15:1–15:58 (2009)
21. Dheepa, V., Dhanapal, R.: Behavior based credit card fraud detection using support vector. Ictact Journal on Soft Computing **02**(4), 391–397 (2012)

22. Bahnsen, C., Stojanovic, A., Aouada, D., Ottersten, B.: Improving credit card fraud detection with calibrated probabilities. In: 2014 SIAM International Conference on Data Mining, Pennsylvania, pp. 677–685 (2014)
23. Pozzolo, A., Caelen, O., Waterschoot, S., Borgne, Y., Bontempi, G.: Learned lessons in credit card fraud detection from a practitioner perspective. Expert Systems with Applications **41**(10), 4915–4928 (2014)
24. Bhusari, V., Patil, S.: Study of hidden markov model in credit Card fraud detection. International Journal of Computer Applications **20**(5), 0975–8887 (2011)
25. Apache JMeter. https://jmeter.apache.org
26. ActiveMQ. https://activemq.apache.org

The Visual Representation of Numerical Solution for a Non-stationary Deformation in a Solid Body

Zhanar Akhmetova, Seilkhan Boranbayev and Serik Zhuzbayev

Abstract One of the most interesting directions in the field of solid body deformation mechanics, having an applied nature, is the solution of half-plane non-stationary deformation problems, on a rectangular base of which the time-varying loads are superimposed. Such problem arises, for example, in the study of well environ stress-strain state, its design and development of piles hammering technology. The main idea of this paper is to show the use of numerical methods in order to visualize the results of solving the tasks about the non-stationary half-plane deformation.

Keywords Investigatory data visualization · Visual representation · Isolines of normal stress · Bicharacteristics method · Half-plane deformation

1 Introduction

One of the best methods for analyzing physical processes and data appears to be the scientific visualization, which is usually based on the construction of physical and mathematical models.

If couple of decades before the visual representation of numerical solutions was highly dependent not only on the available possibilities of computer hardware resources but also on requirements of mathematical modeling and use of numerical methods to the level of visualization, nowadays the great number of software tools and supercomputers allows us to establish and solve even more complex problems.

The investigatory visualization is destined for the analysis and processing of a data set, for instance, to detect patterns in them [15,16,17].

Nowadays digital oscilloscopes are widely used in order to obtain the experimental information.

Z. Akhmetova(✉) · S. Boranbayev · S. Zhuzbayev
Department of Information Systems, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
e-mail: {zaigura,juzbayev}@mail.ru, sboranba@yandex.kz

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_42

In this paper we solve the problem about visualization of half-plane non-stationary deformation. The development of many areas of engineering and construction is associated with the use of highly relevant sections of elasticity. This theory is widely applied in the stress-strain state calculation of the machine parts, composite structures elements in construction, engineering, aviation and space technology.

One of the most interesting engineering problems, from a practical point of view, is the problem of half-plane non-stationary deformation when on the base, which has a rectangular cut, a time-varying load is superimposed. Suchlike problem appears, for example, in the study of well environ stress-strain state, its design and development of hammering piles technology. As studies have shown, the representation of soil as a continuous linearly deformable body allows the use the methods for solving dynamic problems, such as, the construction on the basis of bicharacteristics method involving the idea of splitting the spatial coordinates of explicit difference scheme, their application to the analysis of soil compaction under the piles, the establishment of the area of opening and progress of soil in the well vicinity [18,19,20,21].

Let's consider this method in details.

2 The Bicharacteristics Method

Currently, for solving dynamic problems in elastic medium there widely used the numerical methods of spatial characteristics [10], finite element [1], the boundary integral equation [2] and etc.

One of the most convenient methods in applications is bicharacteristics method with the use of splitting method ideas [5,6,7,8,9]. This method allows approaching dependence domain of the final and differential equation to the dependence area of the initial differential equation as much as possible.

In this paper we reviewed the non-stationary problem solution of the homogeneous isotropic elastic body dynamics in the Cartesian coordinate system using the bicharacteristics method.

Let us introduce the following notations:

x_i – Cartesian coordinates,

t - time,

σ_{ij} – tension tensor,

ϑ_i– velocity vector,

u_i – displacement vector.

Indexes after a comma designate partial derivatives according to the time. For convenience, we used Latin and Greek letters to index, which take on the value 1,2. According to repeated Greek indices the summation is fulfilled in the product, while in repeated Latin indices there is no summation.

Let us consider an elastic semi-strip of final width which in the Cartesian system of coordinates x_1Ox_2 occupies the area of $0 < x_2 < \infty$, $|x_1 < l|$. In the initial time point the body is at state of rest.

$$\vartheta_i = 0, \sigma_{ij} = 0, \textit{(i,j=1,2)} \tag{1}$$

At any other time on the site $N_1 \le x_2 \le N_2$, $x_1 = l$ of border BN the uniformly distributed transient normal load *f(t)* has its influence, which varies according to sine law

$$\sigma_{22}(t) = \begin{cases} -A sin(\omega t)\ O \le t \le S_1 \\ 0, \quad t \ge S_1 \end{cases} \tag{2}$$

$$\sigma_{21}(t) = 0.$$

Where S_1– loadings action time and $\omega = \pi / S_1$.
The other part of the semi-strip border is free from any influence:

$$\sigma_{11}(t) = 0, \quad \sigma_{12}(t) = 0, \quad x_1 = 0, \quad |x_1| \ge l, \sigma_{22}(t) = 0, \quad \sigma_{21}(t) = 0,$$

$$0 \le x_1 \notin (N_1,\ N_2)\ |x_2| = l \tag{3}$$

Under existing conditions it is necessary to investigate an elastic body tension at *t>0.*

2.1 The Defining Equations

In order to solve the problem along with entry and boundary conditions, we used the system of the equations consisting of the movement and ratios equations of the generalized Hooke's law:

$$\sigma_{i\beta,\beta} = \rho \frac{\partial^2 u_i}{\partial t^2}, \tag{4}$$

$$\sigma_{ij} = \lambda u_{\beta,\beta} \delta_{ij} + \mu \left(u_{i,j} + u_{j,i} \right). \tag{5}$$

Where ρ – density,
λ, μ – Lama's constants,
δ_{ij} – Kronecker delta.
On the repeating Greek indexes the summation is carried out.
For convenience, independent immense variables and required sizes are entered [8].

$$\bar{t} = \frac{t c_1}{b}, \quad \overline{x_i} = \frac{x_i}{b}, \quad \overline{v_i} = \frac{1}{c_1} \frac{\partial u_i}{\partial t}, \quad \overline{\sigma_{ij}} = \frac{\sigma_{ij}}{\rho c_1^2}$$

$$\gamma_{12} = \frac{c_2}{c_1}, \quad \gamma_{11} = 1 - 2\gamma_{21}^2 \ \ (i, j = 1, 2)$$

Where b- reference length,
$c_1 = \sqrt{\frac{\lambda+2\mu}{\rho}}$, $c_2 = \sqrt{\frac{\mu}{\rho}}$ – Datum speeds.
Further the line over non-dimensional parameters falls.

After the non-dimensional variables integration, the motion equations (4) and correlations of the generalized Hooke's law differentiated by time (5) take the following form:

$$\begin{aligned}&\left\{\dot{\vartheta}_1 = \sigma_{11,1} + \sigma_{12,\,2},\ \dot{\vartheta}_1 = \sigma_{21,1} + \sigma_{22,\,2}, \dot{\sigma}_{11} = \vartheta_{1,1} + \gamma_{11}\sigma_{2,\,2}, \dot{\sigma}_{22} = \right.\\ &\left.\gamma_{11}\vartheta_{1,1} + \sigma_{2,\,2}, \dot{\sigma}_{12} = \gamma_{12}^2(\vartheta_{1,2} + \vartheta_{2,1})\right.\end{aligned} \tag{6}$$

Indexes after the comma denote the partial derivatives on Cartesian coordinates and the point from the top shows the partial derivatives on time.

2.2 The Equations of Bicharacteristics

In order to obtain bicharacteristics equation and conditions on them, let us split the two-dimensional system (6) on the single-dimensional one. Applying ideas of K.A.Bagrinovski and S.K.Godunov on splitting multidimensional t- hyperbolic systems on single-dimensional systems where $x_k = const$ [9], we will have:

$$\left\{\begin{array}{c}\dot{\upsilon}_i - \sigma_{ij,j} = a_{ij}\\ \dot{\sigma_{ij}} - \lambda_{ij}^2\upsilon_{i,j} = b_{ij}\end{array}\right\} \tag{7}$$

Where

$$a_{ij} = \sigma_{ik,k};; \qquad \lambda_{ij} = \sigma_{ij} + \gamma_{12}\left(1 - \delta_{ij}\right);;$$

$$b_{ij} = (\gamma_{11}\delta_{ij} + \gamma_{12}^2\left(1 - \delta_{ij}\right))\vartheta_{p,k}$$

Hence and in future $i, j, k, p = 1, 2;;$ $p \neq i,$ $k \neq j.$

From here, using notorious methods to obtain differential bicharacteristics equations and conditions on them, we obtain:

$$dx_j = \pm\lambda_{ij}dt,$$

$$d\sigma_{ij} \mp \lambda_{ij}d\upsilon_i = \left(b_{ij} \mp \lambda_{ij}a_{ij}\right)dt \tag{8}$$

2.3 Selection of a Point Scheme and a Pattern

This body is divided into square cells, sides of which are $\triangle x_1 = \triangle x_2 = h$. In the double points, the function values ϑ_i, σ_{ij} are searched at various time points with step of τ. The dot grid (on the basis of which the difference scheme is built, other than those mentioned double points) contains points formed by the intersection of bicharacteristics with hyperplanes $t = const$. Hence, the pattern is accepted, consisting of $\hat{I}$ node and $E_{ij}^{\pm}$ points, separated from the point O to the distance $\lambda_{ij}\tau$. In the future values of the functions at the point O is attributed to the upper mark "0" in the points $E_{ij}^{\pm}$ - with the subindex «ij» and the upper sign «$\pm$» respectively (for example, $\sigma_{ij}^{\pm}$,), and in point A the additional index is not attributed [11-13].

2.4 Resolving Differential Equations

The integration of equations (6) from the point O to the point A and the relations (8) from the point $E_{ij}^{\pm}$ to the point A by trapezoid method allows us to obtain the expression of the following form.

$$\vartheta_i = \vartheta_i^0 + \frac{\tau}{2}(\sigma_{ij,j} + a_{ij} + \dot{\vartheta_i^0}), \tag{9}$$

$$\sigma_{ij} = \dot{\sigma_{ij}^0} + \frac{\tau}{2}(\lambda_{ij}^2 \upsilon_{i,j} + b_{ij} + \dot{\sigma_{ij}^0}),$$

$$\sigma_{ij} - \sigma_{ij}^{\pm} \mp \lambda_{ij}\left(\upsilon_i + \upsilon_i^{\pm}\right) = \frac{\tau}{2}(b_{ij} + b_{ij}^{\pm} \mp \lambda_{ij}\left[a_{ij} + a_{ij}^{\pm}\right]) \tag{10}$$

Excluding from (9) ϑ_i, σ_{ij} by means of (10), we will receive:

$$\lambda_{ij}^2 \upsilon_{i,j} \mp \lambda_{ij}\sigma_{ij,j} = b_{ij}^{\pm} - \sigma_{ij}^0 \pm \lambda_{ij}(\upsilon_i^0 - a_{ij}^{\pm} + \frac{2}{\tau}(\sigma_{ij}^{\pm} - \sigma_{ij}^0 \pm \lambda_{ij}\left[\upsilon_i^0 - \upsilon_l^{\pm}\right]) \tag{11}$$

Values of unknown quantities in non-nodal points of expression (12) are calculated on Taylor's formula near a double point of 0 with accuracy to the second order concerning a step τ, therefore, we will have:

$$\lambda_{ij}^2 \upsilon_{i,j} \mp \lambda_{ij}\sigma_{ij,j} = \lambda_{ij}^2\left(\upsilon_{i,j}^0 + \tau\dot{\upsilon_{i,j}^0}\right) \mp \lambda_{ij}\left(\sigma_{ij,j}^0 + \tau\dot{\sigma_{ij,j}^0}\right) \tag{12}$$

Summing up and subtracting each system equation (13) with identical indexes pair, we will receive

$$\upsilon_{i,j} = \upsilon_{i,j}^0 + \tau\left(\sigma_{ij,jj}^0 + a_{ij.j}^0\right)$$

$$\sigma_{ij,j} = \sigma^0_{ij,j} + \tau(\lambda^2_{ij}\upsilon^0_{i,jj} + b^0_{ij,j}) \tag{13}$$

Procedure of receiving the equations allowing systems in double points of the studied body in time moment $t = t_n + \tau$ is various for internal, boundary and angular points of the studied area.

We should recognize that the given problem in general case should be investigated as part of an axially symmetric deformation. However, the numerous experimental results confirm that during the process of well compaction, any geometric shape of the well base (rectangle, triangle, circle, etc.) affects only the numerical results. Qualitatively, the deformation process is practically repeated almost in all cases. With this in mind, and having developed an algorithm for solving the location of planar linear dynamic problems of deformable bodies, the process of soil compaction was investigated in this paper [14].

3 Statement of Problem

The pile clogging is normally realized by different technological methods. One of them is the establishment of the pile and affecting its free end by the shock loads, driving it deep into the ground. Another way is preliminary prepared wells done by percussion instrument which has mass M. This percussion tool is dropped from a height H, followed by the installation of piles. Finally, the third technique involves the successive application of the first two methods. First, a well is created, a pile is set, and then it is hammered to a certain depth. It is particularly important to know the compaction degree of the soil under the pile. Really, in this case it defines the parameters (stress level, the nature of their effects, etc.) of the subsequent technological operations and influences the choice of the tool material.

In this paper the problem related to soil compaction made by percussion instrument is solved.

Mathematically, the problem is formulated as follows. At the initial moment of time $t_0 = 0$ the soil occupying the half-plane with a cut (see figure 3) is at the resting state, so

$$v_\alpha = 0, \qquad \sigma_{\alpha j} = 0 \qquad (\alpha, j = 1, 2) \tag{14}$$

The percussion tool action is simulated by normal load perceived by the base of the well ($x_1 = H, \quad |x_2| \le 1$) and described by the law:

$$\sigma_{11} = 1.7, \quad if \quad 0 \le t \le 450\tau, and 0, \quad if\ t > 450\tau; ; \sigma_{12} = 0 \tag{15}$$

Both the side surface of the well and the free surface of the soil are not loaded. In the described conditions it is necessary to study the stress-strain state of soil and also to set the area and degree of its compaction.

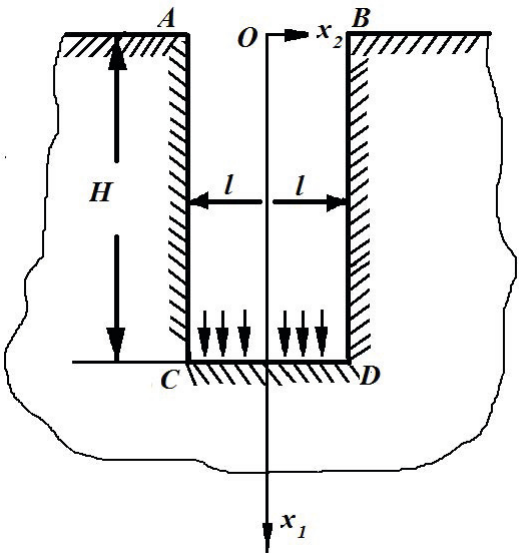


Fig. 1 The study area

Table 1 Experimental data on the compaction of the well bottom hole by projectile

No	The length of the wall, *m*	The diameter of the wall, *m*	The density of the soil, *t/m3*	Poisson's ratio	The mass of the projectile , *t.*
1	0,8	0,6	1,65	0,3	5,5
2	1,2	0,06	1,65	0,3	7,0

4 Final Remarks

The developed finite-difference scheme was performed with the use of the created software at the following initial data:

$$\tau = \triangle t = 0,00035\ c;; \qquad h = \triangle x_1 = \triangle x_2 = 0,03\ sm.$$

The physical and mechanical properties of the soil are:
Elastic modulus is $\AA$=21700Mp;
Poisson's ratio is v=0,3;
Density is $\rho = 1,65\ t/m^3$.

These parameters remain constant in all versions of calculation and the mass of the projectile M, wall diameter d, length l and height of fall H varied. (Table I).

At Figure 4, the curves corresponding to the lines of maximum stress on the time interval of σ_{11} $0 \leq t \leq 5400\tau$, distributed near the leading edge of the boundary wave are shown. At the initial moments of time $0 \leq t \leq 450\tau$ compressive stress drops

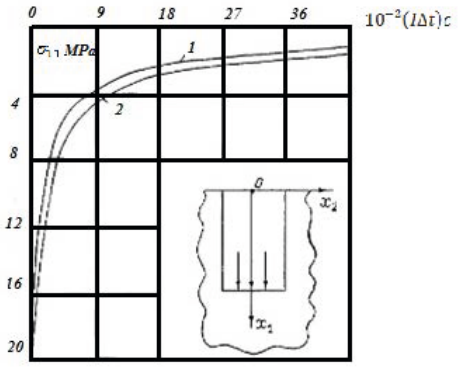


Fig. 2 The lines of maximum stress

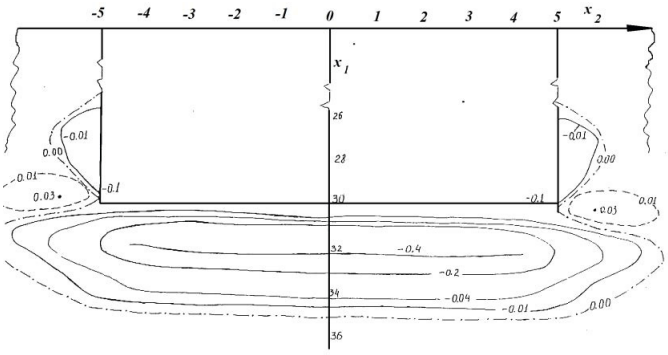


Fig. 3 The normal stress isolines of σ_{11} at the moment of time t = 900τ

sharply immediately after removal of the external load, followed by gradual reduce up to zero during the estimated dynamic field near the base of the well.

The analysis of the dynamic field near the well base shows the great practical interest. Such analysis is normally conducted not only to identify areas of stress concentration, but also to establish the level of soil compaction and its rainfall. Although the results, which are obtained for an elastic medium, have the predictive behavior, we can assume that they might be essential in the other situations. In this regard, in the plane x_1 O x_2 for the moment of time t = 900τ, Figure 5 depicts that the isolines of normal stress σ_{11} in Mp reveal the expected for the task in hand state of stress throughout the disturbed area at the given moment of time.

The calculation results reflect that the maximum compressive stresses are formed near the bottom of the hole. Thus, it can be expected that soil settlement occur in the vicinity of the well bottom base. The extensive area of tensile stress, with some space on the side at the level of the well bottom, represents the room predisposed to destructions.

The results of the issue, which is solved in this paper, have been tested in the designing of the future construction.

5 Conclusions

During the construction in difficult engineering-geological conditions the advantageous structural designs are considered to be the stuffed piles in the punched wells being in the mode of dynamic loading.

Therefore, there is a great need of the «pile-soil» stress-strain state's deep analysis []. Such problems are at immediate interest at the time of pile material durability investigation, in calculation of the bearing ability of soil and at determination of tension concentration in the vicinity of contact surfaces. A reliable estimation of the wells base stress-strain state has a great importance within the development of the design schemes calculating the load transfer made on the soil by piles.

In the case of well bottom shell consolidation by striking tool, the special studies are conducted including tests of the soil stress-strain state and soil behavior in the different modes of impact, and determination of the surface pressure on the contact area of the projectile rammer and soil. The stress nature in the base can be studied by solving the problem of the elastic waves' propagation in the soil.

One of the most complex issues of considering the construction works impact is the study of the interaction between structures and soil.

A theoretical study of this problem is based on the mathematical modeling of real physical and mechanical properties of the "soil – structure" system.

The method, applied to visualize normal stresses appearing from the interaction of structures with the soil, is based on the solution of the inverse problem. This means that we need to find the defining parameters under which there will be a non-stationary deformation of the half-plane. During the repeated calculation of inverse problems the defining parameters range to the prior event (physical effect). The resulting interdependence of the defining parameters is represented visually in the form of isolines to which we can apply the geometric transformations. Considering the space of defining parameters as the isolines set, the inverse problem is formulated as a finding of all sub-areas in the half-plane where the event has occurred.

The explosive energy is widely used during some processes of a consolidation of the constructions bases, tunneling underground workings, construction of embankments, dikes, dams, etc. The effects of explosives on soil, as a special dynamic influence, need to be investigated properly to insure a maximum effect.

In the practice of blasting there widely used not only separate charges but also their highly complex systems representing a great technical interest. For example, a base influence and "soil-pile" interaction are widely investigated, as well as the analysis of strain appearing inside the piles when they are hammered into the soil. The increasing volume of industrial, mining, hydraulic engineering and aviation engineering spheres depicts the need to improve the research methods of the well environ stress-strain state, its design and development of piles hammering technology.

As a result of this work with the utilization of a virtual storage oscilloscope in package VisSim there has been performed the visualization of the problem solution on half-plane non-stationary deformation, that is, the problem associated with soil compaction by the percussion tool.

The graphs of the normal stresses isolines, obtained during the study, allow to identify both the area of precipitation occurrence and soil zones having predisposition to destruction.

References

1. Aytaliyev, Sh.M., Masanov, Zh.K., Baymakhanov, I.B., Makhmetova, N.M.: Numerical methods of the solution of problems of mechanics of a deformable solid body, Karaganda, pp. 3–15 (1987)
2. Aytaliyev, Sh.M., Alekseeva, L.A., Dildabayev, Sh.A., Zhanbyrbayev, N.B.: Method of the boundary integrated equations in problems of dynamics elastic many connected ph. Alma-Ata, p. 238 (1992)
3. Clifton R.D.: Mechanics, Mockow, No. 1, pp. 103–122 (1968)
4. Reker, V.V.: Applied mechanics. Series E, No. 1, pp. 121–129 (1970)
5. Tarabrin, G.T.: Stroit. Mechanics and calculation of constructions, No. 4, pp. 38–43 (1981)
6. Boranbayev S., Altayev S., Boranbayev, A.: Applying the method of diverse redundancy in cloud based systems for increasing reliability. In: Proceedings of the 12th International Conference on Information Technology: New Generations, ITNG 2015, Las Vegas, Nevada, USA, April 13-15, 2015, pp. 796–799 (2015)
7. Boranbayev, S., Boranbayev, A., Altayev, S., Nurbekov, A.: Mathematical model for optimal designing of reliable information systems. In: Proceedings of the 8th IEEE International Conference on Application of Information and Communication Technologies, AICT 2014, Astana, Kazakhstan, October 15-17, 2014, pp. 123–127 (2014)
8. Boranbayev, A.S., Boranbayev, S.N.: Development and Optimization of Information Systems for Health Insurance Billing. In: Proceedings of the 7th International Conference on Information Technology: New Generations, ITNG 2010, Las Vegas, Nevada, USA, April 12-14, 2010, pp. 1282–1284 (2010)
9. Tarabrin G.T.: Metallurgical science, No. 3, pp. 193–199 (1979)
10. Tarabrin, G.T.: Construction mechanics and calculation of constructions, Mockow, No. 6, pp. 53–58 (1980)
11. Karimbayev, T.D., Dzhuzbayev, S.S.: Durability of materials and elements of designs with sound and ultrasonic frequencies of tension. In: Tez. Dokl. Mezhd. Konf., Kiev, p. 18 (1992)
12. Bayteliyev, B.T., Dzhuzbayev S.S.: Tez. Dokl. Vsesoyuzn. Simpoziuma on a rheology of soil, Volgograd, pp. 37–38 (1985)
13. Kim J.-Y.: On the generalized self-consistent model for elastic wave propagation in composite materials. Int. J. Solids Struct., 4349–4360 (2004)

Part IV
Software Engineering

Developing Usability Heuristics for Grid Computing Applications: Lessons Learned

Daniela Quiñones, Cristian Rusu, Silvana Roncagliolo, Virginica Rusu and César A. Collazos

Abstract Several methods allow measuring the degree of usability of a software product. The heuristic evaluation is one of the most popular methods which allow finding more usability problems, in comparison to other methods. However, sets of generic usability heuristics may not evaluate (usability) aspects related to specific software. A set of usability heuristics for Grid Computing was developed using the methodology proposed by Rusu and others, in 2011. The methodology facilitated the heuristics' design and specification, with stages that support the developing process and allow to formally specifying the heuristics. The paper highlights certain deficiencies detected when applying this methodology and suggests some improvements.

Keywords Usability · Heuristic evaluation · Usability heuristics · Grid computing · Methodology

1 Introduction

Different methods allow measuring the degree of usability of interactive software products. The heuristic evaluation is one of the most used methods which allow

D. Quiñones(✉) · C. Rusu · S. Roncagliolo
Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: danielacqo@gmail.com, {cristian.rusu,silvana}@ucv.cl

V. Rusu
Universidad de Playa Ancha de Ciencias de la Educación, Valparaíso, Chile
e-mail: virginica.rusu@upla.cl

C.A. Collazos
Universidad del Cauca, Popayán, Colombia
e-mail: ccollazo@unicauca.edu.co

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_43

finding more usability issues, in comparison to other methods. However, sets of generic usability heuristics may not evaluate (usability) aspects related to specific interactive software systems. Particularly, Grid Computing has specific features which could potentially not be considered by generic heuristics [1].

Grid Computing is a distributed system that allows coordinating different computers in order to process a task that demands a lot of resources and processing power. This kind of distributed system facilitates the sharing, accessing and managing information, through the collaboration of several computers forming the grid [2].

As Grid Computing has specific characteristics, the generic usability heuristics (such as Nielsen's [3]) cannot fully evaluate the usability of such applications [4]. A set of specific Grid Computing heuristics was developed using the methodology proposed by Rusu et al. [5]. The methodology facilitated the heuristics' design and specification, as it establishes stages that support the developing process and allow to formally specifying the heuristics, considering the appropriate validations and refinements. However, certain deficiencies were detected when applying this methodology, so further improvements are necessary.

Section 2 presents concepts related to usability, as its definition and evaluation methods. Section 3 presents what is Grid Computing. Section 4 presents the set of usability heuristics developed for Grid Computing using the above mentioned methodology. Section 5 presents the process of developing Grid Computing usability heuristics, how the methodology was applied; deficiencies detected when applying it; and suggestions for improvements. Section 6 presents conclusions and future work.

2 Usability

2.1 Definition

The ISO 9241-11 standard [6] defines usability as: "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". ISO standards updates still refer to the ISO 9241 usability definition [7].

On the contrary, according to ISO/IEC 9126-1, usability can be defined as: [8]: "The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions". The ISO/IEC 25010:2011 standard replaced ISO/IEC 9126-1 standard adding and changing certain elements.

Nevertheless, a generally accepted usability definition still doesn't exist, as its complex nature is hard to describe in one definition. Lewis highlights two major conceptions of usability evaluations: summative and formative [9].

On the one hand, regarding the summative definition, a product is usable when people can use it for its intended purpose effectively, efficiently, and with a satisfaction feeling. On the other hand, regarding the formative definition, the presence of usability depends on the absence of usability problems [9].

In order to evaluate the usability, several evaluation methods have been developed, essentially classified as: (1) Empirical usability testing, based on users' participation [10]; and (2) Inspection methods, based on experts' judgment [11].

2.2 Usability Evaluation Methods

A method of usability evaluation can be defined as: "a procedure composed by a series of well-defined activities for data recollection related to end user's interaction with a software product and/or how a specific feature of this software product contributes in achieving a certain degree of usability" [12].

Usability evaluation methods, generally, can be classified into two broad categories: (1) usability inspections, which are revisions made by evaluators who inspect and evaluate - with their own judgment - the degree of usability of a product, without the participation of users; and (2) usability tests, which include real users, who evaluate a working system.

Usability inspections are set of methods that are all based on having evaluators inspect the interface. Typically, usability inspection is aimed at finding usability problems in a design, though some methods also address issues like the severity of the usability problems and the overall usability of an entire design [13].

Heuristic evaluation is an inspection method performed by a small set of evaluators who judge a user interface in order to determine whether it meets usability principles or heuristics. Each evaluator judges the interface individually, and when the usability problems are found, they gather to collect the results [14].

Usability tests are methods where different users are asked to run a working system prototype, and evaluate it with the goal of collecting information in order to improve the usability of a software product [15].

3 Grid Computing

Grid Computing can be defined as: "a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities" [16].

Grid Computing have emerged as a global cyber-infrastructure and it distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications, and, in some cases, high-performance orientation [17].

The importance of grids is defined in terms of the amount of work they are able to deliver over a period of time. Grid Computing enables collaboration among multiple organizations for sharing of resources. This collaboration is not limited to file exchange and implies direct access to computing resources. Members of the grid can dynamically be organized into multiple virtual organizations [18].

This kind of distributed system provides: (1) Secure access to all computers, computing capacity, data integrity, access security, and so on; (2) consistent

service based on standards and thus the operations on the grid are defined by these standards; and (3) Access and extract resources from the grid from anywhere and with all the required power [19].

Several scientific communities, such as physics, geophysics, astronomy, and bioinformatics, are using grids to share, manage and process large data sets [20]. Some of the applications of Grid Computing are [18]:

- Distributed computing: simulations, numerical calculation tools, data analysis processes and extracting knowledge of data warehouses.
- Intensive data processing: managers distributed databases.
- Distributed real-time systems: used in medicine for image processing for artificial vision.
- Virtual collaborative environment: tele-immersion.
- Specific services: Applications that provide access to specific hardware to perform remote tasks.

4 Grid Computing Usability Heuristics

One of the most widely used methods to evaluate the usability of interactive software is the heuristic evaluation. As Grid Computing has evolved, the generic usability heuristics (such as Nielsen [3]) cannot fully evaluate the usability of specific applications, so it is necessary to develop a new set of heuristics to adapt to new realities [4]. As Grid Computing has evolved from scripts to portals and web interfaces, is necessary to have usability heuristics for Grid Computing for this new perspectives [21, 22].

In order to develop specific usability heuristics for Grid Computing applications, an iterative methodology proposed by Rusu et al. was followed [5]. This methodology includes six steps (stages) that can be iterating in order to perform several validations and improvement to the set of usability heuristics. The methodology and its stages are presented below.

- STEP 1: An exploratory stage, to collect bibliography related with the main topics of the research: specific applications, its characteristics, general and/or related (if there are some) usability heuristics.
- STEP 2: A descriptive stage, to highlight the most important characteristics of the previously collected information, in order to formalize the main concepts associated with the research.
- STEP 3: A correlational stage, to identify the characteristics that the usability heuristics for specific applications should have, based on traditional heuristics and case studies analysis.
- STEP 4: An explicative stage, to formally specify the set of the proposed heuristics, using a standard template.

- STEP 5: A validation (experimental) stage, to check new heuristics against traditional heuristics by experiments, through heuristic evaluations performed on selected case studies, complemented by user tests.
- STEP 6: A refinement stage, based on the feedback from the validation stage.

STEP 1 explores the specific applications that require new usability heuristic. STEP 2 re-examines the very meaning of usability and its characteristics, in the context of the examined applications. If literature provides no specific and/or related usability heuristics, Nielsen's 10 well known and extensively used heuristics are used as a basis at STEP 3. The standard template used at STEP 4 is the following:

- ID, Name and Definition: Heuristic's identifier, name and definition.
- Explanation: Heuristic's detailed explanation, including references to usability principles, typical usability problems, and related usability heuristics proposed by other authors.
- Examples: Examples of heuristic's violation and compliance.
- Benefits: Expected usability benefits, when the heuristic is accomplished.
- Problems: Anticipated problems of heuristic misunderstanding, when performing heuristic evaluations.

STEP 5 evaluates the set of heuristics defined at STEP 4 against Nielsen's heuristics, in specific case studies. The application is evaluated by two separate groups of evaluators, of similar experience, in equal conditions. One group uses only the set of heuristics defined at STEP 4, while the second group uses only Nielsen's heuristics. Usability problems founded by the two groups are then compared.

STEP 6 refines the set of heuristics defined at STEP 4. Stages 1 to 6 may be applied iteratively. Specific usability checklist may also be developed, detailing usability heuristics and helping heuristic evaluations practice.

Using the above methodology, a set of 12 new heuristics and a checklist were developed, specified, validated and refined in a four-cycle iterative process. The new set of heuristics was presented in [1], and then an improved version of this heuristics – including a checklist for each heuristic – was presented in [23]. The set of usability heuristics were grouped in three categories: (1) Design and Aesthetics, (2) Navigation, and (3) Errors and Help. A summary of the proposed heuristics is presented below, including heuristics' ID and name.

- Design and Esthetics Heuristics: (H1) Clarity; (H2) Metaphors; (H3) Simplicity; (H4) Feedback; (H5) Consistency.
- Navigation Heuristics: (H6) Shortcuts; (H7) Low memory load; (H8) Explorability; (H9) Control over actions.
- Errors and Help Heuristics: (H10) Error prevention; (H11) Recovering from errors; (H12) Help and documentation.

5 The Process of Developing Grid Computing Usability Heuristics

5.1 Applying a Formal Methodology

The set of 12 heuristics proposed for heuristic evaluations of Grid Computing applications was developed using the above methodology proposed by Rusu et al. [5]. Following this methodology, the new set of heuristics was specified, validated and refined in a four-cycle iterative process. Figure 1 presents the iterative process. Iterations are marked as "It. n".

There are few similar methodologies, as the ones proposed by Lechner, Fruhling, Petter and Siy (2013) [24]. However, the detailed analysis of these methodologies is beyond the scope of the paper.

In the first iteration steps 1 to 5 were performed. In step 1 (exploratory stage) bibliography related with usability, usability heuristics and Grid Computing has been collected. In step 2 (descriptive stage) the main concepts have been formalized. In step 3 (correlational stage) we have defined the list of the features that the usability heuristics for Grid Computing should have. Step 4 (explicative stage) formally specified the set of 12 usability heuristics for Grid Computing.

In step 5 (validation stage) the 12 proposed Grid Computing usability heuristics were checked against Nielsen's 10 heuristics, using GreenView as case study [1]. The results seem to prove that Grid Computing heuristics work better than Nielsen's heuristics because more usability problems were captured using Grid Computing than using Nielsen's heuristics; and the average severity of the problems identified using Grid Computing heuristics is higher than the average severity of the problems identified using Nielsen's heuristics [23]. However, the question that arises is: why some usability problems were not identified using Grid Computing heuristics?

In order to answer the previous question, a second iteration was performed. It included only step 5 (validation stage). A usability test was designed and performed, with 5 users. The test was focused on the 6 usability problems identified only by Nielsen's heuristics. All these problems were not in fact perceived as real problems by users, which mean that evaluators using Grid Computing heuristics subjectively ignored the usability problems.

In a third iteration, the 12 proposed Grid Computing usability heuristics were checked against Nielsen's 10 heuristics, using GreenLand as case study (step 5, validation stage) [1]. The results proved once again that Grid Computing heuristics work better than Nielsen's heuristics.

In a fourth iteration, based on the experiments performed, on the nature of the usability problems identified when applying Grid Computing heuristics, and on the problems that some evaluators had when applying such heuristics, step 4 (explicative stage) was performed again, defining a usability checklist. It details Grid Computing heuristics and help their use in heuristic evaluation practice [23].

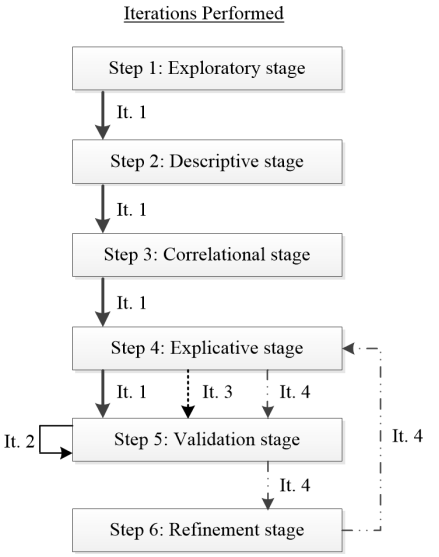


Fig. 1 Methodology applied with four iterations

5.2 Improving the Process of Developing Usability Heuristics

The methodology used to develop the set of heuristics for Grid Computing has been proposed by Rusu et al. [5].

Until now, the methodology has been applied in several research works, in order to establish different sets of specific usability heuristics, as: Touchscreen-based Mobile Devices [24], Interactive digital television applications [25] and Cultural – Oriented Usability Heuristics [26], which it shows the need for a methodology to establish usability heuristics to support the process. Authors of such works have reported the advantages of using the particular set of heuristics against traditional Nielsen's ones.

It is important to use a methodology that formally defines usability heuristics for a specific domain. If heuristics are not designed properly, they could not correctly evaluate the usability of a particular application, or the evaluation instrument could be difficult to implement because:

1. It is difficult to understand by the evaluators,
2. It is not effective and/or efficient,
3. Heuristics' specification is incorrect or incomplete.

The above mentioned methodology facilitated the heuristics' design and specification, as it establishes stages that support the development process and allow to formally specifying the heuristics, considering the appropriate validations and refinements. However, certain deficiencies were detected when applying this methodology. Some suggested improvements needed are listed below.

For step 1: Exploratory stage, no suggestion for now.

For step 2: Descriptive stage, in addition to highlighting the most important features, it is important to connect all this information and give weight (indices) to each element and feature in order to determine which are most relevant and what aspects to regard for the specification of the heuristics.

For step 3: Correlational stage, it is important to make clear the need to consider features of the application in study collected in previous stages, and not only based on traditional heuristics.

For step 4: Explicative stage, the term explicative from itself does not give an enough clear idea about the purpose of this stage. It would be appropriate to use another term that describes better what is done at this stage, such as "Specification stage" or "Formalization stage".

Besides, it would be useful to have a tool to support the specification and management of heuristics, to automate the development process and facilitate their design. For example, prioritize the most important heuristics and report missing elements in the specification of any heuristic elements.

For step 5: Validation stage, it would be appropriate to systematize the validation process. Validation methods should be indicated, suggesting the types of experiments that can be performed to validate the proposal, what they are and what are the benefits of performing a certain experiment.

Also, it would be useful to specify when the validation process has reached a point where the level of specification of the set of heuristics is considered sufficient. This could be determined for example, when there have been several experiments with good results, or when the refinement stage no longer considers relevant changes in the heuristics specification (such as definition or explanation), but only minor elements that do not significantly affect the scope thereof.

Furthermore, this step specifies that the new heuristics work well when they encounter more problems than using Nielsen's heuristics. In that sense, the proposal could be contrasted with other sets of heuristics. If the new set of heuristics works well, it would be important to define what the next steps to follow are.

For step 6: Refinement stage. The proposed methodology ends with this stage, however, is it always necessary to end with this stage? According to the methodology, any iteration should be always completed in Step 6, but in practice this is not always the case. For example, if the heuristics have been refined three times (there have been several iterations) one can finish with Step 5, because after the refinements they are validated and have achieved good results. It would be important to specify in the methodology that not all iterations must end with Step 6 if validations conclude that the set of heuristics works well. In other words, it is possible to finalize the process at a stage that does not necessarily correspond to a new refinement.

The methodology allows applying its steps iteratively. Applying iterations is very useful. It allows the improvement of the set of heuristics, performing new experiments. However, if much iteration is done, this may generate confusion. At what stage should start the next iteration? Is it necessary to refine the set of

heuristics and then begin the new iteration or one can perform various validations on the same set of heuristics? When is worth iterating?

These questions arise when applying the methodology. Therefore we recommend explaining how to apply iterations, when it is appropriate to iterate, and what steps should be repeat depending on the goals to be achieved:

- If adding new aspects that were not considered in the set of heuristics is necessary, one must return to Step 1: Exploratory stage, to investigate other features and aspects not considered,
- If the set of heuristics should be refined (change definitions, add new heuristics, or add new items to checklists), one must return to Step 4: Explicative stage,
- If additional validation of the new set of heuristics from another point of view applying different experiments is necessary, one must return to Step 5: Validation stage.

On the other hand, it would be appropriate to propose a methodology with a higher formalization degree. For example, represent the methodology by diagrams; or using a formal specification language for heuristics and weight indices.

6 Conclusions and Future Work

As new systems and applications emerge, specific usability heuristics are needed in order to evaluate effectively and efficiently their usability. Several sets of specific usability heuristics have been designed, usually working better than generic heuristics. However, there is a need for a formal methodology for developing usability heuristics.

The methodology proposed by Rusu et al. [5] has been used to develop several sets of heuristics for specific applications [25], [26], [27], proving that it is necessary to have a methodology to establish usability heuristics to support the process. The methodology allowed formalizing the process for establishing 12 Grid Computing usability heuristics, in several iterations. The set of usability heuristics for Grid Computing that we proposed, were validated in two case studies (GreenView and GreenLand [1]). They proved that work better than Nielsen's generic heuristics.

However, the methodology applied has deficiencies in explaining stages and generates confusion on how to iterate and how to apply it properly. It is necessary to deepen what means each stage, suggesting procedures to follow, support tools and validation techniques. Also, the methodology should explain better how to iterate, what steps should iterate and when the process may be finalized.

As future work, the methodology proposed back in 2011 will be detailed, improved (as shown above), formalized and experimentally validated.

Acknowledgments The study was developed by members of the "UseCV" Research Group and was highly supported by the School of Informatics Engineering (Escuela de Ingeniería Informática) of the Pontifical Catholic University of Valparaíso (Pontificia Universidad Católica de Valparaíso) – Chile. Daniela Quiñones has been granted the "INF-PUCV" Graduate Scholarship.

References

1. Rusu, C., Roncagliolo, S., Tapia, G., Hayvar, D., Rusu V., Gorgan, D.: Usability heuristics for grid computing applications. In: Proceedings ACHI 2011: The Fourth International Conference on Advances in Computer-Human Interactions, pp. 53–58 (2011)
2. Dong, F., Akl, S.G.: Scheduling algorithms for grid computing: state of the art and open problems. In Techinal report, vol. 504 (2006)
3. Nielsen, J.: Ten Usability Heuristics (1995). http://www.useit.com/papers/heuristic/heuristic_list.html
4. Rusu, C., Gorgan, D., Roncagliolo, S., Rusu, V.: Dezvoltarea de euristici de utilizabilitate Studiu de caz: aplicaţii Grid Computing. Proceedings of the National Conference on Human-Computer Interaction **4**(11), 63–66 (2011)
5. Rusu, C., Roncagliolo, S., Rusu, V., Collazos, C.: A methodology to establish usability heuristics. In: Proceedings ACHI 2011: The Fourth International Conference on Advances in Computer-Human Interactions, pp. 59–62 (2011)
6. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDT's) – Part 11: Guidance on usability. International Organization for Standardization, Geneva (1998)
7. Rusu, C., Rusu, V., Roncagliolo, S., González, C.: Usability and user experience: what should we care about? International Journal of Information Technologies and Systems Approach **8**(2), 1–12 (2015)
8. ISO/IEC 9126-1: Software Engineering - Product Quality (2001)
9. Lewis, J.: Usability: lessons learned … and yet to be learned. International Journal of Human-Computer Interaction **30**(9), 663–684 (2014)
10. Dumas, J., Fox, J.: Usability testing: current practice and future directions. In: Sears, A., Jacko, J. (eds.) The Human - Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 1129–1149. Taylor & Francis, New York (2008)
11. Cockton, G., Woolrych, A., Lavery, D.: Inspection – based evaluations. In: Sears, A., Jacko, J. (eds.) The Human – Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 1171–1189. Taylor & Francis, New York (2008)
12. Fernández, A., Insfran, E., Abrahão, S.: Usability evaluation methods for the web: a systematic mapping study. Journal Information and Software Technology **53**(8), 789–817 (2011)
13. Nielsen, J.: Usability inspection methods. In: Proceeding CHI 1995 Conference Companion on Human Factors in Computing Systems, pp. 377–378 (1995)
14. Scholtz, J.: Usability evaluation. In: National Institute of Standards and Technology (2004)
15. Nielsen, J.: Usability Engineering. Academic Press (1993)
16. Foster, I., Kesselman, C.: The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann Publishers, USA (1999)

17. Foster, I., Kesselman, C., Tuecke, S.: The anatomy of the grid: enabling scalable virtual organizations. International Journal of High Performance Computing Applications **15**(3), 200–222 (2001)
18. Magoulés, F., Pan, J., Tan, K., Kumar, A.: Introduction to Grid Computing. CRC Press, Taylor & Francis Group, UK (2009)
19. Linderoth, J.: An Introduction to the Computational Grid. University of Wisconsin (2007)
20. Yu, J., Buyya, R.: A taxonomy of scientific workflow systems for grid computing. ACM Sigmod Record **34**(3), 44–49 (2005)
21. Dove, M., Walker, A., White, T., Bruin, R., Austen, K.: Usable grid infrastructures: practical experiences from the eMinerals project. In: Proc. UK e-Science All Hands Meeting, pp. 48–55 (2007)
22. Doyle, T.: GridPP: the UK's contribution to the international collaboration building a worldwide Grid, the LHC Computing Grid: is the system usable? University of Glasgow (2006)
23. Roncagliolo, S., Rusu, V., Rusu, C., Tapia, G., Hayvar D., Gorgan, D.: Grid computing usability heuristics in practice. In: Proceedings - 2011 8th International Conference on Information Technology: New Generations, pp. 145–150 (2011)
24. Lechner, B., Fruhling, A., Petter S., Siy, H.: The chicken and the pig: user involvement in developing usability heuristics. In: Proceedings of the Nineteenth Americas Conference on Information Systems (2013)
25. Inostroza, R., Rusu, C., Roncagliolo, S., Rusu, V., Collazos, C.: Developing SMASH: a set of SMArtphone's uSability heuristics. Computer Standards & Interfaces **43**, 40–52 (2016)
26. Solano, A., Rusu, C., Collazos, C., Arciniegas, J.: Evaluating interactive digital television applications through usability heuristics. Ingeniare, Revista chilena de ingeniería **21**(1), 16–29 (2013)
27. Díaz, J., Rusu, C., Pow-Sang, J., Roncagliolo, S.: A cultural – oriented usability heuristics proposal. In: Proceedings of the 2013 Chilean Conference on Human - Computer Interaction (ChileCHI 2013), pp. 82–87 (2013)

Refining Timing Requirements in Extended Models of Legacy Vehicular Embedded Systems Using Early End-to-end Timing Analysis

Saad Mubeen, Thomas Nolte, John Lundbäck, Mattias Gålnander and Kurt-Lennart Lundbäck

Abstract Model and component-based development approaches have emerged as an attractive option to deal with the complexity of vehicle software. Using these approaches, we provide a method to estimate and refine end-to-end timing requirements in vehicular distributed embedded systems that are developed by reusing the models of legacy systems. This method leverages on the early end-to-end timing analysis that can be performed at the highest abstraction level during the development of these systems. As a proof of concept, we conduct a vehicular-application case study to show the process of estimating and refining the timing requirements early during the development. The case study is modeled with the Rubus-ICE tool suite that is used for the software development of vehicular embedded systems by several international companies.

1 Introduction

The vehicle industry has witnessed a significant increase in the size and complexity of vehicle software in the past few decades [7]. In order to deal with such complexity, the research community has proposed to use the principles of Model-Based Software Engineering (MBSE) and Component-Based Software Engineering (CBSE) [9, 12]. MBSE supports the use of models to describe functions, structures and other design

S. Mubeen(✉) · T. Nolte
Mälardalen University, Vasteras, Sweden
e-mail: {saad.mubeen,thomas.nolte}@mdh.se

J. Lundbäck · M. Gålnander · K.-L. Lundbäck
Arcticus Systems AB, Jarfalla, Sweden
e-mail: {john.lundback,mattias.galnander,kurt.lundback}@arcticus-systems.com

The work in this paper is supported by the Swedish Foundation for Strategic Research and ARTEMIS within the projects PRESS and CRYSTAL.

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_44

artifacts. CBSE supports the development of large software systems by reuse and integration of software components and their architectures. A large number of vehicle functions are real-time systems. The developer of such a system is required to guarantee that the system delivers logically correct response within a specified time (e.g., a deadline). There are a large number of schedulability analysis techniques that can provide such guarantees [13, 16]. In this paper we focus on the end-to-end timing analysis [10, 13, 14].

The EAST-ADL methodology [3], renowned in the vehicle domain, describes the software development mainly at four abstraction levels as shown in Fig. 1. The vehicle level captures the features and requirements on the end-to-end functionality of the vehicle in an informal and solution-independent way. The analysis level formally captures the requirements. It supports a high-level analysis for functional verification. At the design level, the artifacts are developed independent of implementation details. These artifacts may also contain middleware abstraction, hardware architecture and software functions to hardware allocation. The implementation level contains the software implementation (software architecture) of the system. There are several development methodologies, models and languages that are used at these levels; some of them are shown in Fig. 1.

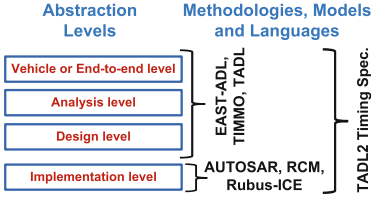


Fig. 1 Abstractions in vehicle software development.

Objective and Paper Contribution. When the top-down software development approach is used, the detailed timing information that is required to perform the end-to-end timing analysis is not available at the higher abstraction levels. In fact, such information is available only at the implementation level. There are several models and tool chains in the vehicle domain that support the timing analysis at the implementation level, e.g., AUTOSAR tool chain [2, 18] and Rubus-ICE [4, 14]. In the vehicle industry, most often, the bottom-up development approach is used as a lot of information, models, artifacts and solutions are reused from other projects and legacy (previously developed) systems. It is estimated that up to 90% of the software can be reused from other projects or previous releases of the vehicle if MBSE and CBSE are used [17].

In our point of view, if a new vehicular distributed embedded system is to be developed using the bottom-up approach (i.e., by reusing the models and communications from legacy systems) then the timing information can be extracted to perform end-to-end timing analysis at the vehicle level. The analysis performed at such a high

level of abstraction can be regarded as the early end-to-end timing analysis. It should be noted that the precision of the analysis depends upon the level of details about the communications and precision of the assumptions about the internal software architectures of the nodes that are not yet developed. The end-to-end timing requirements depend upon the nature of the vehicle feature. However, when the bottom-up development approach is used, the extracted timing information can be utilized for the estimation and refinement of end-to-end timing requirements at the vehicle level. On the other hand, if the timing requirements are already specified then the extracted timing information can help in making design decisions and in refining the software architecture of the new nodes such that the specified requirements are satisfied.

Using this vision, we have recently developed a method for early timing analysis [15]. In this paper, we take a few steps ahead by providing a method to estimate and refine end-to-end timing requirements on the models of vehicular distributed embedded systems that are developed using the models of legacy systems. It leverages on the early end-to-end timing analysis at the vehicle level. As a proof of concept, we show the process of estimating and refining the timing requirements by conducting a vehicular-application case study using Rubus-ICE.

2 Background and Related Work

End-to-end Timing Analysis. The end-to-end timing analysis [10, 14] is a pre-runtime analysis that calculates upper bounds on the end-to-end path delays. Consider a task chain shown on the left side in Fig. 2. The Worst Case Execution Time (WCET) of each task is also shown. Since each task is activated independently with a different clock, there can be several paths through which the data can traverse from input to output. Hence, there can be multiple outputs (values in Reg-4) corresponding to one input (value in Reg-1) of the task chain as shown by the uni-directional curved arrows on the right side in Fig. 2. This results in different end-to-end path delays. Two such delays, namely *age* and *reaction* that find their importance in the controls systems and body electronics domains respectively, are illustrated in Fig. 2. The execution scenario shown in Fig. 2 is created for simplicity and it may not represent exact worst-case scenario. The age delay is equal to the time elapsed between the current non-overwritten release of τ_1 and corresponding last response of τ_3. The reaction delay is equal to the time elapsed between the previous non-overwritten release of τ_1 and

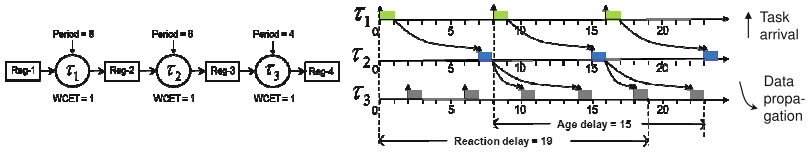


Fig. 2 Example demonstrating the end-to-end path delays.

the first response of τ_3 corresponding to the current non-overwritten release of τ_1. A detailed discussion and calculations for these delays can be found in [1, 10, 14].

Models and Methodologies. There are several frameworks that support timing modeling and analysis such as AADL, SCADE, MAST, SysML and CHESS. However, the focus in the vehicle industry today is on the EAST-ADL language at the higher abstraction levels while AUTOSAR [2] and Rubus Component Model (RCM) [11] are used at the implementation level. We focus only on the vehicle domain. TIMMO-2-USE [5] is an initiative that is jointly driven by the industry and academia to provide AUTOSAR with a timing model. It is based on a methodology and the TADL2 language [1] which expresses timing constraints at all the abstraction levels. TADL2 is inspired by the MARTE [6] profile of UML. The methodology uses EAST-ADL for structural modeling at the top three levels. Whereas, it proposes to use AUTOSAR at the implementation level. AUTOSAR is an industrial initiative to provide a standardized software architecture for the development of vehicle software. AUTOSAR supports end-to-end timing analysis only at the implementation level [2, 18]. A recent work [8] presents a methodology to support high-precision end-to-end timing analysis on the EAST-ADL models of the systems. To the best of our knowledge, none of these approaches support high-precision end-to-end timing analysis at the highest abstraction level.

The Rubus Concept. Rubus [4] is a collection of methods and tools for model-driven development of control functionality in vehicles. It is used by several international companies. The Rubus concept is based around RCM and its development environment Rubus-ICE. The overall goal of Rubus is to be aggressively resource efficient and to provide means for developing predictable, timing analyzable and synthesizable control functions in resource-constrained vehicular embedded systems. The lowest-level hierarchical component in RCM is called Software Circuit (SWC). Its purpose is to encapsulate basic functions. Rubus-ICE supports distributed end-to-end response-time and delay analyses [14, 15] at various abstraction levels shown in Fig. 1.

3 Vehicular-application Case Study

3.1 Modeling of the Adaptive Cruise Control (ACC) System by Reusing a Legacy Cruise Control (CC) System

The ACC system extends the CC system by adapting the vehicle speed according to the outside traffic. It uses proximity sensors such as a radar to create a feedback of distance to and velocity of the preceding vehicle. It controls the vehicle speed based on the feedback. The model of the legacy CC system consists of four nodes namely Cruise Control, Engine Control, Brake Control and Graphical-user interface Control.

These nodes are denoted by CC_ECU, EC_ECU, BC_ECU and GC_ECU in Fig. 3 respectively. The nodes communicate via a CAN bus.

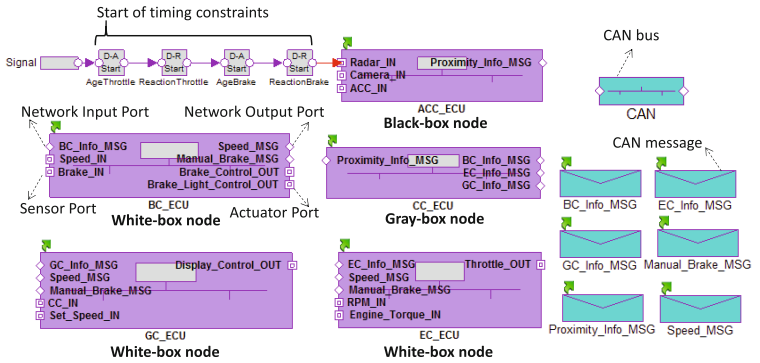


Fig. 3 Model of the Adaptive Cruise Control system at the vehicle level in Rubus-ICE.

The ACC system can be modeled by reusing the models of the four nodes from the CC system and by introducing a new node denoted by ACC_ECU that provides the adaptive cruise control functionality as shown in Fig. 3. We assume that the internal software architectures of EC_ECU, BC_ECU and GC_ECU can be completely reused. That is, the new ACC system does not require any modification in the internal software architectures and communication decisions of these three nodes. Hence, the three nodes are modeled as white-box nodes. However, CC_ECU communicates with the ACC_ECU. This means, the internal software architecture of CC_ECU requires extensions at the implementation level to handle the communication from ACC_ECU. Hence, CC_ECU is modeled as a gray-box node. Since, the internal software architecture of ACC_ECU is not available at the vehicle level, it is modeled as a black-box node. We assume that the decisions concerning the external communications for ACC_ECU can be made at the vehicle level. The speed of the CAN bus is 250 Kbit/s. The model of the CAN bus contains six messages as shown in Fig. 3. The attributes of the messages are shown in Table 1. The priority of each CAN message is assumed to be equal to its ID. Also, all messages are periodic.

Model of ACC_ECU. The black-box model of ACC_ECU is shown in Fig. 3. There are three sensor ports and one network output port that are specified on the black-box model. The sensor ports receive radar signals, camera feed and the driver's input corresponding to the ACC settings. The network output port sends a CAN message to ACC_ECU. This message carries the information about other vehicles in the vehicle's proximity. The start objects for the four end-to-end timing constraints are specified on the Radar_IN sensor port. We will come back to the discussion of timing constraints in the next subsection.

Table 1 Attributes of the CAN messages.

Message	ID	Size (Bytes)	Period (ms)	Source ECU	Destination ECUs
Proximity_Info	0	8	10	ACC	CC
Speed	1	8	10	BC	EC,GC
EC_Info	2	8	10	CC	EC
Manual_Brake	3	2	10	BC	EC,GC
BC_Info	4	8	10	CC	BC
GC_Info	5	8	10	CC	GC

Model of CC_ECU. The gray-box model of CC_ECU is shown in Fig. 3. There are three network output ports that are reused from the legacy model of CC_ECU. There is a new network input port that is specified on this model. This port is responsible for handling communication from ACC_ECU. The internal software architecture of this gray-box node is modeled with three SWCs denoted by CAN_Input_SWC, CC_Mode_Control_SWC and CC_Processing_SWC as shown in Fig. 4. These SWCs are responsible for receiving Proximity_Info_ MSG from ACC_ECU; performing CC mode control calculations; and sending three CAN messages denoted by BC_Info_MSG, EC_Info_MSG and GC_Info_MSG.

Model of BC_ECU. There are two sensor ports; one network input port; two network output ports; and two actuator ports that are specified on the white-box model of BC_ECU as shown in Fig. 3. The internal software architecture of this node is modeled with five SWCs as shown in Fig. 4. Two SWCs that are denoted by Speed_SWC and Brake_SWC handle the sensor values. BC_Processing_SWC is responsible for processing all the inputs including the sensor values and the CAN message, denoted by BC_Info_MSG, that is received from CC_ECU. The main purpose of this SWC is to perform the calculations for the brake control functionality. It also sends two CAN messages, namely Speed_MSG and Manual_Brake_MSG, that carry information regarding the vehicle speed and manual brake status (if the brake peddle is pressed or not). Brake_Control_SWC and Brake_Light_Control_SWC are responsible for computing the control signals for the brake actuators and brake light controllers.

Model of EC_ECU. There are two sensor ports; one actuator port; and three network input ports that are specified on the legacy white-box model of EC_ECU as shown in Fig. 3. The internal software architecture of EC_ECU is modeled with four SWCs as shown in Fig. 4. Engine_Torque_SWC and RPM_SWC are responsible for handling the sampled sensor inputs that correspond to the engine torque and rotational speed of the vehicle. EC_Processing_SWC is responsible for processing sensor inputs and the CAN messages (EC_Info_MSG, Speed_MSG and Manual_Brake_MSG). It also performs the calculations for the engine control functionality. Throttle_Control_SWC is responsible for computing the control signals for the engine throttle actuator.

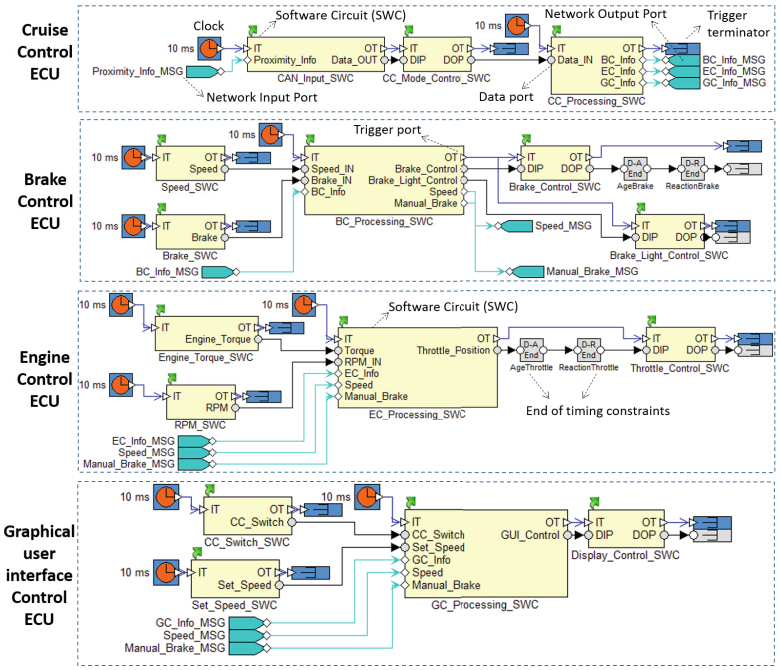


Fig. 4 Internal software architectures of the CC, BC, EC and GC ECUs.

Model of GC_ECU. There are two sensor ports; one actuator port; and three network input ports that are specified on the legacy white-box model of GC_ECU as shown in Fig. 3. The internal software architecture of GC_ECU is modeled with four SWCs as shown in Fig. 4. CC_Switch_SWC and Set_Speed_ SWC are responsible for acquiring and processing the driver's inputs. GC_Pro-cessing_SWC performs the calculations for the graphical-user interface functionality based on the sensor values and the CAN messages (GC_Info_MSG, Speed_MSG and Manual_Brake_MSG). The information to be shown on the display is processed by Display_Control_SWC.

3.2 Modeling of Timing Information and Requirements

The WCETs of all SWCs (range: 60μs-200μs) in the software architectures of all white-box nodes (EC_ECU, BC_ECU, GC_ECU) are taken from the legacy models. The existing SWCs in the software architecture of CC_ECU perform the cruise control functionality based on one extra input, i.e., a CAN message (Proximity_Info_MSG) that is received from ACC_ECU. The effect of this extra input can be compensated by inflating the WCETs of all SWCs in CC_ECU by a factor that can be estimated by expert designers and integrators at the vehicle level. Based

on our experience from various industrial case studies [14], we inflate the WCETs of all SWCs in CC_ECU by 30%. Accordingly, the WCETs of CAN_Input_SWC, CC_Mode_Control_SWC and CC_Processing_SWC are 100 μs, 100 μs and 120 μs. These WCETs can be refined at the implementation level.

We specify four delay constraints on the system. The start objects of these constraints are specified on the Radar_IN senor port of ACC_ECU as shown in Fig. 3. These constraints are denoted by AgeThrottle, ReactionThrottle, AgeBrake and ReactionBrake. The end objects of AgeThrottle and ReactionThrottle are specified on the software architecture of EC_ECU as shown in Fig. 4. These two constraints specify the timing requirements on the age and reaction delays between the radar sensor input in ACC_ECU and the production of the engine throttle actuation signals in EC_ECU. The end objects of AgeBrake and ReactionBrake are specified on the software architecture of BC_ECU as shown in Fig. 4. These two constraints specify the timing requirements on the age and reaction delays between the radar sensor input in ACC_ECU and the production of the brake actuation signals in BC_ECU. The two Distributed Chains (DCs), on which the age and reaction constraints are specified, are shown below.

$$DC_1 := \text{Radar\_IN}_{ACC\_ECU} \rightarrow \text{Proximity\_Info\_MSG}_{ACC\_ECU} \rightarrow \text{Proximity\_Info\_MSG}_{CC\_ECU} \rightarrow \text{CAN\_Input\_SWC}_{CC\_ECU} \rightarrow \text{CC\_Mode\_Control\_SWC}_{CC\_ECU} \rightarrow \text{CC\_Processing\_SWC}_{CC\_ECU} \rightarrow \text{EC\_Info\_MSG} \rightarrow \text{EC\_Processing\_SWC}_{EC\_ECU} \rightarrow \text{Throttle\_Control\_SWC}_{EC\_ECU}$$

$$DC_2 := \text{Radar\_IN}_{ACC\_ECU} \rightarrow \text{Proximity\_Info\_MSG}_{ACC\_ECU} \rightarrow \text{Proximity\_Info\_MSG}_{CC\_ECU} \rightarrow \text{CAN\_Input\_SWC}_{CC\_ECU} \rightarrow \text{CC\_Mode\_Control\_SWC}_{CC\_ECU} \rightarrow \text{CC\_Processing\_SWC}_{CC\_ECU} \rightarrow \text{BC\_Info\_MSG} \rightarrow \text{BC\_Processing\_SWC}_{BC\_ECU} \rightarrow \text{Brake\_Control\_SWC}_{BC\_ECU}$$

3.3 Estimating Timing Constraints by Early Timing Analysis

Our goal is to estimate the lower bounds on the specified timing constraints at the vehicle level. The estimates are based on the information from the models of the nodes and network communication that is available from the legacy CC system. In order to perform the end-to-end timing analysis at the vehicle level, we first make a few assumptions about the black-box model of ACC_ECU.

The node-level ports (i.e., three sensor ports and one network output port) that are specified on ACC_ECU are assumed to be triggered by independent periodic clocks. Since all the SWCs that comprise the software architectures of the three white-box nodes and one gray-box node are triggered by the clocks with a periodicity of 10 ms each, we assume that all the node-level ports in ACC_ECU are also triggered periodically with periods of 10 ms. We also assume that there is a direct

data connection between the Radar_IN sensor port and the Proximity_Info_MSG network output port. The data path delay along this connection can be estimated by expert designers and integrators. Using our experience from previous industrial case studies [14], we specify this delay as 1.5 ms. It should be noted that when the internal software architecture of ACC_ECU is available at the implementation level, the WCETs of all the SWCs in the software architecture must comply with the estimated data path delay of 1.5 ms.

We perform the end-to-end path delay analysis of the ACC system using the recent extensions in Rubus-ICE [15]. The response times of CAN messages Proximity_Info_MSG, Speed_MSG, EC_Info_MSG, Manual_Brake_MSG, BC_Info_MSG and GC_Info_MSG calculated by the analysis engines are 2140 μs, 5240 μs, 2480 μs, 6080 μs, 3320 μs and 3320 μs respectively. The age and reaction delays calculated at the vehicle level are shown in the second row (Results I) of Table 2. The analysis engines took 56.9 seconds to calculate the results on a computer with a 64-bit, Core i7, 2.0 GHz processor. The results represent the worst-case delays without considering the software architecture of ACC_ECU. In reality, ACC_ECU can include several SWCs (instead of the direct data connection) along DC_1 and DC_2 at the implementation level where its software architecture will be available. Therefore, more realistic value of a delay constraint can be estimated by summing up the periods of all clocks that trigger the SWCs in each black-box node along the distributed chain and adding it to the calculated delay. Mathematically, it can be represented as:

$$\text{Delay Constraint} = \text{Calculated delay} + \Sigma(\text{Clks\_DC}_{\text{Black-box node}}) \quad (1)$$

where $\Sigma(\text{Clks\_DC}_{\text{Black-box node}})$ is the summation of the periods of all clocks that trigger the SWCs in the black-box node(s) along the distributed chain under analysis. In order to get this value, we need to anticipate the total number of SWCs and the periods of their triggering clocks in ACC_ECU along DC_1 and DC_2. The judgements of expert designers and integrators as well as experiences from the previous projects and case studies can be helpful in this regard. In our case study, we anticipate (based on our experience from several industrial case studies [14]) that each node-level port in ACC_ECU will be handled by a separate SWC. There will be two more SWCs along DC_1 and DC_2 that will be triggered by clocks of periods 10 ms each. This means, according to (1) we need to add 20 ms to each calculated delay to get the first estimation of the age and reaction constraints as shown in the third row (Results II) of Table 2.

Table 2 Calculated age and reaction delays.

Calculated delay(μs)	AgeThrottle	ReactionThrottle	AgeBrake	ReactionBrake
Results I	55540	65540	53700	63700
Results II	75540	85540	73700	83700
Results III	65540	75540	63700	73700

3.4 High-precision Timing Analysis of the ACC System

In order to evaluate the results obtained from the early end-to-end timing analysis, we perform the high-precision end-to-end timing analysis of the ACC system at the implementation level using the analysis framework in Rubus-ICE [14]. At this level, the internal software architecture of ACC_ECU is available as shown in Fig. 5. The WCETs of all the SWCs in ACC_ECU are selected in such a way that the internal delay in ACC_ECU along DC_1 and DC_2 does not exceed the estimated data path delay of 1.5 ms. The calculated age and reaction delays for a 10 ms clock that triggers ACC_Mode_Control_SWC and ACC_Processing_SWC in ACC_ECU are shown in the fourth row (Results III) of Table 2.

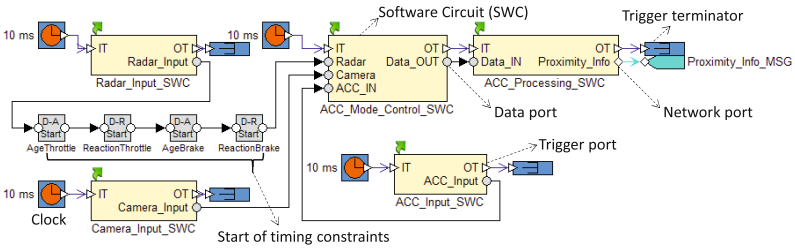


Fig. 5 Internal software architecture of the ACC ECU at the implementation level.

3.5 Comparative Evaluation and Discussion

We compare the calculated worst-case age and reaction delays in the fourth row (Results III) of Table 2 with the estimated timing constraints in the third row (Results II) of Table 2. It is evident that all the timing constraints that are specified on the ACC system, which is modeled by reusing the legacy CC system, are satisfied. This comparison also shows that the timing constraints that are calculated using (1) can be refined by subtracting the period of the clock that triggers the first object in the distributed chains DC_1 and DC_2. This means that the lower bounds on the estimated age and reaction constraints can be obtained by subtracting the period of the clock that triggers the Radar_IN sensor port in ACC_ECU. This generally applies to all cases as long as the clock periods of the SWCs in the black-box node are equal to the period of the clock that triggers the first object in the distributed chain. Otherwise, the estimated age and reaction delays in the second row (Results I) of Table 2 provide lower bounds on the corresponding timing constraints. It should be noted that these lower bounds on the timing constraints strongly depend upon accuracy of the internal delay of the black-box node along the distributed chain that we estimated in Subsection 3.3. If this delay cannot be estimated accurately, it is better to make an over-estimated assumption about it. Hence, it is important that the WCETs of all SWCs in each black-box node (when the software architecture is available at the implementation level) must comply with the estimated internal delay.

4 Conclusion

In this paper we have provided a method to estimate and refine end-to-end timing requirements in vehicular distributed embedded systems that are developed by reusing the models of legacy systems. This method exploits early timing analysis to estimate and refine timing requirements at the highest abstraction level. As a proof of concept, we have conducted a vehicular case study using an existing industrial component model and tool suite. Using this method, we have calculated the end-to-end delays in the case study at the highest abstraction level to estimate and refine the timing requirements. The comparison between the estimated timing requirements at the highest abstraction level and the actual end-to-end delays at the lowest abstraction level shows that our method provides correct lower bounds on the end-to-end timing requirements (that correspond to the worst-case age and reaction delays). The proposed method can be used complementary to EAST-ADL. We believe, the tools that implement our method may prove helpful for the software development organizations in the vehicular domain to decrease the costs for development and testing.

References

1. Timing Augmented Description Language (TADL2) syntax, semantics, metamodel Ver. 2, Deliverable 11, August 2012
2. AUTOSAR Techincal Overview, Release 4.1, Rev. 2, Ver. 1.1.0., The AUTOSAR Consortium, October 2013. http://autosar.org
3. EAST-ADL Domain Model Spec., V2.1.12. http://www.east-adl.info/Specification/V2.1.12/EAST-ADL-Specification_V2.1.12.pdf
4. Rubus models, methods and tools. http://www.arcticus-systems.com
5. TIMMO-2-USE. https://itea3.org/project/timmo-2-use.html
6. The UML Profile for MARTE: Modeling and Analysis of Real-Time and Embedded Systems, January 2010. http://www.omgmarte.org/
7. Broy, M., Kruger, I., Pretschner, A., Salzmann, C.: Engineering automotive software. Proceedings of the IEEE **95**(2), 356–373 (2007)
8. Bucaioni, A., Mubeen, S., Cicchetti, A., Sjödin, M.: Exploring timing model extractions at east-adl design-level using model transformations. In: 12th International Conference on Information Technology: New Generations, April 2015
9. Crnkovic, I., Larsson, M.: Building Reliable Component-Based Software Systems. Artech House Inc., Norwood (2002)
10. Feiertag, N., Richter, K., Nordlander, J., Jonsson, J.: A compositional framework for end-to-end path delay calculation of automotive systems under different path semantics. In: CRTS Workshop, December 2008
11. Hänninen, K., et al.: The rubus component model for resource constrained real-time systems. In: IEEE Symposium on Industrial Embedded Systems (2008)
12. Henzinger, T.A., Sifakis, J.: The embedded systems design challenge. In: 14th International Symposium on Formal Methods (FM) (2006)
13. Joseph, M., Pandya, P.: Finding response times in a real-time system. Computer Journal **29**(5), 390–395 (1986)

14. Mubeen, S., Mäki-Turja, J., Sjödin, M.: Support for end-to-end response-time and delay analysis in the industrial tool suite: Issues, experiences and a case study. Computer Science and Information Systems **10**(1) (2013)
15. Mubeen, S., Sjödin, M., Nolte, T., Lundbäck, J., Gålnander, M., Lundbäck, K.L.: End-to-end timing analysis of black-box models in legacy vehicular distributed embedded systems. In: 21st International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), August 2015
16. Audsley, N.C., et al.: Fixed priority pre-emptive scheduling: an historic perspective. Real-Time Systems **8**(2/3), 173–198 (1995)
17. Thorngren, P.: Keynote Talk: Experiences from EAST-ADL Use, EAST-ADL Open Workshop, Gothenberg, October 2013
18. Henia, R., et al.: System level performance analysis - the symta/s approach. Computers and Digital Techniques **152**(2), 148–166 (2005)

Integrated Metrics Handling in Open Source Software Quality Management Platforms

Julio Escribano-Barreno, Javier García-Muñoz and Marisol García-Valls

Abstract Software quality is of vital importance in software development projects. It influences every aspect of the system such as the functionality, reliability, availability, maintainability, and safety. In critical software projects, quality assurance has to be considered at each level of the initial concept to the software engineering process: from specification to coding and integration. At the lowest coding level, there are several tools that enable the monitoring and control of software quality. One of them is SonarQube, an open source quality management platform, used to analyse and measure technical quality. I can be extended through plugins for customization and integration with other tools. The specific conception and development of these plugins is a significant design effort that ensures the correct handling of the different phases involved in the software quality process. We present an initial design and development of an integrated analyser component for extending the functionality of the open source framework for software quality management.

Keywords Software quality · Monitoring of verification · Software design

1 Introduction

Every software development process needs information about almost every aspect of the software development phase, like achievement of objectives, monitoring and control of activities, project costs and technical quality. Metrics are of key importance in all engineering disciplines and, in particular, for software development (see [13]),

J. Escribano-Barreno
Indra, Alcobendas, Spain
e-mail: jebarreno@indra.es

J. Escribano-Barreno · J. García-Muñoz · M. García-Valls(✉)
Universidad Carlos III de Madrid, Leganés, Spain
e-mail: {jgmunoz,mvalls}@it.uc3m.es

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_45

providing a vital insight into the development process to assess maintainability, reliability and even development progress. Metrics provide reproducible indicators useful to estimate the quality, performance, management, and cost within a project. They bring in benefits like the possibility of analysing the data to understand, improve, and predict future behaviours for undertaking corrective actions on time.

Metrics has been studied and developed through years in order to improve software and systems development. They have increased their relevance due to their applied use and the contrasted benefits to define baseline quality indicators for serveral purposes. For example, the SEI Capability Mature Model Integration (CMMI) [9] for development, relies on the usage of metrics (see [17]), and is used for evaluating the maturity process of the organizations.

The collection and processing of the metrics can involve a significant human effort, that highlights the need of automating the specification of metrics and subsequent data collection that must later be processed. By using automatic analysis tools (such as [4] for static code analysis), the metrics collection effort is significantly reduced. One of the platofms that supports this process is Few open source platforms support the automatic metrics management for this process. Among them, the most popular quality management platform is SonarQube [10][43]. It enables continuous inspection and it supports a number of languages, including Java, C, C++, C#, PHP, and JavaScript. SonarQube provides some *basic* metrics like complexity, duplicated code detection, or lines of code counting, among others. However, this is a very generic functionality that requires to be enhanced for software development projects of a certain complexity. Highly complex projects such as critical software projects require that these basic metrics be enhanced to provide the information required by each particular project, as each project may have to adjust to specific normd. It is then useful and needed to take information from other sources, like other external tools, and stablish the adequate methodology in order to enhance the functional power of this framework in order to provide the required information with a suitable design that makes it easily customizable.

In this paper, we present an approach to enrich the metrics management and presentation to verification engineers based on the SonarQube framework. This has been performed by designing an *integrated metrics analyser* component that provides the enhanced functionality for the platform in order to integrate its own analysis results with the ones from external analysis tools in a single presentation space. The design is a modular one that allows to easily customize the integration of any external code analysis tool. The result is the achievement of a more complete set of metrics that can be managed in the projects to control and monitor the software development process according to the needs of each specific project. The information from the analysis of th project code is then centralized in the platform that is, at the same time, a collaborative environment that allows the remote work of teams of verification engineers. We validate the component by implementing the specific integration with external information sources that provide static analysis metrics (such as Understand [53]); this external tool provides metrics and rule-checks against both, custom and published standards. In our work, the rules and metrics can also be extended to

complement the provided ones. We exemplify technical metrics, related to software quality through actual static analyses on a real critical software project.

The paper is structured as follows. Section 1 includes a brief introduction and motivation for this work. Section 2 describes the related work in what concerns norms and practices related to technical quality and metrics. Section 3 describes the baseline framework offered by SonarQube. Section 4 presents our contribution in the form of a new functionality for the analysis of metrics from different tools, presenting the design details on the enabling plugin for this functionality. Section 5 validates the design through a concrete implementation for an external tool and we show its usage in the context of a real-world software project. Section 6 concludes this paper and presents the continuation work.

2 Background and Related Work

This section describes selected work most related to the objective of the present contribution, mainly concerning the existing norms and regulations for critical software development, the engineering processes that describe the steps to the objective software development, and current tools for software code analysis.

2.1 Norms and Software Quality Tools and Frameworks

Technical metrics are collected through static analysis techniques. Software static analysis is required in several norms related to critical systems, some of which are described here. The norms selected here require the use of static analysis techniques to comply with their objectives. For example, DO-178C [46] and DO-278A [47] introduces the use of metrics to be specified in the Software Quality Assurance Plan. The metrics collection and analysis, additionally supports the compliance of some norms, like North Atlantic Treaty Organization (NATO) AQAP-2210 [2], and its spanish version PECAL-2210 [39].

DO-178C is the norm for *Software Considerations in Airborne Systems and Equipment Certification*, [46]. It is one of the most accepted international standards. Includes additional objectives and it is complemented with the supplements [49], [50], [51] and [52]. The previous version of this norm is DO-178B [48]. DO-278A is the norm for *Guidelines for Communication, Navigation, Surveillance and Air Traffic Management* [47]. It provides guidelines for non-airborne Communication, Navigation, Surveillance and Air Traffic Management (CNS/ATM) systems, including the guidelines for the software assurance activities to be conducted with non-airborne CNS/ATM systems. IEC 61508 is the norm for *Functional safety of electrical/electronic/programmable electronic safety-related systems* [31]. It is the standard for industry automation, intended to be a safety standard applicable to all kinds of industry that includes the complete safety life cycle. Used as basis for other documents, as

railway (CENELEC 50128 [34]), automotive industries (ISO 26262 [31]) or nuclear power plants (IEC 61513 [32]).

The metrics and quality information with respect to a complex software project may easily require different analysis techniques that generate results and data from different sources, possibly also collecting data in different ways. In most projects, more than purely the source code is analyzed, e.g. dependencies among packages such as [14]. Information about the software is collected in different ways by means of different tools such as the ones presented as follows. *Understand* [53] and LDRA [37] are commercial tools for static code analysis, supporting multi-language. PC-Lint [41] is also a commercial solution for static code analysis supporting C and C++ languages. Splint [44] is a GNU licensed tool that supports static code analysis for C programs. PMD [42] supports static code analyser for Java, JavaScript, XML and XSL. SonarQube [10, 43] is an open source quality management platform, developed under LGPL v3 license.

As not a single tool is capable of providing all required data, it is typically required to set up a tool chain for code analysis setting also the procedures and principles for information collection and interpretation. in the form of a collaborative environment. As a result, the SonarQube platform has appeared as a platform to support advanced analysis; however, SonarQube appears in a similar way to a blank sheet of paper, so that the required techniques and methods to implement the needed functionality have to be designed and integrated in it.

The tools mentioned in section 2 are used to statically analyse the code quality. However, the metrics that they yield do not meet all the requirements across different projects. In each project, different metrics can be required, or different implementation languages can be used that vary from a project to another. This situation can make that one tool that is suitable for a project is not valid to other one.

For example, SonarQube is a quality management platform that can reflect and present the different quality aspects of several projects, coded in a number of different languages. This is carried out by the design and implementation of modules that can extend the functionality of the platform. Apart from the very basic set of metrics and data that SonarQube provides for any given project, a mechanism to customize it for particular projects that have different information requirements is needed. This characteristic can be complemented with the metrics available in other external code analysers (such as Understand) to enrich and extend the capability of SonarQube.

One other missing element in the current status of SonarQube is that each user needs to have the tool installed locally; as such, the run-time environment has little orientation to be a collaborative environment that can simultaneously support their access to several projects with a number of users.

We overcome these missing characteristics with the design and implementation of enhancements to SonarQube to support the customizable integration with external tools for code analysis. We design the required plugins to enrich the basic metrics presentation functionality of the platform, combining them in an unique repository that provides collaborative access to different (though shared) software projects.

2.2 *Software Architectures in Critical Systems in Emerging Domains*

An essential element of critical systems is the software architecture as it directly impacts the source code quality and the complexity of the final development. Critical software systems validation focuses, among other aspects, on temporal behavior applying real-time techniques [1, 6–8, 15]. These differ to some extent in the thorough application and verification of the temporal properties, rather providing quality of service mechanisms that are embedded in the software logic. These mechanisms allow dynamic execution preserving timely properties [11, 21–23, 25, 45]. Software quality should account for the verification of the properties of distributed software also related to newer domains as cloud [18], the characteristics of the middleware are integrated in the model [19, 20]. Critical software systems only tolerate off-line and design time verification; however newer domains such as cyber-physical systems require verification techniques to be applied on-line [5, 24]. For distributed environments based on middleware such as [29, 30], on-line decision on correctness of the system composition is applied (e.g. [26–28]). In such emerging domains, the software quality frameworks will have to device new ways of considering properties that will only emerge at execution time.

3 SonarQube Overview

SonarQube is a software quality management platform, multi-language, capable of performing simple analysis over the source code. Basically, it provides information about duplicated code, unit tests, coding standards, code coverage, code complexity, comments, and software design and architecture. The functionality and capabilities of SonarQube can be extended with new modules, namely *plugins*. This allows to integrate support for additional programming languages, additional metrics, or the integration with other tools that bring in new functionalities.

The underlying logic of SonarQube is based on a *source code analyzer* component that performs basic analysis activities (such as counting lines of code) and an *application server* that graphically displays the data that results from the analysis in a browser front-end. The following elements are key to the internal function of the SonarQube framework:

- *Widgets* are the components that configure the graphical display of data resulting from the source code analysis, i.e., it enables the customization of the analysis results presentation to the user. A widget supports the specification of the visual format and display locations of the presented data. Each widget yields one of the square boxes that are shown. Each box contains a number of data items whose display location and characteristics is indicated in the widget code. For each new analysed project, SonarQube creates a *project dashboard* for selecting, adding, or removing the available widgets.

- *Sensors* are components that access the specific source code to be checked and that support the implementation of analysis functions over the source code to extract the metrics.
- *Decorators* are the components that support the programming of additional processing over the initial metrics provided by sensors in order to derive more complex metrics.

The framework comprises a key component, Sonar Runner, that controls the sequence of steps to launch the source code analysis. In order to execute a SonarQube analysis, it is needed to initially launch *Sonar Runner* that is determines the sequence of invocations of the functions provided by the *Sensors* and *Decorators*. Once the process is completed, the *runner* stores in a database all the collected data.

4 Metrics Integration in the Quality Framework

This section presents the design and implementation of the integrated metrics analyser that provides the enhanced functionality for SonarQube to integrate its own analysis results with the ones from external analysis tools in a single presentation space. The enhancement has been done via a plugin that integrates the external data and displays the integrated metrics in the project dashboard. The design is flexible and modular in order to support the integration of any external tool with minor modifications.

4.1 Addition of Metrics

SonarQube provides basic analysis facilities, yielding very basic metrics over the code. In software projects for critical systems, the specific standards that must be applied require more complex metrics over the code. A few examples of these are: complexity, nesting levels, function parameters, and other values derived from the previous ones such as maximum, minimum and mean values for each of the previous metrics. These are not provided by SonarQube. However, there are other specific external tools that do provide a broad range of complex metrics.

With the enhancement of SonarQube quality management platform, users view richer information over a software project code by using SonarQube as the single front-end, presenting a number of metrics in the project dashboard, as an additional widget.

Following, the structure of a sensor is provided. The `Sensor` interface is a tagging components, i.e., a class extending this interface is automatically a sensor as it is obligued to implement the `analyse` method to provide a customized functionality for the sensor.

```
public interface Sensor extends BatchExtension, CheckProject {
    void analyse(Project module, SensorContext context);
}
```

The design of the integration plugin has to allocate modules to provide the logic for: (*i*) storage of the analysis results from the external tool (the external analysis data); (*ii*) use and extend a sensor template to locate and acces the external analysis data; (*iii*) overwrite the `analyse` method to scan and collect the external analysis data; and optionally (*iv*) design and implement a *decorator* that computes additional metrics from the data collected by the sensor.

4.2 Software Design

The architecture of the analyzer software module is explained below, containing the following clases:

- *ExternalToolPlugin* is the class containing the specification of the properties of the analyser module. An arbitrary number of properties can be specified. For the integrated metrics analyser, `sonar.externaltool.metrics` is the basic property to specify the path to the external analysis results. Other possible properties are programming language and language.
- *ListMetrics* is the class that specifies all metrics to be used (displayed) by the plugin. Metrics should be specified by name, type, description, qualitative or quantitative and domain. Sensors later assign values to each metric of the list as a result of the source code analysis done by SonarQube or by some other external tool. Here, this class should contain all metrics provided by the external tool, that are precisely the external analysis results.
- *ExternalToolMetricSensor* class contains the functionality to scan the analysis data produced by the external analysis tool in order to collect the metrics that it provides. This class is invoked when the *Runner* component is executed.
- *MetricsRubyWidget* class contains the definition of the properties of the widget, the title of the widget, and the the design and display characteristics of the data to be included in the project dashboard. Precisely, the file containing the information about the design and display characteristics (`html.erb`) that contains the template for such a design and the positioning of the data.

5 Implementation Details for Validation

For the validation, a real project developed under norms [46] and [2] was analysed with our software module. Understand was selected as the external tool to validate the analyser module that integrates different sources of analysis results.

The first step to integrate Understand analysis into the SonarQube framework was to develop an extension of the Understand metrics to provide the required files for SonarQube.

Input files were provided by the external tool and contained all the metrics that the integrated analyser module is able to detect. The text marked as *Free text to include comments* will not be analysed by the module.

The class *ExternalToolMetricSensor* of the integrated analyser module is extended to derive the class *UnderstandMetricSensor* that supports the specific characteristics of this specific external analysis tool. Consequently, when the *Runner* component is executed, this class reads the required output file from Understand to derive its analysis metrics.

For the class *MetricsRubyWidget* that defines the properties to customize the widget and the data display, the specific data for Understand is given:

- *getId()* returns an identifier to the external tool Understand that is *UnderstandMetrics*
- *getTitle()* returns *Understand Metrics*.
- *getTemplatePath()* that definesto the file *html.erb*.

6 Conclusions

The paper has presented the design and implementation of a modular integrated metrics analyser for the SonarQube framework. Its execution inside the SonarQube framework results in the integration and connection of both tools that improves their capabilities, yielding a single colloborative remote working space that supports the interaction of verification teams working over specific projects.

The module is applicable to all projects with the need of technical metrics collection where external tools are mandatory to collect some metrics not provided initially by the SonarQube platform, used as quality management plattform; and it has been tried in a real project that requires compliance with norms related to the development of critical software such as DO-178C [46] and software quality such as AQAP-2210 [2].

Acknowledgment This work has been partly supported by the Spanish national project REM4VSS (TIN 2011-28339) and the Technology and Product Management department of Indra (Spain) under contract no. 2004/00476/001.

References

1. Alonso, A., García-Valls, M., de la Puente, J.A.: Assessment of timing properties of family products. In: ARES Workshop – Development and Evolution of Software Architectures for Product Families. LNCS, vol. 1429, pp. 161–169. Springer (1998)
2. AQAP 2210. NATO Supplementary Software Quality Assurance Requirements to AQAP 2110, 1st edn., November 2006

3. AQAP 2110. NATO Quality Assurance Requirements for Design, Development and Production, 2nd edn., November 2006
4. Balachandran, V.: Reducing human effort and improving quality in peer code reviews using automatic static analysis and reviewer recommendation. In: Proc. of International Conference on Software Engineering (ICSE) (2013)
5. Bersani, M.M., García-Valls, M.: The cost of formal verification in adaptive CPS. an example of a virtualized server node. In: Proc. of 17th IEEE High Assurance Systems Engineering Symposium (HASE), January 2016
6. Bouyssounouse, B., et al.: Programming languages and real-time systems. In: Embedded Systems Design: The ARTIST Roadmap for Research and Development. Springer (2005)
7. Bouyssounouse, B., et al.: QoS management. In: Embedded Systems Design: The ARTIST Roadmap for Research and Development. Springer (2005)
8. Bouyssounouse, B., et al.: Adaptive real-time systems development. In: Embedded Systems Design: The ARTIST Roadmap for Research and Development. Springer (2005)
9. CMMI Product Team. CMMI for Development, version 1.3. Improving processes for developing better products and services. CMU/SEI-2010-TR-033 (2010)
10. Campbell, G.A., Papapetrou, P.P.: SonarQube in Action. Manning Publications (2013). ISBN-9781617290954
11. Cano Romero, J., García-Valls, M.: Scheduling component replacement for timely execution in dynamic systems. Software: Practice and Experience **44**(8), 889–910 (2014)
12. CENELEC. Railway applications - Communications, signalling and processing systems. CENELEC (2001)
13. Coleman, D., Ash, D., Lowther, B., Oman, P.: Using metrics to evaluate software system maintainability. IEEE Computer **27**(8), 44–49 (2002)
14. di Ruscio, D., Pelliccione, P.: A model-driven approach to detect faults in FOSS systems. Journal of Software: Evolution and Process **27**(4), April 2015
15. Duenas, J., Alonso, A., Lopes Oliveira, W., Garcia, M., Leon, G.: Software architecture assessment. In: Software Architecture for Product Families: Principles and Practice. Addison-Wesley (2000)
16. Duvall, P.M., Matyas, S., Glover, A.: Continuous integration: improving software quality and reducing risk. Pearson Education (2007)
17. Fenton, N., Bieman, J.: Software metrics: a rigorous and practical approach. CRC Press (2014)
18. García Valls, M., Cucinotta, T., Lu, C.: Challenges in real-time virtualization and predictable cloud computing. Journal of Systems Architecture **60**(9), 736–740 (2014)
19. García Valls, M., Baldoni, R.: Adaptive middleware design for CPS: considerations on the OS, resource managers, and the network run-time. In: Proc. 14th Workshop on Adaptive and Reflective Middleware (ARM) (2015)
20. García-Valls, M., Fernández Villar, L., Rodríguez López, I.: iLAND: An enhanced middleware for real-time reconfiguration of service oriented distributed real-time systems. IEEE Transactions on Industrial Informatics **9**(1), February 2013
21. García-Valls, M., Alonso, A., de la Puente, J.A.: A Dual-Band Priority Assignment Algorithm for QoS Resource Management. Future Generation Computer Systems **28**(6), 902–912 (2012)
22. García-Valls, M., Alonso, A., de la Puente, J.A.: Mode change protocols for predictable contract-based resource management in embedded multimedia systems. In: Proc. of IEEE Int'l Conference on Embedded Software and Systems (ICESS), May 2009
23. García-Valls, M., Alonso Munoz, A., Ruíz, J., Groba, A.: An architecture of a quality of service resource manager middleware for flexible multimedia embedded systems. In: Proc. of 3rd Intern'l Workshop on Software Engineering and Middleware. LNCS, vol. 2596 (2003)
24. García-Valls, M., Perez-Palacin, D., Mirandola, R.: Time sensitive adaptation in CPS through run-time configuration generation and verification. In: Proc. of 38th IEEE Annual Computer Software and Applications Conference (COMPSAC), pp. 332–337, July 2014
25. García-Valls, M., Basanta-Val, P., Estévez-Ayres, I.: Real-time reconfiguration in multimedia embedded systems. IEEE Transactions on Embedded Consumer Electronics **57**(3), 1280–1287 (2011)

26. García-Valls, M., Basanta-Val, P.: A real-time perspective of service composition: key concepts and some contributions. Journal of Systems Architecture **59**(10), 1414–1423 (2013)
27. García-Valls, M., Basanta-Val, P.: Comparative analysis of two different middleware approaches for reconfiguration of distributed real-time systems. Journal of Systems Architecture **60**(2), 221–233 (2014)
28. García-Valls, M., Uriol-Resuela, P., Ibánez-Vázquez, F., Basanta-Val, P.: Low complexity reconfiguration for real-time data-intensive service-oriented applications. Future Generation Computer Systems **37**, 191–200 (2014)
29. García-Valls, M., Basanta-Val, P.: Usage of DDS Data-Centric Middleware for Remote Monitoring and Control Laboratories. IEEE Transactions on Industrial Informatics **9**(1), 567–574 (2013)
30. García-Valls, M., Basanta-Val, P., Estévez-Ayres, I.: Adaptive real-time video transmission over DDS. In: Proc. of 8th IEEE International Conference on Industrial Informatics (INDIN), July 2010
31. IEC 61508. Functional safety of electrical/electronic/programmable electronic safety-related systems, April 2010
32. IEC. Nuclear power plants. Instrumentation and control important to safety. General requirements for systems. IEC 61513 Ed.2.0., August 25, 2011
33. IEC. Medical Device Software–IEC, May 2006
34. ISO. Road Vehicles - Functional Safety. ISO-26262, November 11, 2011
35. Jenkins. Information. http://jenkins-ci.org/ (last retrieved, February 19, 2015)
36. Krutchen, P.: Contextualizing agile software development. Journal of Software: Evolution and Process **25**, 351–361 (2013)
37. LDRA. Information. http://www.ldra.com/ (last retrieved, February 19, 2015)
38. Maven. Information. http://maven.apache.org/ (last retrieved, February 19, 2015)
39. PECAL-2210. Requisitos OTAN de aseguramiento de la Calidad del software, suplementarios a la PECAL 2110, 1st edn., November 2007
40. PECAL-2110. Requisitos OTAN de aseguramiento de la Calidad para el diseño, el desarrollo y la producción, 2nd edn., November 2006
41. PC-Lint. Information. http://www.gimpel.com/html/index.htm (last retrieved, February 19, 2015)
42. PMD. Information. http://pmd.sourceforge.net/ (last retrieved, February 19, 2015)
43. SonarQube. Information. http://www.sonarqube.org/ (last retrieved, May 04, 2015)
44. Splint. Information. http://www.splint.org/ (last retrieved, February 19, 2015)
45. Otero Pérez, C.M., Steffens, L., van der Stok, P., van Loo, S., Alonso, A., Ruíz, J., Bril, R.J., García Valls, M.: QoS-based resource management for ambient intelligence. In: Ambient Intelligence: Impact on Embedded Sytem Design, pp. 159–182. Kluwer Academic Publishers (2003)
46. RTCA Inc. Software Considerations in Airborne Systems and Equipment Certification. RTCA Inc. DO-178C, December 13, 2011
47. RTCA Inc. / EUROCAE. Software Integrity Assurance Considerations for Communication, Navigation, Surveillance and Air Traffic Management (CNS/ATM) Systems. DO-278A, December 13, 2011
48. RTCA Inc. DO-178B. Software Considerations in Airborne Systems and Equipment Certification. RTCA Inc. DO-178B (1992)
49. RTCA Inc. Software Tool Qualification Considerations. DO-330, December 13, 2011
50. RTCA Inc. Model-Based Development and Verification Supplement to DO-178C and DO-278A. DO-331, December 13, 2011
51. RTCA Inc. Object-Oriented Technology and Related Techniques Supplement to DO-178C and DO-278A. DO-332, December 13, 2011
52. RTCA Inc. Formal Methods Supplement to DO-178C and DO-278A
53. Scitools. Scitools Understand. Information. https://scitools.com/ (last retrieved, February 19, 2015)

Supporting the Development of User-Driven Service Composition Applications

Alex Roberto Guido, Antonio Francisco do Prado, Wanderley Lopes de Souza and Eduardo Gonçalves da Silva

Abstract One of the software engineering challenges is the development of applications that can adapt to the heterogeneous needs of users. Technical Dynamic Composition of Services Driven by User is a solution for developing applications capable of overcoming these challenges. This type of application which will call *User-Driven Service Composition Application (UDSCA)* allows to compose services during its execution, thus meeting the needs of users. But the lack of guidance on how to develop UDCAs can make development a complex task, because it may aggregate unknown solutions by developers, thus damaging the development team and the developed application. Therefore, the objective of this work is to be able to guide developers during the development of this kind of application. To develop the approach, it has been defined which activities should be undertaken during the development as well as the concepts, techniques, artifacts, technologies and tools needed to perform these activities. To evaluate the approach one conducted a case study in which a UDSCA was developed in the field of building maintenance services. The resulting application of the approach was shown to be able to adapt to the heterogeneous needs of the user, also the approach provided artifacts that promoted reuse. In conclusion, the approach guides the developer during the UDSCAs development and provides artifacts that reduce efforts for development.

A.R. Guido(✉) · A.F. do Prado · W.L. de Souza
Department of Computer Science,
Federal University of São Carlos (UFSCar), São Carlos, SP, Brazil
e-mail: {alex.guido,prado,desouza}@dc.ufscar.br

E.G. da Silva
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS),
University of Twente, Enschede, The Netherlands
e-mail: e.m.g.silva@utwente.nl

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_46

1 Introduction

With the evolution of mobile devices and the Internet becoming ubiquitous, it is increasingly common for people to use the help of the computer to perform daily tasks. One of the current challenges of software engineering is to produce applications that can adapt to the heterogeneous needs of users [3]. This heterogeneity is due to the different users preferences and running applications in different situations [14]. The use in different situations requires the application to be able to adapt to changes occurring in the environment, on the devices and because the needs may change depending on the user [4]. A situation that can illustrate this scenario is a trip planning. Due to the fact users have different preferences, this planning can be made in several ways, for example, while a person may be interested in visiting historic monuments, another person may be only interested in meeting people in the region; while a person is planning a trip using a stationary device, another may be using a mobile device.

This reality has led to the increasing need for applications that can adapt to meet the different needs of users [6]. In practice, when implementing this type of application, developers encountered some difficulties, such as: impossibility to develop the functionalities because these functionalities change as the users' needs [14]. So building applications to meet the different needs of users is a complex task. Aiming to overcome these difficulties and simplify the development of this type of application, many researches are being carried by the scientific community [12, 14, 15].

One solution that has been successfully shown is the reuse of services [9], extensively explored in *Service-Oriented Computing* (SOC). SOC allows the creation of applications quickly at low cost, because of combining their own services or third-party services in an interoperable way in heterogeneous environments [10]. Following this idea, using the technique of combining services (also known as service composition), it is an interesting solution to meet the need to develop applications that need to adapt to keep up with changes that may occur in the environment and the needs of users [1, 14]. This technique when performed dynamically and focused on user needs *(Dynamic Composition of Services Driven by User)* makes it possible for the development of User-Driven Service Composition Applications (UDSCA), that is, custom applications to meet the different needs of users. This is due to the high number of services and compositions provided by several providers, thus increasing the chances of finding services and compositions that meet the different users' preferences.

However, the lack of guidance on how to develop UDSCAs can make development a complex task, because it may aggregate unknown solutions by the developers, thus damaging the development team and the developed application. Therefore, the aim of this work is to present an approach to guide developers during the development of such applications and provide artifacts that promote reuse, thus reducing development efforts.

2 Background

In this section two important topics are presented so that there is a better understanding of the proposed approach: Framework A-DynamiCoS and HATEOAS.

2.1 Framework A-DynamiCoS

A Dynamic Composition of Services Driven by the user aims to provide on-demand services compositions for heterogeneous users, that is, the composition should be performed when using the application, from different devices in order to meet users with different needs and contexts [14].

The Framework A-DynamiCoS [14] provides support for this type of composition. Your goal is to perform dynamically the composition services, to deliver services that meet the different needs of users.

To have the information about the available services to be used in the composition process, A-DynamiCoS accesses a repository of services described semantically in SPATEL [2]. A-DynamiCoS is provided as a Web service, enabling its interaction with different platforms. This interaction occurs via predefined commands, each command encapsulates a request for a particular behavior that should be delivered by A-DynamiCoS. The commands provided by the A-DynamiCoS are *ServType-Discovery, SeleServ, ValidateInputs, ResolveServ, ExecServ e AddToContext*. These commands are physically combined in order to meet the requirements of an application, the combination of these commands generates a flux which is denominated *Command Flow*. The purpose of the command flow is represented as the commands are related, that is, the order to use the commands.

2.2 HATEOAS

Hypermedia as the Engine of Application State (HATEOAS), is a restriction of REST architecture presented by Fielding [5]. The principle is that a client application does not need any prior knowledge of how to communicate with an application or server on the network, and should only make communication using hypermedia. An example of using this principle is the Web, when a Web page is requested, it is not necessary to have knowledge of the implemented logic on the server, it's just a request through the Hypertext Transfer Protocol (HTTP) sending the URL of the page and then the server returns content and the links with the next actions to be performed.

3 Approach

The approach is intended to guide the development of User-Driven Service Composition Applications (UDSCA), that is, applications that use the Dynamic Composition of Services Driven by User [14] to meet the heterogeneous needs of users. The approach defines which activities should be undertaken to develop this type of application and suggests tools to assist in these activities.

Several works to perform service composition have been reported in literature [13]. Since some of these works provide the responsible component for performing service composition, the need arises to insert it into the development of applications, so that you can reuse it in a simple and easy way, thus reducing additional development efforts. The resulting component of the work [14] named *A-DynamiCoS* is used in the presented approach. This component was chosen because it differentiates itself from others in the literature for not having a hard life cycle (that is, discover, compose, deploy and run exactly in that order). Because of this flexibility, *A-DynamiCoS* makes it feasible for the development of applications that require adaptation to meet the needs of the user.

An overview of the approach can be seen in the SADT(*Structured Analysis and Design Technique*) diagram [11] in Figure 1. The defined activities in the approach are divided into Domain and Application Engineering. In the Domain Engineering, activities are responsible for generating reusable artifacts from a domain, however in Application Engineering, activities are responsible for generating the final application, which makes reuse of the artifacts generated by Domain Engineering.

To make the understanding of the approach easier, its activities are presented through a case study developed in the field of building maintenance services. In this domain people are able to advertise and search for maintenance services in a especific

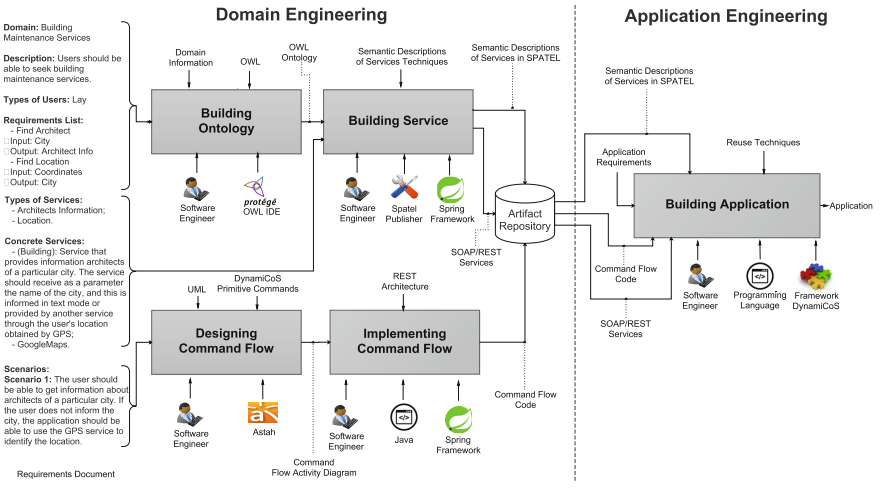


Fig. 1 Overview of the Approach.

urban area. Followed by the presentation of the approach beginning with Ontology Building activity.

3.1 Building Ontology

In this activity the software engineer creates the ontologies that are used as a basis to describe and seek the services in a semantic way. To support the creation of these ontologies, is provided in the beggining of the activity, one *Requirements Document*[1] containing information about the domain and application requirements, it is suggested that these ontologies are created using OWL [8] because it is a very promising language and also to be a pattern adopted by W3C[2]. The following ontologies should be created:

- Ontology Goals: this ontology must contain objectives that can be achieved within a domain. These goals will be used to describe a feature that will provide a particular service, as well as the objectives that a particular user may want to achieve when using the application within the domain.
- Ontology parameters: this ontology will contain the input/output parameters of the available services in an domain.

There are several tools for creating ontologies in OWL, it is suggested to use the *protégé IDE*[3], because of its simplicity of use.

The results of this activity are the ontologies in OWL for the domain where the application is used. An example of these ontologies can be seen in Figure 2, which were created for the domain of building maintenance services.

In Figure 2(a) it can be seen that the ontology contains the objectives provided in the domain of building maintenance services such as *FindArchitect*. To achieve this goal (*FindArchitect*) some inputs must be provided (ex: *Coordinates*), which results in output (*InformationAboutArchitect*) after application processing developed to support this domain. These inputs and outputs can be observed in the ontology of Figure 2(b).

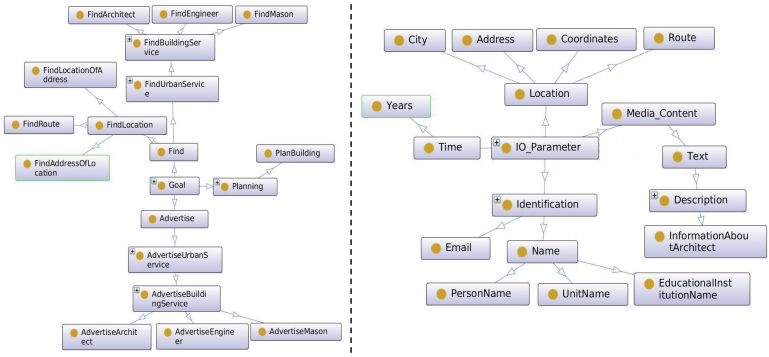
3.2 Building Services

In this activity the services that will be used by the application should be built/reused, described semantically and published in a repository by the software engineer. First, one must seek existing services, in order to reuse them and meet the functionalities requested in the requirement's documents. If these services are not found, it must be built based on the information available in the requirements document. For this

[1] Part of the definition of this requirements document is based on the work presented by Vlieg [15]

[2] http://www.w3.org/ accessed on 01/10/2015

[3] http://protege.stanford.edu/ accessed on 17/09/2015



(a) Ontology Goals (b) Ontology Inputs / Outputs

Fig. 2 Fragments of the Ontologies Goals and Parameters.

construction it is suggested to use the *Spring Framework*[4], because it supports the construction of SOAP/REST services and also for its simplicity of use.

Later, one must describe semantically the services built/reused. To make this semantic description, one should use a language that allows to write as follows: input parameters, output, preconditions, effects, service goals and non-functional properties, and these notes should be grounded in the ontologies provided as input to the activity. SPATEL [2] is a language that enables you to note all of these operations, because of this it is suggested that this language to be used.

A semantically described service example in SPATEL can be seen in the Code Snippet 1.1. This service aims *FindArchitect*, so the semantic note of service objectives can be observed in the attribute `name="OperationGoal"` and `semType="Goals.owl:FindArchitect"` (Lines 15-16). The semantic notes refers to the input parameters and service output that can be observed in the tags `<ownedParameter />` (Lines 5-13). As can be observed by the values attribute `semType=""` these are fundamental in the ontologies' objectives and parameters created in *Building Ontology* activity.

```
1 <spatel:ServiceLibrary>
2   <service name="FindArchitect">
3     <ownedOperation name="FindArchitect">
4
5       <!-- Input -->
6       <ownedParameter name="name" semType="IOTypes.owl:PersonName" />
7       <ownedParameter name="city" semType="IOTypes.owl:City" />
8       <ownedParameter name="expTime" semType="IOTypes.owl:Years" />
9       <ownedParameter name="coord" semType="IOTypes.owl:Coordinates" />
10
11      <!-- Output -->
12      <ownedParameter name="return" semType="IOTypes.owl:InformationAboutArchitect" />
13
14      <semTag xmi:id="id" name="OperationGoal" semType="Goals.owl:FindArchitect" />
15
16    </ownedOperation>
17  </service>
18 </spatel:ServiceLibrary>
```

Code Snippet 1.1 Service Semantically Described in SPATEL

[4] http://projects.spring.io/spring-framework/ accessed on 02/10/2015

After finishing the semantic description of services, these descriptions should be published in a service repository *Universal Description, Discovery and Integration (UDDI)*. To accomplish this publication it is suggested to reuse the project *SPATEL Publisher*[5]. SPATEL Publisher is a project based on Java API and facilitates the import of semantic descriptions for a repository services. To make the publication of semantic descriptions through SPATEL Publisher it must provide the following parameters:

- Description: A brief description in natural language about the service;
- URL of the service: the service address on the network, either an internet or intranet network;
- URL Semantic description: File address created to describe semantically the service, for example, the file in SPATEL format.

Because of the need for a repository that allows the semantic descriptions publishing of services, it is suggested to reuse the extension registry jUDDI[5] provided by Silva [7]. This extension includes the operations necessary to write down the services semantically, thus making it an attractive option.

One of this activity's results relates to services that will be reused for the final application. These services will be stored on a server that allows its reuse by applications over the internet or intranet.

Another result of this activity is the semantic descriptions of published services in a repository. These descriptions will be required for the A-DynamiCoS Framework to be able to find relevant services to meet the needs of users, since this user will use a particular application to achieve their goals within a domain.

3.3 Designing Command Flow

In this activity the control flow is modeled. This command flow will be responsible for defining the application's behavior, that is, the call order of the primitive commands provided by the A-DynamiCoS Framework. The definition of this command flow should be based on the *requirements document* provided as input to the activity, since the created command flow must be able to cover all scenarios specified in this document.

The command flow must have activities (primitive commands) and the call order of these activities, and thus the activity diagram of Unified Modeling Language (UML)[6] suggestion is to perform modeling of this command flow. It is also suggested to assist with modeling, the use of *Astah Tool*[7], because this tool gives the software engineer a graphical interface, thus enabling the activity diagram being created to be displayed in a practical and simple way.

The resulting artifact of this activity is the Activities Diagram of Command Flow.

---

[5] http://sourceforge.net/projects/dynamicos/files/ accessed on 17/09/2015

[6] http://www.uml.org/ accessed on 24/09/2015

[7] http://astah.net/ accessed on 24/09/2015

3.4 Implementing Command Flow

In this activity the command flow is encoded and supplied as a RESTful API. This encoding is as follows: (1) the software engineer analyzes the activity diagram provided as input to the activity, in order to identify the primitive commands (Diagram activities) provided by the Framework A-DynamiCoS will be reused; (2) for each identified primitive command the necessary code is generated so that there is re-use, the approach opted for the use of the Java language for encoding; (3) the generated code for reuse for each primitive command is exposed as a REST resource, it is suggested to use the Spring Framework due to the fact that it provides ways to expose Java code as REST resources in a simple and easy way; (4) the software engineer again analyzes the activity diagram to identify what the invocation order of primitive commands (Diagram activities); (5) based the on information obtained in the previous step, the responses of each resource must be added, from which the next feature can be called. It is suggested to use Spring Framework HATEOAS[8] to make this step, because devising mechanisms that allow you add the addresses to the responses to of resources the next few resources that can be invoked on the HATEOAS principle's; and (6) that resources should be available through a *Uniform Resource Identifier (URI)* so they can be reused by applications on different platforms, this reuse will occur through the set of operations provided by the HTTP protocol.

The resulting artifact of this activity is the command flow of the code provided as a RESTful API. The main goal of this artifact is to promote communication between the Framework A-DynamiCoS and application that is reusing the artifact. The reuse of this artifact is an independent platform, thus the developer does not need to develop a communication device for each platform, reducing development workload.

A slice of this artifact can be seen in the Code Snippet 1.2, which refers to the reuse of the original command ServTypeDiscovery. The code to reuse this primitive command is exposed as a REST resource (Line 5) and asks as an argument `serviceType`. The responsible code for communication with the A-Dynamicos Framework is mounted `ServTypeDiscovery` (Line 9) and interaction with A-DynamiCoS occurs through `DynamicosIntegration` class (Line 11). The services found at A-DynamiCoS are returned through `getServices()` method and added to the `ServTypeDiscovery` class through `addServiceList()` method (Line 13). Finally, added in the resource reply, is the next feature that can be called (Line 23) and then the resource response is returned (Line 25).

```
1 @RestController
2 @RequestMapping("/command_flow")
3 public class CommandFlowController {
4
5     @RequestMapping("/serv_type_discovery")
6     @ResponseBody
7     public HttpEntity<ServTypeDiscovery> servTypeDiscovery(@RequestParam(value = "
          servicetype", required = true) String serviceType) {
8
9         ServTypeDiscovery servTypeDiscovery = new ServTypeDiscovery(serviceType);
10
11        DynamicosIntegration dynamicos = new DynamicosIntegration(servTypeDiscovery);
12
13        servTypeDiscovery.addServiceList(dynamicos.getServices());
```

[8] http://projects.spring.io/spring-hateoas/ accessed on 10/10/2015

```
14
15          if(servTypeDiscovery.getServiceList().isEmpty()){
16
17              servTypeDiscovery.add(linkTo(methodOn(CommandFlowController.class).
                    servTypeDiscovery()).withSelfRel());
18
19              return new ResponseEntity<ServTypeDiscovery>(servTypeDiscovery, HttpStatus.
                    NOT_FOUND);
20
21          }else{
22
23              servTypeDiscovery.add(linkTo(methodOn(CommandFlowController.class).seleServ())
                    .withSelfRel());
24
25              return new ResponseEntity<ServTypeDiscovery>(servTypeDiscovery, HttpStatus.OK)
                    ;
26          }
27      }
28
29      @RequestMapping("/sele_serv")
30      @ResponseBody
31      public HttpEntity<SeleServ> seleServ(@RequestParam(value = "service", required = true)
            String service) {}
32
33  }
```

Code Snippet 1.2 Command Flow Code

A flow command response code example is shown in the Code Snippet 1.3. In this example, the response refers to `servTypeDiscovery` request to the resource `servTypeDiscovery`, which was performed through HTTP GET operation on the parameter `serviceType` set as FindArchitect. The service found for the defined parameter is located in the content of the response (Line 2-5) together with the next features that can be invoked (Line 6-8).

```
1 {
2   "service": [ {
3         "service_id": 1,
4         "name": "FindArchitect",
5         "goal": "FindArchitect",
6         "links": [ {
7             "rel": "sele_serv",
8             "href": "http://localhost:8080/sele_serv"
9         } ]
10      }
11 }
```

Code Snippet 1.3 Response Resource servTypeDiscovery

3.5 Building Application

This activity application is developed to be used by the end user. To develop this application it is suggested to do the following activities: reuse artifacts provided by the domain engineering, look for reducing efforts in the development; develop the graphical interface to interact with the user; a specific application code.

Since the artifacts for reuse are available as RESTful API, to reuse them, simply access them via HTTP operations such as: POST, GET, PUT and DELETE. The content returned by the artifacts should be handled under the application developer's responsibility and deciding which procedures to accomplish.

The graphical development for user interface interaction and specific code (for example, user control) can be developed on any platform, at the discretion of the development team to decide on the most suitable for achieving their goals. This flexibility

occurs because of artifacts that will provide the content for the application being developed such web services, making them accessible via any platform.

The result of this activity is the final application that can be used by the domain users for which it was developed.

4 Case Study

In this section it is presented the case study to evaluate the applicability of the proposed approach. The study case consisted in the development of a UDSCA domain of building maintenance services with the UDSCA announcing, finding and planning property maintenance services in a given urban area. To achieve these goals, the application uses 25 Web Services built using REST / SOAP, and 3 Web Services responsible for providing information of localization.

The purpose of a UDSCA is to meet the different needs of users, so UDSCA development combines these Web Services during its execution, In order to adapt to the user who is using it on that moment. For example, when advertising or searching one building maintenance service you may need help when it comes to its location, so the Web Services responsible for advertising and seeking building maintenance services can be composed with Web Services responsible for providing information of location to assist users. The search and composition of this service are performed by the Framework A-DynamiCoS.

The development was performed in two UDSCA ways: with and without the help of the proposed approach, in order to compare and obtain the significant contributions provided by the approach.

The use of the approach shows a reduce of the development time of the application in approximately 52.39% (with approach: 240 hours, without approach: 504 hours), this is due to fact that developers use the available suggestions in the approach, not just spending time researching which activities should be undertaken to develop the application, and what tools and technologies are needed to support the activities. The development time was also reduced because the approach is providing reusable artifacts, thereby reducing the amount of developed code. The fact that these reusable artifacts are independent of platform, also reduces workload for developing when the application needs to be developed in different platforms. The results are considered significant, proving the applicability the proposed approach.

5 Related Work

Vlieg [15] presents different patterns of use for User-Driven Service Composition and develops a case study in the entertainment domain. But does not show which activities and tools are needed to develop the application to support this area.

Santos [12] developed a User-Driven Service Composition Application to support the public services domain. But also does not show how the application was developed and the required artifacts to develop the application.

6 Conclusion

This paper presents an approach to support the development of User-Driven Service Composition applications, that is, applications that can adapt to meet the different needs of users. The approach defines which activities should be undertaken to develop this type of application, as well as the concepts, techniques, devices, technologies and tools needed to perform these activities.

In order to verify the applicability of the proposed approach, an application was developed in the field of building maintenance services. The development of this application was based on the activities defined in the approach, which was enough for the development of the application, thus showing the applicability of the approach.

Another approach contribution refers to reusable artifacts produced by Domain Engineering. These artifacts reduced effort on the development because developer does not need to develop all the source code, and only complement it.

For future work tools will be developed for the automation of some activities of the approach, which will promote the automatic generation of code in order to also reduce development workload. Furthermore, the case study will be replicated in other fields and for different languages and platforms in order to compare results and verify the effectiveness of the approach.

References

1. A dynamic service composition model for adaptive systems in mobile computing environments. In: Franch, X., Ghose, A., Lewis, G., Bhiri, S. (eds.) Service-Oriented Computing. LNCS, vol. 8831 (2014)
2. Almeida, J.P., Baravaglio, A., Belaunde, M., Falcarin, P., Kovacs, E.: Service creation in the spice service platform. In: Proceedings of the 17th Wireless World Research Forum Meeting (WWRF17), pp. 1–7 (2006)
3. Bertolino, A., Blake, M., Mehra, P., Mei, H., Xie, T.: Software engineering for internet computing: Internetware and beyond. Software, IEEE **32**(1), 35–37 (2015). [guest editors' introduction]
4. Cirilo, C., do Prado, A., de Souza, W., Zaina, L.: Model driven richubi - a model-driven process to construct rich interfaces for context-sensitive ubiquitous applications. In: 2010 Brazilian Symposium on Software Engineering (SBES), pp. 100–109, September 2010
5. Fielding, R.T.: Architectural styles and the design of network-based software architectures. Ph.D. thesis (2000)
6. Gartner: Highlights Key Predictions for IT Organisations and Users in 2010 and Beyond. Tech. rep. (2010)
7. da Silva, E.M.G.: User-centric Service Composition - Towards Personalised Service Composition and Delivery. Ph.D. thesis. University of Twente, Enschede, May 2011

8. Smith, M.K., McGuiness, R.V.D., Welty, C.: Web ontology language (owl) guide. Tech. rep. (2002)
9. Oreizy, P., Medvidovic, N., Taylor, R.N.: Runtime software adaptation: Framework, approaches, and styles. In: Companion of the 30th International Conference on Software Engineering, pp. 899–910. ICSE Companion (2008)
10. Papazoglou, M.P., Traverso, P., Dustdar, S., Leymann, F.: Service-oriented computing: A research roadmap. International Journal of Cooperative Information Systems **17**(02), 223–255 (2008)
11. Ross, D.: Structured Analysis (SA): A Language for Communicating Ideas. Software Engineering, IEEE Transactions on SE **3**(1), 16–34 (1977)
12. dos Santos, J.H.: Public service improvement using runtime service composition strategies (2010). http://essay.utwente.nl/59920/
13. Sheng, Q.Z., Qiao, X., Vasilakos, A.V., Szabo, C., Bourne, S., Xu, X.: Web services composition: A decade's overview. Information Sciences, 218–238
14. da Silva, E., Pires, L., van Sinderen, M.: A-DynamiCoS: A Flexible Framework for User-centric Service Composition. In: 2012 IEEE 16th International Enterprise Distributed Object Computing Conference (EDOC), pp. 81–92, September 2012
15. Vliegs, E.: Usage patterns for user-centric service composition, November 2010. http://essay.utwente.nl/59753/

Mining Source Code Clones in a Corporate Environment

Jose J. Torres, Methanias C. Junior and Francisco R. Santos

Abstract Many researches around code clone detection rely on Open Source Software (OSS) repositories to execute their studies. These cases do not reflect the corporative code development scenario. Big Companies repositories' are protected from the public's access, so their content and behavior remain as a black box on the researchers' viewpoint. This article presents an experiment performed on systems developed in a large private education company, to observe and compare the incidence of cloned code on proprietary software with other studies involving open source systems, using different similarity thresholds. The results indicate that the closed-source repository presents similar clone incidence as the OSS ones.

Keywords Proprietary software · Mining software repositories · Clones · Experimental software engineering · Closed-source projects

1 Introduction

The demand for speeding up software development allied to the lack of patterns and the inexistence of internal policies to implement best practices triggers a series of issues related to coding organization. Software development teams have to achieve business deadlines, so they adopt the bad practice to copy-and-paste code. In this way, clones populate software repositories and hinder the improvement or maintenance of systems. The most part of legacy systems code is the result of reusing existing code, so, developers who want to implement a new feature find

J.J. Torres(✉) · M.C. Junior
Federal University of Sergipe – UFS, Sao Cristovao, Sergipe, Brazil
e-mail: jorgesamango@gmail.com, mjrse@hotmail.com

F.R. Santos
Federal Institute of Sergipe – IFS, Lagarto, Sergipe, Brazil
e-mail: frchico@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_47

some code snippet similar to the desired one then make a copy and modify it. A clone also result from identical instructions that works only with different data types – this indicate the failure to use Abstract Data Types. [1].

Construction of device drivers also generates many similarities between codes, as much of this type of program geared to the same platform is virtually identical [2]. Moreover, another reason for the existence of clones is called "reinventing the wheel", because some developers do not bother to look if there is already a piece of code for something that was requested to the team [3].

The copy-and-paste way may cause a serious maintenance problem, for example, in case of bugs. If a bug was found in a piece of code that has been cloned in several other pieces, all of these clones should be corrected too [4].

Most studies around clone code theme make use of the same concepts. For example, the main types of code clones are [8]: Exact clones or program fragments identical to each other; Parameterized clones, are fragments with the same structure except for changes in data types, identifiers, layout and comments; Near-miss clones, program fragments copied with a few modifications inside; Semantic clones, blocks of code textually different but producing a same computation.

Other authors bring some terminologies concerning the relationships between clones [22]. A Clone Pair is a pair of code fragments identical or similar to each other. A Clone Class is a set of code fragments in which any two of the members can form a clone pair. In short, a clone class is the union of all clone pairs who shares code fragments in common. Clone Family, also known as Super Clone, is the group of all clone classes belonging to the same domain.

Despite code clones are considered harmful [15], for all the reasons we presented earlier, in some cases they may be a great deal. Introducing a new feature inside existing software can be eased by replicating the code and making the modifications. When the modified version of the code is tested in a sandbox or something similar, it can be applied in the production environment. This way minimizes the risk of instabilities in the stable version.

Some studies suggest that code clones may be avoided by adopting good design techniques and development methodologies, including refactoring on the development process [22]. Many efforts [9,11,12,13,14,16] show that code refactoring as part of the package of a clone detection tool may be a desirable feature in some situations. Thus, considering a corporate environment with a well-defined software process, this paper aims address the following research question: "*Have corporate software environments lower clone incidence than Open Source Systems?* To answer, our experimental evaluation analyzed large-scale Closed-Source Systems and compared with OSS Systems published. The results showed similar numbers for both.

The remainder of this paper is structured as follows: Section 2 discusses related work. Section 3 is dedicated to the understanding of the main tool used in our experiment. Section 4 presents the experiment planning and definition. Section 5 describes the experiment execution among with the environment used to explore the clone code detection. Section 6 describes, analyzes and discusses the validity of the obtained results. Finally, Section 7 contains conclusions and final remarks.

2 Related Works

As this work focuses on clone incidence inside software repositories, this section presents studies about this subject. The main peculiarity of these articles regarding our work is that they were performed inside Open Source Software (OSS) environments.

Many works [4,10,11] were concerned about evaluating source code mining techniques and tools, identifying their strengths and weakness. Khatoon, Mahmood and Li [4] try to extract positive and negative aspects from cloned detection tools and techniques to help future researchers and developers.

Schwarz, Lungu and Robbes [7] focused on large code bases, combining three lightweight clone detection techniques to evaluate performance on a real-world ecosystem. The techniques are directed to three types of clones. The type 1 are Hashes of Source Code. Type 2 are defined as Hashes of Source Code With Renames. A clone is considered a type-2 if it is a type-1 even after every sequence of alphabetical letter be replaced by the letter "t" and all sequence of digits replaced by number 1. Type 3 or "Shingles", are defined as a consecutive sequence of tokens in a document, after the transformations defined by rules of type-2 clones.

Roy and Cordy [5] motivate our work. Like them, we run the NICAD tool in a software repository with the difference that is a protected repository belonging to a private corporation. Their study was about finding function clones inside C and Java systems code repository, with projects varying in size from 4K LOC to 6265K LOC. All non-empty functions with a minimum of 3 LOC were considered, that includes the function header with opening and ending bracket and at least one code line. They used similarity thresholds, also known as Unique Percentage of Items (UPI) thresholds to present their results. The validation of NICAD results was done by hand and using Linux diff tool to check the textual similarities. The same researchers provided in another work [9] a description of commonly used terms, review of existing clone taxonomies, detection approaches and experimental evaluations of clone detection tools. At last, a list of some problems related to clone detection for future research is presented and discussed.

Clone Mining research needs substantial infrastructure support, particularly with respect to adopting a standard experimental process, described in some Mining Software papers [18] and in this paper, with the goal of effectively replicating clone studies. The barrier and cost for experimentation with Clones Mining are considerably low compared to other software engineering techniques (e.g., on-line experiments with participants). In other words, research projects and papers can conceive an experience factory and demonstrate true value of this area for practitioners.

3 NICAD Tool

NICAD Clone Detector [6] is an Open Source implementation of the NICAD method. It supports five languages, C, C#, Java, Python and WSDL, thru two granularity levels: blocks or functions. It is possible extend language support by adding a new TXL parser or extractor for the new language or granularity.

This is possible though plugin architecture that also allows to add custom normalization templates. The TXL [16] is a hybrid functional/rule-base programming language, designed to support computer software analysis and source transformation tasks.

The NICAD method involves three stages [6]. The first one called Parsing, is about extract fragments of a given granularity, like functions or blocks to export them to a pretty-printed textual form, normalizing spacing and line breaks and removing comments to expose identical clones as textually identical fragments. Normalization is the second stage that concerns to normalize, filter or abstract the extracted fragments before comparison. They can be renamed to adopt some standard or have some declarations removed. Lastly, comparison concerns to compare each fragment line using a LCS (longest common subsequence) algorithm to detect the clones. It is possible to parametrize NICAD with four different similarity thresholds, from exact to near-miss clones.

4 Experiment

Our work is presented here as an experimental process. It follows the guidelines by Wohlin et al. in [17]. In this section, we start introducing the experiment definition and planning. The following sections, will direct to the experiment execution and data analysis.

4.1 Goal Definition

Our goal is to compare clone findings of a private source code repository with another work that evaluates an Open Source software repository, using the same similarity thresholds.

To achieve this, we are going to execute an experiment in a controlled environment using the same tools, concepts and metrics than our main related work. This comparison test attempts to answer questions about clone incidence related to code freedom paradigms.

The goal is formalized using the GQM Goal template proposed by Basili and presented in [20]:

- **Analyze** our corporate projects
- **with the purpose of** evaluation (against published OSS projects)
- **with respect to** code clone manifestation
- **from the point of view of** the programmers
- **in the context of** software repositories

4.2 Planning

Context Selection. The experiment will be off-line and executed with the NICAD clone detector inside a Java code repository containing about seven different

systems of a corporate environment. The selected subject organization is an educational-purpose company active in market since the 60s, with more than 2,000 employees and around 50,000 customers.

Hypothesis Formulation. The research question for this experiment is: have corporate software environments lower clone incidence than Open Source Systems?

We will compare some extracted statistics of Java systems from open-source projects reported in Roy and Cordy [5] work with seven other private systems from our target corporation using the same extraction tool, respecting similarity thresholds between comparisons. To assure the reliability of our hypothesis test, we will calculate the average between the proportional results of exact similarity for each system (*S)*, where UPI threshold is 0.0. The proportion (*P)* is calculated by dividing Clone Pairs or Clone Classes Findings *(C)* by its respective KLOC.

When defining the variables for the formal test, the systems size was considered, because just the clone numbers does not imply conditions to evaluate a greater propensity to lower abstraction. Besides, the UPI threshold as 0.0 indicates an identical clone, evidencing more reliably the possibility of a type of Technical Debt (DT) [19] such as failure to code reuse or failure to use Abstract Data Types (ADT). Capture of Clone Classes were included in our experiment in order to identify repositories storing methods that are cloned in excess. Keeping this idea, we will try to confirm the following hypothesis:

HYPOTHESIS 1.

- **Null hypothesis H_0^{CP}:** Source Code Repositories in the context of our corporate projects (1) have same incidence of clone pairs of the Open Source Projects (2) reported in the literature.
  - H_0^{CP}: $\mu 1_{(\text{Clone Pairs Proportion})} = \mu 2_{(\text{Clone Pairs Proportion})}$
- **Alternative hypothesis H_1^{CP}:** Source Code Repositories in the context of our corporate projects (1) have lower incidence of clone pairs than the Open Source Projects (2) reported in the literature.
  - H_1^{CP}: $\mu 1_{(\text{Clone Pairs Proportion})} \neq \mu 2_{(\text{Clone Pairs Proportion})}$

HYPOTHESIS 2.

- **Null hypothesis H_0^{CC}:** Source Code Repositories in the context of our corporate projects (1) have same incidence of clone classes of the Open Source Projects (2) reported in the literature.
  - H_0^{CC}: $\mu 1_{(\text{Clone Classes Proportion})} = \mu 2_{(\text{Clone Classes Proportion})}$
- **Alternative hypothesis H_1^{CP}:** Source Code Repositories in the context of our corporate projects (1) have higher incidence of clone classes than the Open Source Projects (2) reported in the literature.
  - H_1^{CC}: $\mu 1_{(\text{Clone Classes Proportion})} \neq \mu 2_{(\text{Clone Classes Proportion})}$

Independent Variables. NICAD method; Java OSS Projects reported in the literature; Our Java Industrial Projects. Moreover, the metrics used to evaluate this experiment were reused from the previously mentioned study [5]. We have used four UPI thresholds for the whole work: 0.0, 0.1, 0.2 and 0.3.

Dependent Variables. The proportions and averages between results of clone pairs and classes and their respective KLOCs will be used as dependent variables. They are described as follows:

- **Proportion**: $P_S = C_S/KLOC_S$
- **Final Average**: $\mu = (P_{S1} + P_{S2} + \ldots + P_{Sn}) / n$

Objects Selection. The private code projects size varied from a 6K LOC to a 35K LOC application. This selection was done by convenience. We have used some corporate projects which we were clone consultants for. The analysis is non-intrusive to developers as the data were drawn directly from the code repository, they did not know which source code would be extracted. Our repository contained only Java systems, to compare with the results of Open Source Java Systems obtained from the previously referenced study.

Instrumentation. We have used NICAD tool described in section 3. Results are printed to the standard output. Additional information results are exported to XML and HTML files in the same directory of the original system source.

5 Experiment Operation

In this section, we describe the whole experiment execution. The detection tool was configured to consider only functions or methods with a minimum of 3 LOC. We do not analyze in this work clone distribution and localization over files or directories.

5.1 Execution

First, we extracted clone information for the whole repository to compare with the Open Source results using the Percentage of Total Cloned Methods (TCMp). Then, each project was analyzed individually and every clone-related discovered information was recorded and analyzed.

At first, we can confirm that inside our Java repository there are fewer clone manifestations than the Open Source Java repository studied in Roy and Cordy work [5]. We may see in Figure 1 an incidence from 15% to almost 25% of TCMp, meanwhile in our experiment we did not reach 15% even using the highest UPI threshold value.

Our results also show that going from UPI value 0.0 to 0.3, the TCMp metric do not grow as the Java Open Source code. They show a 10% growth in clone appearance while in our repository clones increases only by 4% with the same test. The next step is to evaluate all projects individually.

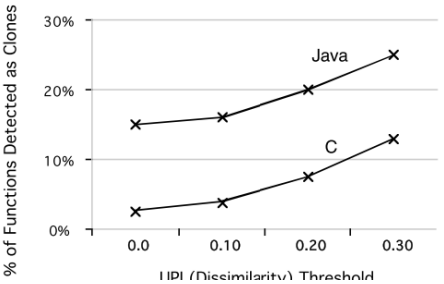


Fig. 1 TCMp values extracted from an OSS experiment [5].

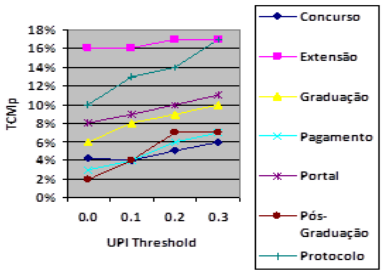


Fig. 2 TCM Percentage for Individual Systems

Running the NICAD clone detection tool for each system, we identified the projects with more and few clone incidences. We present these results in Figure 2.

Analyzing the graph, we note that clone incidence is not related to the project size. The bigger the worst does not apply here, since we have Graduação with less than 10% TCMp inside 35K LOC versus Extensão presenting more than 16% TCMp for only 7K LOC.

5.2 Data Validation

The NICAD clone detection tool generated HTML reports where we extracted the cloned methods to validate by hand the clone pairs with low similarity. Also, using Linux diff tool we compared the XML output file generated by NICAD to determine textual similarities of clones with the original source. This two-step process was the same used by the referenced work.

To ensure analysis, interpretation and validation, we used two types of statistical tests: Shapiro-Wilk Test and the Mann-Whitney test. Shapiro-Wilk test was used to verify normality of the samples. The Mann-Whitney test was used to check our hypothesis. All statistical tests were performed using the SPSS tool [21].

6 Results

To answer our experiment question, we executed all individual tests and created a table showing data to compare with the results obtained from OSS experiment. The Table 1 shows several statistics collected from the analysis of our experiment.

6.1 Analysis and Interpretation

Clone detection statistics from all the Java OSS analyzed are also present on Table 1. The “T” values on “Clone Pairs” and “Clone Classes” are representing the UPI thresholds. The OSS represent significantly much more LOC than corporate systems. The “PROP.” column represents the proportional values for T=0.0.

Ant project showed excellent results in comparison to the Academic System, both having about 35 KLOC. Ant returned less CPs and more CCs, which indicates a much lower clone incidence with the usage of methods abstraction. The worst performance was with Swing where proportionally presented more exact clones than Academic System.

For the OSS, Spule, JHotDraw and Jdtcore were the more clone-free projects. Analysis of Spule returned only 4,62% of clones. Following the same KLOC value, we can exemplify with the Protocol System, which we found 12,93% Clone Pairs. The corporate system with less Clone Pairs was the Payment System, with among 2% of exact clones. Post-grad System holds better performance about Clone Pairs, with 0,38% of incidence. The only OSS with clone incidence below 1% was Spule. The final average found for the OSS and Corporate Clone Pairs and Classes can be found on Table 2.

Table 1 Nicad statistics for OSS and Private code Repositories.

REPOSITORY	PROJECT NAME	KLOC	CLONE PAIRS T=					CLONE CLASSES T=				
			0.0	P_S (0.0)	0.1	0.2	0.3	0.0	P_S (0.0)	0.1	0.2	0.3
Open Source Software	Ant	35	363	10,37	365	374	426	92	2,63	94	101	119
	EIRC	11	117	10,64	117	121	149	35	3,18	35	36	47
	Spule	13	60	4,62	64	68	113	11	0,85	13	15	19
	Javadoc	14	193	13,79	197	240	304	80	5,71	82	95	110
	JHotDraw	40	291	7,28	295	377	598	137	3,43	141	170	208
	Jdtcore	148	1427	9,64	1553	2126	4378	323	2,18	377	518	660
	Swing	204	8115	39,78	8203	9978	11209	516	2,53	558	687	843
Corporate Software	Concurso (Contest System)	6	31	5,17	33	38	40	10	1,67	10	13	15
	Extensão (Extension System)	7	120	17,14	120	121	127	53	7,57	53	54	55
	Pagamento (Payment System)	8	16	2	20	47	89	5	0,63	7	10	11
	Pós-Graduação (Post-grad System)	8	30	3,75	32	38	40	3	0,38	5	9	9
	Portal (Web Portal System)	9	194	21,56	205	213	258	19	2,11	21	24	25
	Protocolo (Protocol System)	14	181	12,93	217	243	266	31	2,21	38	43	51
	Graduação (Academic System)	35	1280	36,57	1315	1401	1460	44	1,26	60	70	78

Table 2 Final average results

FINDINGS	REPOSITORY	
	OSS	CORPORATE
CLONE PAIRS	13,73	14,16
CLONE CLASSES	2.93	2.26

Based on these results, we observe that was not significant difference between the two kinds of repository. The Corporate Systems showed a little more proportional clone incidence than OSS. With this data, is not possible yet to make any assumption about results without sufficiently conclusive statistical evidence.

Firstly, we applied the Shapiro-Wilk test with a significance level of 0.05, analyzing the distribution normalization. The Sig variables (also known as p-values) for Clone Pairs were 0,003 on OSS samples and 0,370 on Industry samples. For Clone Classes, the p-values were 0,006 on Industry Systems and 0,541 on OSS. The numbers of at least one sample for each hypothesis were below the significance level, so, we assume that data distribution is not normal for all samples.

Applying the Mann-Whitney non-parametric test, we obtained a Sig. result of 0,110 for Clone Classes samples and 0,949 for Clone Pairs, both above the significance level of 0.05. Thus, the final decision is do not reject the null hypothesis.

In fact, for Clone Pairs there was a strong retention for the null hypothesis H_0^{CP} ($\mu 1$(Clone Pairs Proportion) = $\mu 2$(Clone Pairs Proportion)). In real terms, there is a probability of almost 95% that we will mistakenly reject the similar clone incidence, although closed-source coding is easy to control, has stricter methodologies and more controlled development teams.

Systems with high incidence of clone pairs and much lower incidence of clone classes at the same time as the Academic System have methods that are excessively replicated. This represents a lot of code reuse and lack of using abstraction.

From the data extracted by NICAD, we calculated TCMp values for all projects. For each UPI threshold, we obtained the number of Clone Classes (CC). The CC were divided by the Potential Clone (PC) methods to obtain finally the percentage of TCM metric. The table also shows a Potential Clone Lines statistic capture by the tool, already filtered in pretty-printed format.

6.2 Threats to Validity

In spite of the fact that our corporate systems are a mature, real world, large projects, and our results seem to be quite consistent with the systems sizes, our study shows threats to its validity that we must consider:

- We cannot conclude that all closed-source projects will present similar results as ours. Process maturity can play a large role on code clone manifestation;
- Other software characteristics such as complexity and programming paradigm may affect the results. We have not test for those variables;
- Adoption of design patterns also may influence on code clone manifestation;
- The profile of the development team (team size, age, experience) also can represent a change on the final sample.

7 Conclusions and Future Work

We found evidences in our experiment that clone incidence is not directly related to the size of code. In fact, the studied corporate systems had similar cloning incidence as the OSS ones. We encourage more research inside private environments to test hypothesis only studied on Open Source Software systems. In addition, our corporate systems had very few KLOCs than the Open Source ones. It is important to replicate this experiment inside several other private repositories to check if they present the same behavior. The more the systems are tested more we assure external validity.

As mentioned before, we adapted the software engineering experimental process described in Wohlin et al [12] to clones mining experiments. We believe that the studies, applications, and tools for software clone mining can benefit from this type of approach. Rigorous experimental description facilitates replication of studies and the executing of systematic reviews and other types of secondary analysis.

As future work, we have in mind a few projects related to clone incidence. The first one is to verify if the human profile of development team has some direct effect on clone appearance. Data like age, experience and qualification may be extracted and combined from several sources to mount this profile. Other insight is to explore code comments to find out words that indicate something that was purposely implemented missing some pieces (for many reasons) and this will have to be done some time, indicating a Technical Debt (TD) issue.

Acknowledgements This study could only be developed due to the support of Tiradentes University – UNIT, along with the Technology Information Department – DTI, who provided the repository used in our experiment.

References

1. Baxter, I.D., Yahin, A., Moura, L., Sant'Anna, M., Bier, L.: Clone detection using abstract syntax trees. In: ICSM, pp. 368–377 (1998)
2. Ma, Y., Woo, D.: Applying a code clone detection method to domain analysis of device drivers. In: Proceedings of the 14th Asia Pacific Software Engineering Conference (APSEC 2007), pp. 254–261, Nagoya, Japan, December 2006
3. Marcus, A., Maletic, J.I.: Identification of high-level concept clones in source code. In: 16th IEEE Intern. Conf. on Auto. Soft. Eng., pp. 107–114, San Diego, CA, USA, November 2001
4. Khatoon, S., Mahmood, A., Li, G.: An evaluation of source code mining techniques. In: Proc. – 8th Int. Conf. Fuzzy Syst. Knowl. Discov., FSKD 2011, vol. 3, pp. 1929–1933 (2011)
5. Roy, C.K., Cordy, J.R.: An empirical study of function clones in open source software. In: 15th Work. Conf. Reverse Eng., pp. 81–90 (2008)
6. Cordy, J.R., Roy, C.K.: The NiCad clone detector. In: 2011 IEEE 19th Int. Conf. Progr. Compr., no. Figure 3, pp. 219–220 (2011)
7. Schwarz, N., Lungu, M., Robbes, R.: On how often code is cloned across repositories. In: Proc.–Int. Conf. Softw. Eng., pp. 1289–1292 (2012)
8. Rattan, D., Bhatia, R., Singh, M.: Software clone detection: A systematic review, vol. 55(7). Elsevier B.V. (2013)
9. Kim, M., Sazawal, V., Notkin, D.: An empirical study of code clone genealogies. ACM SIGSOFT Softw. Eng. Notes **30**, 187 (2005)
10. Roy, C.K., Cordy, J.R., Koschke, R.: Comparison and evaluation of code clone detection techniques and tools: A qualitative approach. Sci. Comput. Program. **74**, 470–495 (2009)
11. Sarala, S., Deepika, M.: Unifying clone analysis and refactoring activity advancement towards C# applications. In: 4th IntConf. Comput. Commun. Netw. Technol, ICCCNT (2013)
12. Duala-Ekoko, E., Robillard, M.P.: Clonetracker: tool support for code clone management. In: Icse 2008 Proc. Thirtieth Int. Conf. Softw. Eng., pp. 843–846 (2008)
13. Duala-Ekoko, E., Robillard, M.P.: Tracking code clones in evolving software. In: Proc.–Int. Conf. Softw. Eng., pp. 158–167 (2007)

14. Kapser, C.J., Godfrey, M.W.: 'Cloning considered harmful' considered harmful: Patterns of cloning in software. Empir. Softw. Eng. **13**, 645–692 (2008)
15. The TXL Programming Language. http://www.txl.ca
16. Wohlin, P. Runeson, M. Host, M. C. Ohlsson, B. Regnell, Wesslén, A.: Experimentation in software engineering: an introduction. Kluwer Academic Publishers (2000). ISBN: 0-7923-8682-5
17. Colaço, M., Mendonça, M., M. André, M., Farias, D.F., Henrique, P.: A neurolinguistic-based methodology for identifying OSS developers context-specific preferred representational systems. Context, no. c, pp. 112–121 (2012)
18. Guo, Y., Seaman, C., Gomes, R., Cavalcanti, A., Tonin, G., Silva Da, F.Q.B., Santos, A.L.M., Siebra, C.: Tracking technical debt - An exploratory case study. In: IEEE Int. Conf. Softw. Maintenance, ICSM, pp. 528–531 (2011)
19. van Solingen, R., Berghout, E.: The Goal/Question/Metric Method: A practical guide for quality improvement of software development. McGraw-Hill (1999)
20. SPSS, IBM Software. http://goo.gl/eXfcT3
21. Roy, C.K., Cordy, J.R.: A survey on software clone detection research. Queen's Sch. Comput. TR **115**, 115 (2007)

Fuzzy Resource-Constrained Time Workflow Nets

Joslaine Cristina Jeske de Freitas and Stéphane Julia

Abstract The underlying proposal of this work is to express in a more realistic way the resource allocation mechanisms when human behavior is considered in Workflow activities. In order to accomplish this, fuzzy sets delimited by possibility distributions are associated with the Petri net models that represent human type resource allocation mechanisms. Additionally, the duration of activities that appear on the routes (control structure) of the Workflow process, are represented by fuzzy time intervals produced through a kind of constraint propagation mechanism. New firing rules based on a joint possibility distribution are then defined. Finally, the model is built using CPN (Colored Petri net) language supported by CPN Tools.

Keywords Petri net · Workflow net · Fuzzy · Time constraints · CPN tool

1 Introduction

The purpose of Workflow Management Systems is to execute Workflow processes. Workflow processes represent the sequence of activities that have to be executed within an organization to treat specific cases and to reach a well-defined goal. Of all notations used for the modeling of Workflow processes, Petri nets are very suitable [1], as they represent basic routing of business processes. Moreover, Petri nets can be used for specifying the real time characteristics of Workflow Management Systems (in the time Petri net case) as well as complex resource allocation mechanisms. As a matter of fact, late deliveries in an organization are generally due to resources overload.

J.C. Jeske de Freitas(✉) · S. Julia
Universidade Federal de Uberlândia, Uberlândia, MG, Brazil
e-mail: joslaine@gmail.com, stephane@facom.ufu.br
http://www.facom.ufu.br

J.C. Jeske de Freitas—would like to thank FAPEMIG for its financial support.

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_48

Workflow nets have been identified and widely used as a solid model in Workflow processes by several authors [2–5]. In [5], an extension of Workflow nets is presented. This model is called time Workflow net and associates time intervals with the transitions of the corresponding Petri net model. In [6], an extended Workflow Petri net model is defined. Such a model allows for the treatment of critical resources which have to be used for specific activities in real time. In [7], a resource-oriented Workflow net (ROWN) based on a two-transition task model was introduced for resource-constrained Workflow modeling and analysis. Considering the possibility of task failure during execution, in [8], a three-transition task model for specifying the start, end and failure of a given task was proposed. Additional research can be found in [9–12].

The majority of existing models put their focus on the process aspect and do not consider important characteristics of the Workflow Management System. In [2, 3] for example, the resource allocation mechanisms are represented only in an informal way. In [5, 6] resource allocation mechanisms are represented by simple tokens in places, as it is generally the case in production systems [13]. However a simple token in a place will not represent in a realistic context human employees who can treat different cases simultaneously in a single day, as it is usually the case in most Business processes.

Colored Petri Net (CPN) language was introduced by Kurt Jensen [14] and supported by CPN Tools. CPNs allow for the modeling of systems in which communication, synchronization, and resource sharing play an important role. The CPN language combines the strengths of Petri nets with the strengths of high-level programming languages. Petri nets provide the primitives for process interaction, whereas the CPN language provides the primitives for the definition of data types and the manipulation of data values. One of the main advantages of CPNs is their strong theoretical basis and the existence of tools to support the modeling and analysis of CPN models.

The core proposal of this work is to build a model using CPN language to express in a more realistic context resource allocation mechanisms when human behavior is considered. To reach this goal, fuzzy sets delimited by possibility distributions [15] are associated with the Petri net models that represent human type resource allocation mechanisms and fuzzy time intervals are associated to activity durations. A firing mechanism using a joint possibility distribution is then defined in order to associate through a single formalism explicit time constraints as well as resource availability information.

The remainder of this paper is as follows. Section 2 introduces the principal concepts of fuzzy sets and possibility measures, Workflow modeling and CPN (Colored Petri Net). Section 3 presents resource allocation mechanisms, a fuzzy time constraint propagation mechanism, new firing rules that consider fuzzy time constraints as well as fuzzy resource allocation mechanisms and the CPN model. Finally, section 4 concludes the paper and provides references for additional works.

2 Background

2.1 Fuzzy Sets and Possibility Measures

The notion of fuzzy set was introduced by [16] in order to represent the gradual nature of human knowledge. A certain degree of belief can be attached to each possible interpretation of symbolic information and can simply be formalized by a fuzzy set F of a reference set X that can be defined by a membership function $\mu_F(x) \in [0, 1]$. In particular, for a given element $x \in X$, $\mu_F(x) = 0$ denote that x is not a member of the set F, $\mu_F(x) = 1$ denotes that x is definitely a member of the set F, and intermediate values denote the fact that x is more or less an element of F. Normally, a fuzzy set is represented by a trapezoid $A = [a1, a2, a3, a4]$ where the smallest subset corresponding to the membership value equal to 1 is called the core, and the largest subset corresponding to the membership value greater than 0 is called the support.

When considering two distincts fuzzy sets A and B, the basic operations are as follows [17]:

- the fuzzy sum $A \oplus B$ defined as:
  $[a1, a2, a3, a4] \oplus [b1, b2, b3, b4] =$
  $[a1 + b1, a2 + b2, a3 + b3, a4 + b4]$,
- the fuzzy subtraction $A \ominus B$ defined as:
  $[a1, a2, a3, a4] \ominus [b1, b2, b3, b4] =$
  $[a1 - b4, a2 - b3, a3 - b2, a4 - b1]$,
- the fuzzy product $A \otimes B$ defined as:
  $[a1, a2, a3, a4] \otimes [b1, b2, b3, b4] =$
  $[a1.b1, a2.b2, a3.b3, a4.b4]$.

A fuzzy set F can be delimited by a possibility distribution [15, 18] Π_f, such as: $\forall x \in X, \Pi_f(x) = \mu_F(x)$. Given a possibility distribution $\Pi_a(x)$, the measure of possibility $\Pi(S)$ and necessity $N(S)$ that a data a belongs to a crisp set S of X are defined by $\Pi(S) = sup_{x \in S} \Pi_a(x)$ and $N(S) = inf_{x \notin S}(1 - \Pi_a(x)) = 1 - \Pi(\overline{S})$. If $\Pi(S) = 0$, it is impossible that a belongs to S. If $\Pi(S) = 1$, it is possible that a belongs to S, but it also depends on the value of $N(S)$. If $N(S) = 1$, it is certain (the larger the value of $N(S)$, the more the proposition is believed in). In particular, it exists a duality relationship between the modalities of the possible and the necessary which postulates that an event is necessary when its contrary is impossible. Some practical examples of possibility and necessity measures are presented in [15].

Given two data a and b characterized by two fuzzy sets A and B, the measure of possibility and necessity of having $a \leq b$ are defined as:

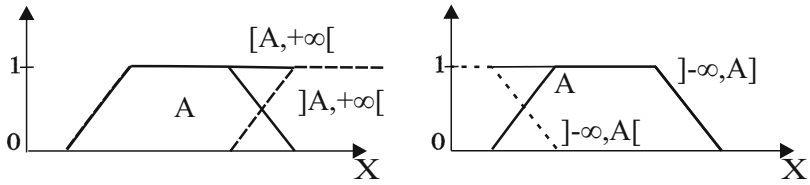
$$\Pi(a \leq b) = sup_{x \leq y}(min(\Pi_a(x), min(\Pi_b(y))) = max([A, +\infty[\cap] - \infty, B]) \quad (1)$$

and

$$N(a \leq b) = 1 - sup_{x \leq y}(min(\Pi_a(x), min(\Pi_b(y)))). \quad (2)$$

Given a normalized possibility distribution π_a, [19] defines the following fuzzy sets of the time point that are:

- possibly after a: $\mu_{[A,+\infty[}(X) = sup_{x \in X}\pi_a(s)$ (see figure 1(a));
- necessarily after a: $\mu_{]A,+\infty[}(X) = inf_{x \in X}(1 - \pi_a(s))$ (see figure 1(a));
- possibly before a: $\mu_{]-\infty,A]}(X) = sup_{x \in X}\pi_a(s)$ (see figure 1(b));
- necessarily before a: $\mu_{]-\infty,A[}(X) = inf_{x \in X}(1 - \pi_a(s))$ (see figure 1(b)).



(a) Possibly/Necessarily after a. (b) Possibly/Necessarily before a.

Fig. 1 Possibly/Necessarily

A visibility time interval $[a, b]$ is a period of time between two dates a and b. In the case where a and b are fuzzy dates A and B (delimited by π_a and π_b) respectively, the interval $[a, b]$ is represented by the the following pair of fuzzy sets:

- $[A, B]$, the conjunctive set of time instants that represents the set of dates possibly after A and possibly before B;
- $]A, B[$, the conjunctive set of time instants that represents the set of dates necessarily after A and necessarily before B.

The joint possibility admits as upper bound in [15]:

$$\forall x \in X \quad \forall y \in Y \quad \pi(x, y) = min(\pi_X(x), \pi_Y(y)) \quad (3)$$

when the reference sets are non-interactive (the value of x in X has no influence on the value of y in Y, and vice versa).

2.2 *Workflow Modeling*

A Petri net that models a Workflow process is called a Workflow net [1]. A Workflow net satisfies the following properties [20]:

- It has only one source place, named *Start* and only one sink place, named *End*. These are special places such that the place *Start* has only outgoing arcs and the place *End* has only incoming arcs.

– A token in *Start* represents a case that needs to be handled and a token in *End* represents a case that has been handled.
– Every task t (transition) and condition p (place) should be on a path from place *Start* to place *End*.

According to Aalst [1] the task are modeled by transitions. However, an activity of a Workflow Process can also be represented by a specific place of an ordinary Petri net [21] with an input transition which shows the beginning of the activity and an output transition which shows the end of the activity.

2.3 Concepts of CPN

CPN is a graphical modeling language [14], which combines the strengths of Petri nets and of functional programming languages. The formalism of Petri nets is well suited for describing concurrent and synchronizing actions in distributed systems. Programming languages can be used to define data types and manipulation of data. An introduction to the practical use of CPN can be found at [22].

The CPN are designed to reduce the size of the model, allowing individualization of tokens, using colors assigned to them, so different processes or resources can be represented in the same network. Colors do not mean just colors or patterns. They can represent complex data types [14].

CPN models are formal, in the sense that the CPN modeling language has a mathematical definition of its syntax and semantics. This means that they can be used to verify system properties, i.e., prove that certain desired properties are fulfilled or that certain undesired properties are guaranteed to be absent.

Large and complex models can be built using hierarchical CPN in which modules, which are called pages in the CPN terminology, are related to each other in a well-defined way. Without the hierarchical structuring mechanism, it would be difficult to create understandable CPN models of real-world systems [14].

The practical application of CPN modeling and analysis heavily relies on the existence of computer tools supporting the creation and manipulation of models. CPN Tools is a tool suite for editing, simulating, providing state space analysis, and providing performance analysis of CPN models. The user of CPN Tools works directly on the graphical representation of the CPN model. The graphical user interface (GUI) of CPN Tools has no conventional menu bars or pull-down menus, but is based on interaction techniques, such as tool palettes and marking menus. A license for CPN Tools can be obtained free of charge via the CPN Tools web pages http://www.cpntools.org.

3 Approach

3.1 Resource Allocation Mechanism

There exists different kinds of resources in Workflow processes. Some of which are of the discrete type and can be represented by a simple token. On the contrary, some other resources cannot be represented by simple tokens. This is generally the case with human type resources. As a matter of fact, it is not unusual for an employee who works in an administration to treat several cases simultaneously. For example, in an insurance company, a single employee can normally treat several documents during a working day and not necessarily in a pure sequential order. In this case, a simple token could not model human behavior in a proper manner. Fuzzy allocation mechanisms were presented in [23].

3.2 Fuzzy Time Constraint Propagation Mechanism

As the actual time required by an activity in a Workflow Management System is non-deterministic and not easily predicted, a fuzzy time interval can be assigned to every Workflow activity.

The static definition of a fuzzy time Workflow net is based on fuzzy static intervals $[a1, a2, a3, a4]s$ which represent the permanency duration (sojourn time) of a token in places. Before duration $a1$ the token is in the non-available state. After $a1$ and before $a4$, the token is in the available state for the firing of a transition. After $a4$, the token is again in the non-available state and cannot enable any transition: it therefore becomes a dead token. In a real time system case, the "death" of a token has to be seen as a time constraint that is not respected. A transition cannot be fired with dead tokens as this would correspond to an illegal action or behavior: a constraint violation.

In a Workflow Management System, a visibility interval depends on a global clock associated to the entire net which calculates the passage of time from date zero, which corresponds to the start of the system operations. In particular, the existing waiting times between sequential activities can be represented by visibility intervals whose minimum and maximum fuzzy boundaries will depend on the earliest and latest delivery dates of the considered case. Through correct knowledge of the beginning date of the process and the maximum duration of a case, it is possible to calculate estimated visibility intervals associated with each token in each waiting place using constraint propagation techniques very similar to the ones used in scheduling problems based on activity-on-arc graphs without circuits [24].

Figure 2 shows "Handle Complaint Process" presented in [25]. This figure shows basic routings, activity time, visibility intervals and resource allocation mechanisms. Resources $R1$, $R2$, $R4$ and $R5$ are fuzzy resources and $R3$ is discrete resource. The fuzzy static intervals (intervals of fuzzy durations) associated to the activity places of

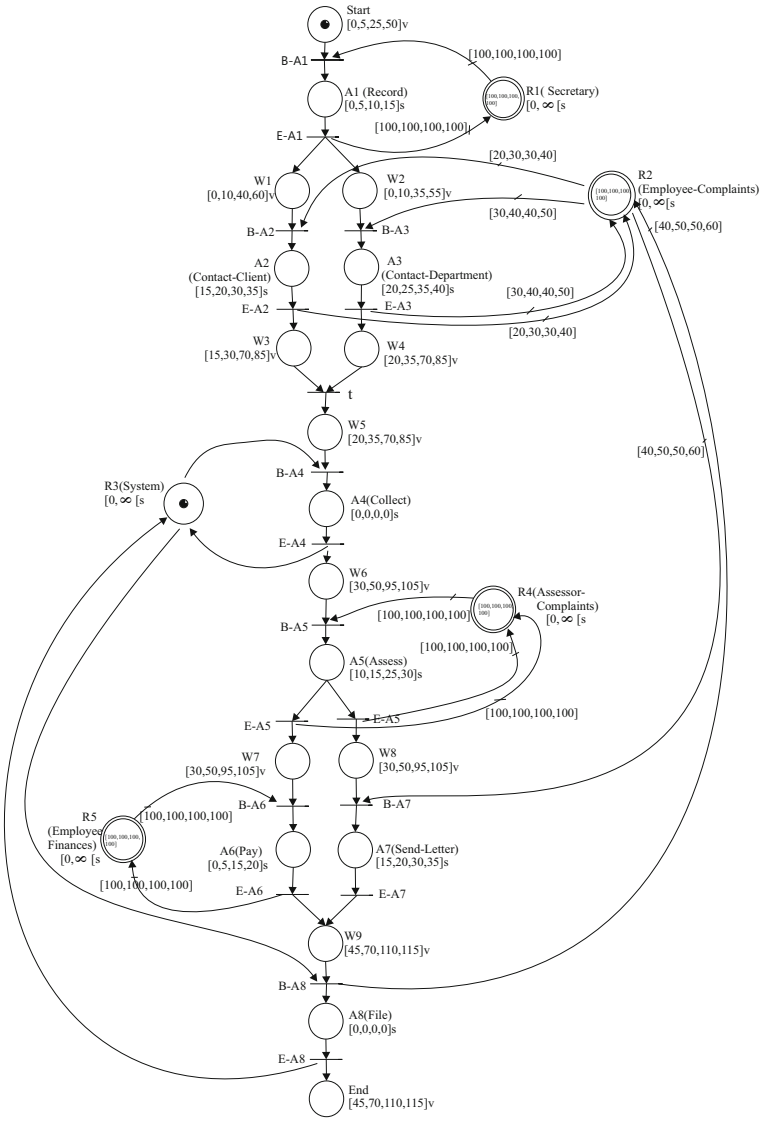


Fig. 2 "Handle Complaint Process".

the process and the fuzzy visibility intervals (intervals of fuzzy dates) associated with the waiting places (condition places of the Workflow net). It is important to note that there is no time restriction on resources - static interval defined for each resource is $[0, \infty[s$. The minimal fuzzy bounds of the estimated visibility intervals attached to the waiting places are calculated applying a forward constraint propagation technique applied to the different kinds of routings associated with the "Handle Complaint

Process", and the maximum fuzzy bounds of the estimated visibility intervals are calculated by applying a backward constraint propagation techniques to the different kinds of routings considering the latest delivery dates of the case. The complete definition of Fuzzy Time Constraint Propagation Mechanism for Workflow Nets can be found in [26].

3.3 Firing Rules and Token Player Algorithm

If a transition has n input places and if each one of these places has several tokens in it, then the enabling time interval $[a1, a2, a3, a4]$ of this transition is obtained by choosing for each one of these n input places a token, the visibility interval associated with it. In this paper, there exists no time restriction on the resources (the static interval attached to the resource places is always $[0, \infty[s)$ and, as a consequence, the enabling time interval of a transition will simply be equal to the visibility interval associated with the case to be treated by the corresponding transition. For example, knowing that the visibility interval attached to the case represented by a token in place $W1$ is equal to $[0, 10, 40, 60]v$, the enabling time interval of the transition $B\text{-}A2$ will be $[0, 10, 40, 60]v$ too.

For firing a transition, it is necessary that the arrival date of the token in the input place of the transition belongs to the fuzzy visibility interval associated with the input place of the transition ($\mu > 0$) and the resource availability (equation (1)) necessary to realize the activity initiated by the firing of the transition must be greater than 0 ($\Pi(a \leq b) > 0$). To evaluate the availability of resource and time simultaneously, the joint possibility presented in equation (3) must be calculated, where $\pi_X(x)$ corresponds to the resource availability and $\pi_Y(y)$ to the time location of the corresponding activity.

3.4 "Handle Complaint Process" - CPN model

The modeling approach through the use of the CPN Tools is hierarchical, where each transition can be replaced by a specific module. The most abstract level of the modeling of the "Handle Complaint Process" is depicted in figure 3.

Observing figure 3, note that:

- A token in place *Start* enables the transition that defines the $A1$ activity for firing.
- The colset *PLACE* is a set of colors that defines the types associated with the places of the model. In addition, there exists two types of resources defined in the model: RESOURCE that defines discrete resource and RESOURCEF that defines the fuzzy resources.
- Each of the transitions with double borders denominated as "abstract transitions" creates a link with one of the sub-networks (it corresponds to the hierarchy concept)

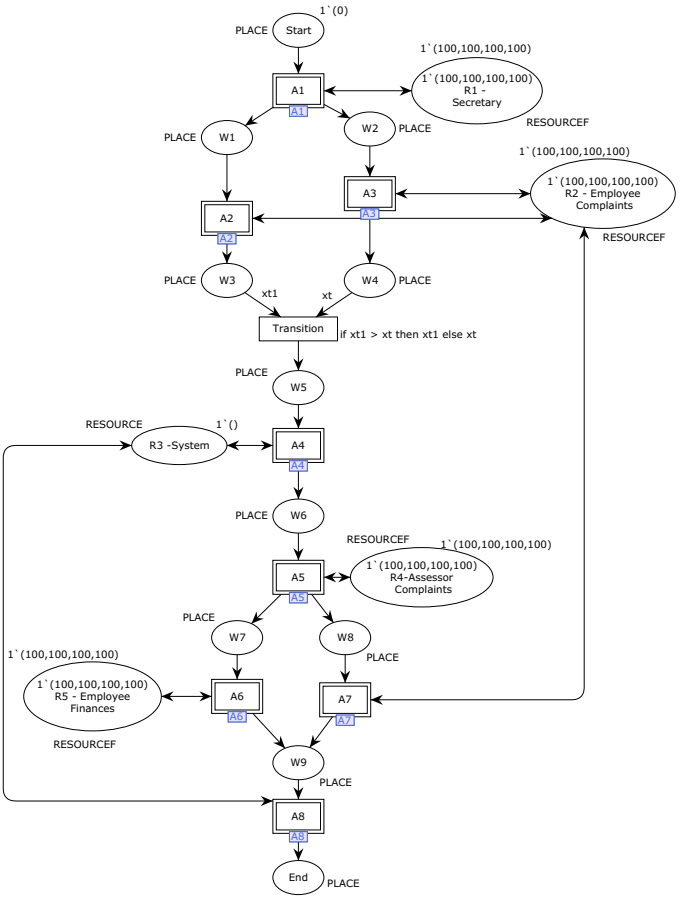


Fig. 3 The most abstract level of "Handle Complaint Process".

$A1$, $A2$, $A3$, $A4$, $A5$, $A6$, $A7$ and $A8$. As already mentioned, $A2$ and $A3$ are parallel activities. At the end of these activities, a synchronization must be made. In this case, it is necessary to compare the end time of each activity and use the higher value. In the proposed model, this condition is defined with the subscription expression $if\ xt1 > xt\ then\ xt1\ else\ xt$ in the output arc.

Due to limited space the modeling of all activities cannot be shown. As the modeling of each activity is similar, only the activity $A1$ will be presented. (Figure 4).

Basically, each activity is defined as an input transition bounded by the guard condition which should guarantee that there exists resource available and time is within the visible interval. The first part $(\#1(RBA1)) \leq (\#4(r1))$, checks if there is resource availability and the second part $((xt \geq 0) and also (xt \leq 50))$, checks if the time is within the visible interval. If the transition is fired then the *Startproc function* is called and the resource is used. The *SartProc function* takes as its input

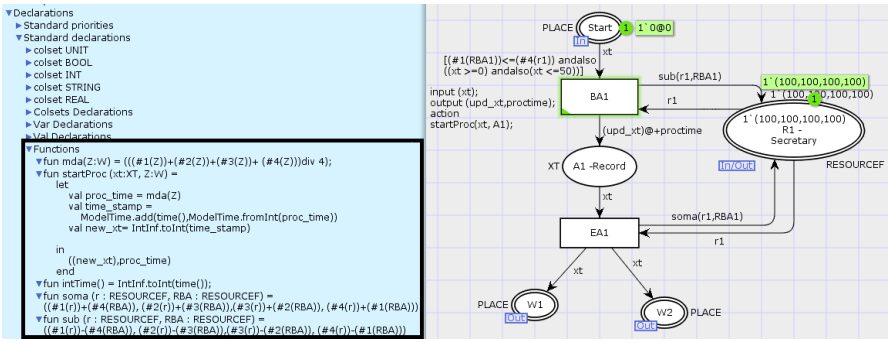


Fig. 4 Modeling Activity A1.

the current time value and the interval that defines the duration of the activity, and returns the updated time value after the end of the activity. When the resource is used, the *sub function* is called to calculate the new value of resource availability. When the activity ends, the value of the resource is returned by *soma function*.

4 Conclusion

This article presented how to model fuzzy hybrid resources in Workflow nets with fuzzy time intervals associated to the activities using CPN Tools. Besides this, through the definition as well as use of a joint possibility distribution, it was possible to define a transition firing definition. This definition takes into consideration the time constraints associated with the cases of the process, as well as the availability of the resources used to execute the activities. CPN is an interesting modeling approach as it is capable of describing the software error detection and obtaining greater confidence in terms of model correctness.

As a future work proposal, it will be interesting to use this mechanism in Resource Workflow net to simulation in CPN Tools for qualitative and quantitative analysis with the various proposals for resource allocation mechanism.

References

1. Aalst, W., van Hee, K.: Workflow Management: Models, Methods, and Systems. MIT Press, Cambridge (2004)
2. Aalst, W.: Verification of workflow nets. In: Proceedings of the 18th International Conference on Application and Theory of Petri Nets, ICATPN 1997, pp. 407–426. Springer, London (1997)
3. van Hee, K., Sidorova, N., Voorhoeve, M.: Resource-constrained workflow nets. Fundam. Inf. **71**(2,3), 243–257 (2006)

4. Martos-Salgado, M., Rosa-Velardo, F.: Dynamic soundness in resource constrained workflow nets. In: Bruni, R., Dingel, J. (eds.) FMOODS/FORTE. Lecture Notes in Computer Science, vol. 6722, pp. 259–273. Springer (2011)
5. Ling, S., Schmidt, H.: Time Petri nets for workflow modelling and analysis. In: 2000 IEEE International Conference on Systems, Man and, Cybernetics, vol. 4, pp. 3039–3044. IEEE (2000)
6. Kotb, Y.T., Badreddin, E.: Synchronization among activities in a workflow using extended workflow petri nets. In: CEC, pp. 548–551. IEEE Computer Society (2005)
7. Wang, J., Tepfenhart, W.M., Rosca, D.: Emergency response workflow resource requirements modeling and analysis. IEEE Transactions on Systems, Man, and Cybernetics, Part C **39**(3), 270–283 (2009)
8. Wang, J., Li, D.: Resource oriented workflow nets and workflow resource requirement analysis. International Journal of Software Engineering and Knowledge Engineering **23**(5), 677–694 (2013)
9. Adogla, E.G., Collins, J.W.: Managing resource dependent workflows, US Patent 8,738,775, May 27, 2014
10. He, L., Chaudhary, N., Jarvis, S.A.: Developing security-aware resource management strategies for workflows. Future Generation Computer Systems **38**, 61–68 (2014)
11. Deng, N., Zhu, X.D., Liu, Y.N., Li, Y.P., Chen, Y.: Time management model of workflow based on time axis. Applied Mechanics and Materials **442**, 458–465 (2014)
12. Guo, X., Ge, J., Zhou, Y., Hu, H., Yao, F., Li, C., Hu, H.: Dynamically predicting the deadlines in time-constrained workflows. In: Web Information Systems Engineering-WISE 2013 Workshops, pp. 120–132. Springer (2013)
13. Lee, D.Y., DiCesare, F.: Scheduling flexible manufacturing systems using petri nets and heuristic search. IEEE T. Robotics and Automation **10**(2), 123–132 (1994)
14. Jensen, K., Kristensen, L.: Coloured Petri Nets. Springer (2009)
15. Dubois, D., Prade, H.: Possibility theory. Plenum Press, New-York (1988)
16. Zadeh, L.A.: Fuzzy sets. Information and Control **8**, 338–353 (1965)
17. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall Inc., Upper Saddle River (1995)
18. Cardoso, J., Valette, R., Dubois, D.: Possibilistic petri nets. IEEE Transactions on Systems, Man, and Cybernetics, Part B **29**(5), 573–582 (1999)
19. Dubois, D., Prade, H.: Processing fuzzy temporal knowledge. IEEE Systems, Man and Cybernetics, 729–744 (1989)
20. Aalst, W.: The Application of Petri Nets to Workflow Management. Journal of Circuits, Systems, and Computers **8**(1), 21–66 (1998)
21. David, R., Alla, H.: Discrete, Continuous, and Hybrid Petri Nets, 2nd edn. Springer Publishing Company, Incorporated (2010)
22. Kristensen, L.M., Jrgensen, J.B., Jensen, K.: Application of coloured petri nets in system development. In: Desel, J., Reisig, W., Rozenberg, G. (eds.) Lecture on Concurrency and Petri Nets. LNCS, vol. 3089, pp. 626–685. Springer (2004)
23. Jeske, J.C., Julia, S., Valette, R.: Fuzzy continuous resource allocation mechanisms in workflow management systems. In: XXIII Brazilian Symposium on Software Engineering: SBES 2009, pp. 236–251. IEEE (2009)
24. Gondran, M., Minoux, M., Vajda, S.: Graphs and Algorithms. John Wiley & Sons Inc., New York (1984)
25. Julia, S., de Oliveira, F.F., Valette, R.: Real time scheduling of Workflow Management Systems based on a p-time Petri net Model with Hybrid Resources. Simulation Modelling Practice and Theory **16**(4), 462–482 (2008)
26. Jeske de Freitas, J.C., Julia, S.: Fuzzy time constraint propagation mechanism for workflow nets. In: 12th International Conference on Information Technology - New Generations (ITNG), pp. 367–372 (2015)

Performance Indicators Analysis in Software Processes Using Semi-supervised Learning with Information Visualization

Leandro Bodo, Hilda Carvalho de Oliveira, Fabricio Aparecido Breve and Danilo Medeiros Eler

Abstract Software development process requires judicious quality control, using performance indicators to support decision-making in the different processes chains. This paper recommends the use of machine learning with the semi-supervised algorithms to analyze these indicators. In this context, this paper proposes the use of visualization techniques of multidimensional information to support the labeling process of samples, increasing the reliability of the labeled indicators (group or individual). The experiments show analysis from real indicators data of a software development company and use the algorithm bioinspired Particle Competition and Cooperation. The information visualization techniques used are: Least Square Projection, Classical Multidimensional Scaling and Parallel Coordinates. Those techniques help to correct the labeling process performed by specialists (labelers), enabling the identification of mistakes in order to improve the data accuracy for application of the semi-supervised algorithm.

Keywords Software processes · Software quality · Performance indicators · Machine learning · Information visualization · MPS-SW · BSC

L. Bodo(✉) · H.C. de Oliveira · F.A. Breve
Department of Statistics, Applied Mathematics and Computer Science,
Unesp - Universidade Estadual Paulista, São Paulo, Brazil
e-mail: lebodo@gmail.com, {hildaz,fabricio}@rc.unesp.br

D.M. Eler
Department of Mathematics and Computer Science,
Unesp - Universidade Estadual Paulista, São Paulo, Brazil
e-mail: daniloeler@fct.unesp.br

L. Bodo · H.C. de Oliveira · F.A. Breve
Unesp - Universidade Estadual Paulista, Rio Claro, Brazil

D.M. Eler
Unesp - Universidade Estadual Paulista, Presidente Prudente, Brazil

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_49

1 Introduction

Software development companies have been facing problems related to quality, time and cost in the software development process for decades [1]. The control of processes contributes to minimize these problems and it requires measurements at multiple frequencies and scales. However, this control occurs only in organizations with some pre-established maturity level. Thus, it can check whether the results are in accordance with previously defined quality criteria. Key Performance Indicators (KPI) are essential resources for the quality control, since they quantify the performance of the processes, activities and products, enabling the comparison with the established goals [2].

Generally, the KPIs can be analyzed individually or in groups, to provide relevant information for quick and effective decision making. The indicator groups can support the evaluation of some situations that involve one or more processes/projects. However, indicators that make up a group can be characterized by different granularity, types and frequency of data collection. Thus it is important that the analysis of indicators groups do not use mathematical operations such as average. These operations have limited the indicators understanding, requiring different techniques with higher semantic expression for the indicator analysis.

Machine learning techniques have been used to support the analysis of large data volumes generated by KPIs in different application areas. In the Software Engineering area, Bodo et al. [3] propose the use of a semi-supervised learning algorithm called "Particle Competition and Cooperation" (PCC), in order to analyze KPIs individually or in group in software development processes. According to the authors, the algorithm was effective in analyzing real indicators of software processes related to the Project Management and Requirements Management, according to the software quality model MPS-SW [4].

Moreover, the use of a semi-supervised learning technique requires that the labeled samples are reliable, considering that only some of the samples have to be labeled. Observe that the labeling process identifies the samples belonging to a specific data set with a "label" (e.g.: color, letter, word, etc.). However, the labeling process is a manual activity prone to errors, even though it is done by specialized personnel, such as project managers, for example.

In this context, this paper proposes the use of multidimensional information visualization techniques to support the KPI labeling process in software engineering processes. The goal is to enable the visualization of the existent relationship between instances of the indicators in a multidimensional plan. Thus, people who is labeling the data can identify gaps in the process and correct them before applying the machine learning algorithm. The failures can be identified from visual analysis, observing the distance between the different instances of the data sets, which are represented by dots with different colors for each label in the set. The result can be viewed on the screen or printed, in order to facilitate the visual identification by the specialist conducting the training process. Overall, the infor-

mation visualization techniques allow experts to visually recognize structures and patterns present in the data [5].

In order to present the proposal, this paper considered the PCC semi-supervised learning algorithm used by Bodo et al. [3]. A real KPI data set was considered as well, which was provided by a company that develops software for public management.

The PCC algorithm is a bioinspired algorithm developed by Breve [6] based on ant communities. This algorithm has a low computational complexity when compared to other algorithms of the same structure [7]. The algorithm provides benefits on the labeling process for indicators analysis in software processes for two basic reasons: (1) it requires less human effort in the labeling process - which helps lowering its costs; (2) it considers the instance labels - which are relevant for the identification of clusters, formed by KPIs.

Thus this paper was organized into seven sections. *Section 2* presents a literature review on topics related to the purpose of this paper, which address the machine learning use on performance indicators and information visualization techniques. *Section 3* presents the recommended information visualization techniques for the purpose of this paper, before applying the PCC algorithm, which is presented in *Section 4*.

The experiments presented in this paper used a real historical data of performance indicators. Two reference models presented in [3] were used to support the selection and grouping of these indicators. These models are presented briefly in the *Section 5*, along with the group of performance indicators used.

Section 6 presents experiments that demonstrate how useful the visualization techniques are for the process of labeling samples. Three multidimensional information visualization techniques were used, presenting failures, which could be corrected by the project manager (labeler) before running the PCC algorithm. Lastly, *Section 7* presents the final considerations, with future research prospects.

2 Related Work

Machine learning has been used to analyze performance indicators in different application areas. For example, Melo et al. [8] show how the supervised machine learning technique Multilayer Perceptron (MLP) contributed to the monitoring and prevention measures on vehicle traffic in concession roads. On the other hand, Neto et al. [9] present the use of Kohonen Feature Maps, which is a non supervised learning technique, to group the socioeconomic indicators of the agricultural cooperatives. In the software engineering area, Bodo et al. [3] introduce the use of semi-supervised learning, with the technique called Particle Competition and Cooperation (PCC), for KPI analysis in the software processes. The authors also present a reference model to support the definition of KPIs for processes related to levels G and F of the MPS-SW model: Requirements Management, Project Management, Measurement, Configuration Management, Acquisition, Quality Assurance and Project Portfolio Management. Another reference model is proposed by

Bodo et al. [3] based on the perspectives of the strategic model Balanced Scorecard (BSC). The purpose is to assist the composition of the indicator groups which must be analyzed together.

According to Bodo et al. [3], semi-supervised learning techniques are more appropriate to the control of software processes than supervised and unsupervised learning techniques. Firstly, because they consider the opinion of experts that are working in the processes, unlike what happens with unsupervised learning techniques. Moreover, they do not demand much time of those specialists for the execution of the labeling samples process, like the supervised learning techniques. Using semi-supervised learning, in some scenarios it is possible to obtain a classification so accurate as the classification obtained with supervised learning, but using less labeled samples and consequently reducing the specialists work.

It is important to point out that errors may occur in the labeling process, especially when the experts are labeling groups of two or more KPIs. Thus, this paper proposes the use of information visualization techniques to support the labeling process in order to enable the identification of mistakes made by specialists (labelers). Therefore, the labeling process can be revised in order to eliminate the identified errors and, as a consequence, increase the accuracy of the labeled samples before applying the semi-supervised machine learning algorithm.

Generally, visualization techniques aid data analysis in several application fields, transforming geometric symbolic data in order to show what is not explicitly visible. Thus, the information display helps to visualize data that have no physical representation predefined. Visualizations that encode multivariable data (more than one variable) are called multidimensional visualizations.

Multidimensional projection techniques have been used in visual data mining [5]. Each data instance is mapped to a visual element in a visualization space 1D, 2D or 3D (e.g.: a point or circle). The relative position of each visible element in the plan represents some kind of relationship between data instances. According to Paulovich and Minghim [10], similarity or neighborhood is the most common type of relationship in data analysis. For example, Paiva et al. [11] use multidimensional projection techniques in visual data mining to support the visual classification of data set. The authors emphasize that experts in the domain were able to quickly identify results of "false positive" and "false negative", as well as inconsistencies and discrepancies in data classification.

3 Information Visualization Techniques to Support Data Labeling

Information visualization techniques can be used in different types of problems that require the support of a graphical representation (visual) of the data analyzed (e.g.: text, music, social networking, etc.). The purpose is to provide visual conditions to identify and understand the relationships between the instances of the data sets.

These techniques exploit the visual substrate using color and metaphorical figures to represent information.

The information visualization techniques can be classified according to the characteristics of the data to be visualized: one-dimensional, two-dimensional, three-dimensional and multidimensional (four or more dimensions) [12]. In this paper the number of dimensions corresponds to the amount of KPIs from a group of indicators analyzed (two KPIs or more). Generally, multidimensional projection techniques preserve the similarity relations of the original space in a smaller dimension space (e.g., 2D). The intention is to create graphical representations which assist the user to explore the relationship between the data instances. For example, nearby points on a plane forming clusters indicate instances with high similarity level.

This paper recommends the use of three multidimensional projection techniques: (1) Least Squares Projection; (2) Classical Multidimensional Scaling; (3) Parallel Coordinates. The software tools used in this work are freeware and developed in Java. The first two techniques were applied with the PEx-Image (Projection Explorer for Images) tool, proposed by Eler et al. [13]. The third technique was applied using a tool that combines multidimensional projection and parallel coordinates, developed by Oliveira et al. [14].

The Least Squares Projection (LSP) technique uses the neighborhood control points to establish the relation of similarities among the data instances, preventing the calculation of similarity between any data set [15]. The application of this technique creates groups of points with level of similarity in 2D space. According to Paulovich et al. [15], this information visualization technique is one of the most accurate.

The Classical Multidimensional Scaling (CMDS) technique is based on linear data models and shows a geometric structure to view the distances among the data instances [16]. Generally, the similarity between data set instances is used as input. According to Birchfield and Subramanya [17], the application of this technique in environments with linear characteristics has been shown to be relevant, although it also depends on the database analyzed.

The Parallel Coordinates (PC) technique assigns an axis for each attribute (dimension) and shows these parallel axes in a plan. Each instance of a set of n-dimensional data is represented by a line passing through the parallel axes. It is possible to see the behavior of attributes (dimensions) of data instances through parallel coordinates. This allows the comparison between the instances from different data set clusters.

4 Semi-supervised Learning Technique Used

The Particle Competition and Cooperation (PCC) algorithm is bioinspired and belongs to the swarm intelligence category. The algorithm is inspired by ant communities, that work in a “cooperative” way in search for food and in a

"competitive" way in the conquest of new territories. The algorithm is based on a connected network graph, which has various particles (ants) spread by the network.

The particles are organized in the form of teams (groups) and walk throughout the network in a cooperative and competitive way. The same team particles aim to spread their label and walk throughout the network cooperating with each other. On the other hand, particles from different teams compete with each other to determine the borders of the classes and rejecting intrusive particles. The central idea of the PCC algorithm is that the particles walk along the graph in the competition process for the nodes, trying to dominate the largest possible amount of nodes and protecting the nodes that have been dominated. At the end of the iterative process, the particles team (cluster) that dominated the node is used to label this node.

Thus, every particle chooses a neighboring node to visit through the "random x greedy" rule in each iteration. The "random" walk objective is the acquisition of new nodes: each particle randomly chooses any neighbor to visit. The "greedy" walking objective is the defense of the territory: each particle visits the nodes conquered for its team and the nodes that are closer to their home nodes (nodes to the reference of the distance from the neighbors) [6].

The team domain level is increased whenever a particle of this team selects a neighboring node. Each team tries to avoid the invasion of particles from other teams. At the end of the iterative process, each unlabeled data item will be labeled with the team's label with greater domain. Every node on the network has an array of elements representing the level of each particles team. At first, the input data are transformed into an undirected network, without weights. Each node corresponding to an unlabeled sample will have its domain level configured with the same value $1/c$, where c is the number of classes (teams).

Figure. 1(a) illustrates an initial scenario with four classes of unlabeled nodes and domain levels equal to 0.25: [0.25 0.25 0.25 0.25]. Then, a group of particles with labeled nodes is placed in the network and its domain level is set to the highest value [6]. Fig. 1(b) displays this situation for the four classes, but with labeled node equal to 1: [1 0 0 0].

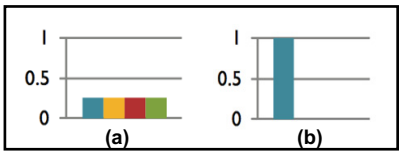


Fig. 1 Initial domination level: (a) unlabeled sample; (b) labeled sample [6].

5 Performance Indicators and Reference Models Used

This section presents the performance indicators that were used in the labeling process prior to the application of the PCC algorithm. The Brazilian company that provided such data has been operating in the software market for public administration for over 30 years and has hundreds of municipal governments as customers.

The data are part of a historical basis of performance indicators and were collected for about three years, through software monitoring processes. The company is certified in level G of MPS-SW model [4], presented in the *Subsection 5.1*.

A group of five indicators collected daily was considered in this paper, as shown in the *Section 6*. These indicators were chosen with the support of two project managers and the indicators are associated with two processes of levels G and F of the MPS-SW model: Management Requirements and Quality Assurance. The indicators (numeric types) are: (1) "Customer Requests", which identifies the number of customer requests and reports; (2) "Service level", which indicates the relationship between the requests made by customers and the amount of people available for service; (3) "Open Requirements", which identifies the amount of requirements open on the day; (4) "Unimplemented requirements", which indicates the amount of not implemented requirements at end of period; (5) "Impediment requirements", which displays the total requirements that aborted the project.

The indicators were grouped with the support of a reference model based on the perspectives of the Balanced Scorecard (BSC). The grouping of indicators was made by the customer's perspective and strategic aspects of the company. *Subsection 5.2* provides an overview of this model proposed by Bodo et al. [3].

5.1 Quality Model MPS-SW

The software quality model MPS-SW was created in 2003 and is part of Brazilian Software Process Improvement Program (MPS.BR). The model is maintained by the Association for Promoting the Brazilian Software Excellence (Softex) and is designed to benefit mainly micro, small, and medium software enterprises (MSME). The MPS-SW model is based on the ISO/IEC 12207, ISO/IEC 15504, ISO/IEC 20000 and the CMMI (Capability Maturity Model Integration).

The MPS-SW model is incremental and has seven maturity levels, starting at level G (the lowest level) to the level A (the highest level). Each maturity level has processes that must be added to the processes of the previous level. Thus, the implementation of the process is gradual up to level A that values the continued improvement in such processes [18]. It is important to notice that almost 90% of certifications in MPS-SW model are for levels G or F. There are two processes in the level G: Requirements Management and Project Management. There are five processes in the level F: Measurement, Configuration Management, Acquisition, Quality Assurance and Project Portfolio Management.

5.2 Reference Model to Group Indicators

Performance indicators can be grouped according to defined criteria by the own software development company. However, the company can use a reference model for grouping the indicators proposed by Bodo et al. [3]. This model has four classes of indicator groups that correspond to the four strategic Customers of BSC technique: (1) Financial; (2) Customer: (3) Internal business processes; (4) Learning and growth [19].

The "status" of the set of indicators is represented by three clusters, according to the metaphor of a traffic light. The green light indicates that it is to continue with the processes involved, because the situation is satisfactory. The yellow sign warns to pay attention, because the values indicate a situation with regular level of satisfaction. The red light indicates that the processes involved must be stopped because the displayed result is unsatisfactory, far away the established target. The five indicators selected for this paper were grouped according to the Customer perspective.

6 Information Visualization in the Labeling Process

This section shows how the information visualization techniques can increase the reliability of labeling process in the performance indicators. Thus, *Subsection 6.1* shows the implementation of the LSP and CMDS techniques using the PEx-Image tool [13], as well as the PC technique application, using the proposed tool by [14]. The concept of Euclidean distance was used in all techniques, without reducing dimensions. Five attributes were used, which were represented by the five indicators presented in the *Section 5*. The instances were extracted from the indicators monitoring database, totaling 660 instances. Three output clusters were considered: C1 (satisfactory) with 16% of samples, C2 (regular) with 14% of samples and C3 (unsatisfactory) with 70% of samples.

The two project managers of the partner company were instructed to learn how to do the data analysis from the historical bases of performance indicators. The intention was to search a semantic result for the "exit", considering three status possibilities of the process: satisfactory, regular and unsatisfactory. Thus, the labeling of the 660 samples was made by the project managers, who manually analyzed the historical data, line-by-line of the measurement. It should be noted that 100% of the data were labeled in the experiment to be able to analyze the accuracy rate of the PCC algorithm, presented in the *Subsection 6.3*. Then, the three multidimensional information visualization techniques have been applied to evaluate if the data labeling process was correctly performed. The results of information visualization techniques have enabled the correction of samples labeling, as presented in the *Subsection 6.2*. Therefore, the semi-supervised learning PCC algorithm was applied more safely, as shown in the *Subsection 6.3*. Note that each attribute had its value normalized before being used in PCC algorithm.

6.1 Application of Information Visualization Techniques

The first technique applied to the set of 660 labeled samples was the Least Squares Projection (LSP). The colors that represent the cluster are: green (cluster 1) to status "satisfactory", yellow (cluster 2) to status "regular" and red (cluster 3) to status "unsatisfactory". Each point corresponds to an instance of the indicators, i.e., the value obtained in measurements. Each point is identified by the letter "A" followed by the instance number.

Figure. 2 shows a partial view of the results obtained from the LSP technique where the grouping of the class "regular" (yellow) is in evidence. The A130 and A147 instances had its labels assigned to the "unsatisfactory" cluster (red). However, it is possible to identify that they are closer to the "regular" cluster (yellow) than the "poor" cluster (red). This proximity with a different class of the original class indicates that an error may have occurred in the labeling process. On the other hand, the A186 instance is at the limit of the border of the classes "regular" and "unsatisfactory", though it has been attributed the "regular" label (yellow) for this instance. Therefore, the A186 sample could belong to the cluster "regular" or "unsatisfactory".

Thus the visual analysis using the LSP technique shows that errors may have occurred in the labeling process. Nevertheless, it is not possible to say that really happened with the A130, A147 and A186 instances using only this technique. Then the other two techniques presented in the *Section 3* were used to better evaluate the behavior of the attributes of these three instances.

Due to proximity in the visual representation of the A130 and A147 samples (cluster "unsatisfactory") with the samples labeled as "regular", the first step was to evaluate the behavior of these instances applying the Parallel Coordinates (PC) technique. It is worth remembering that the colors correspond to the classes, the axes represent the attributes (indicators), the lines represent the data instances (indicators values) and each line crosses the parallel axis to the corresponding value to the attribute. The A130 and A147 instances are represented in the Fig. 3 by the color red. It is possible to see that these instances present a very similar behavior to the instances of the class "regular", which are represented by yellow lines. This reinforces the possibility of an error has occurred in the labeling of the A130 and A147 samples.

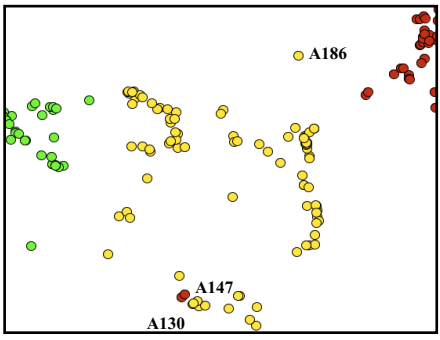


Fig. 2 View of three probable errors in the labeling process using the LSP technique.

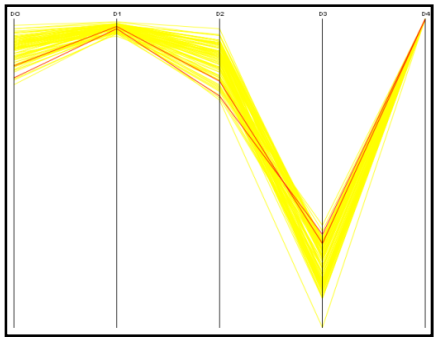


Fig. 3 View of probable errors in the labeling process of the instances A130 e A147 using the PC technique.

The Classical Multidimensional Scaling (CMDS) technique was used to evaluate in another way the relationship between the instances presented in the Figs. 2 and 3. The result shown in the Fig. 4 strengthened the possibility to have occurred errors in the labeling of the samples A130 and A147. Also, the results showed the necessity of questioning on the label attributed to the A186 sample, which is much closer to the "unsatisfactory" class (red) than the "regular" class (yellow).

6.2 Correcting of the Mislabeled Samples

Once visualization techniques evidenced the possibility of three samples labeled with errors, the two project managers reviewed the labeling of these samples by means explanation of the impact of this problem on an analysis of indicators. After a few meetings and discussions, the project managers admitted isolated errors in the labeling of the samples A130, A147 and A186 and they made the necessary correction.

Then, the three information visualization techniques have been applied to the revised data set. Figs 5 and 6 show the new application's results of LSP and CMDS techniques, respectively. It is possible to see that the distribution of classes is much more uniform, because potential conflicts in the identification of the points in the classes are not shown in the plan. On the other hand, the result of applying the PC technique is shown in the Fig. 7. It is important to point out that is not possible to identify with the naked eye the discrepancies between the A130 and A147 instances and the rest of the instances.

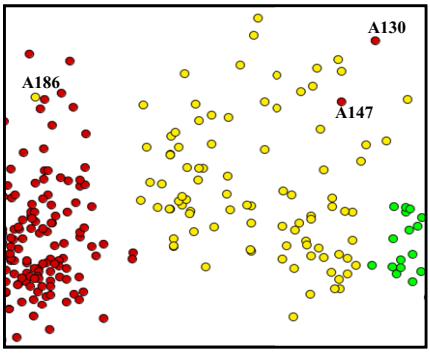


Fig. 4 Emphasis on the occurrence of three errors in the labeling process using CMDS technique.

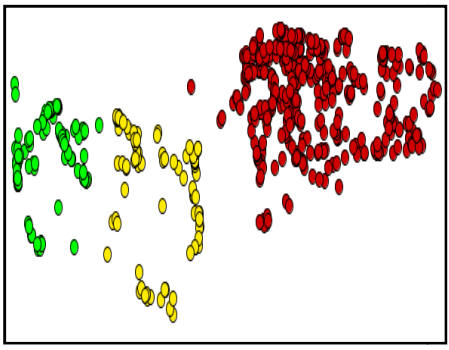


Fig. 5 LSP technique application on the set of samples after labeling correction.

6.3 Application of the PCC Algorithm

The machine learning with the semi-supervised algorithm PCC was applied twice: on the set of labeled samples initially and on the set of samples with corrected labels (after analysis of the information visualization techniques). Thus, the two

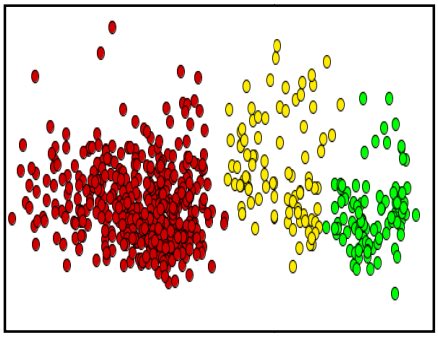


Fig. 6. CMDS technique application on the set of samples after labeling correction.

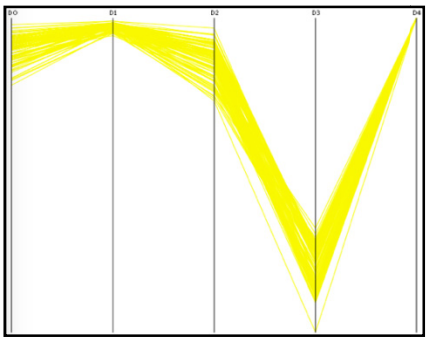


Fig. 7. PC technique application on the set of samples after labeling correction.

applications of the algorithm used the same historical databases and the same parameter values, including the amount of "*k*-neighbors" (*k*=9). Table 1 shows the results of the accuracy rate per cluster, which were obtained before the application of information visualization techniques, where: C1 corresponds to "satisfactory" cluster, C2 "regular" and C3 "unsatisfactory". It could be noted that the PCC algorithm correctly classified all the samples C1. However, the algorithm missed the classification of fourteen C2 cluster samples (they were classified as C3) and three samples of C3, (they were classified as C2).

Thus, the software development company may have a wrong orientation using the Table 1. The table indicates the necessity of corrective actions for fourteen samples that may have been wrongly labeled as "unsatisfactory" (C3). The other case is worse because the table indicates that three samples must be labeled as "regular" (C2), although they have been labeled as "unsatisfactory" (C3). On the other hand, Table 2 shows the result of applying the algorithm PCC after the application of information visualization techniques. It could be noted that the PCC algorithm correctly classified all the samples C1 and C3. However, the algorithm has mistakenly classified three samples of the C2 class (they were classified as C3). Nevertheless, the algorithm has classified the three samples more appropriately to the analysis of the indicators, i.e., those were classified as "unsatisfactory" (before they were classified as "regular"). It is noteworthy that in cases of doubt the algorithm must classify the indicators in a lower class, so that they are treated and possible classification errors can be detected. In this work, it was essential that the algorithm has not classified these three samples as "satisfactory" because this would indicate that no corrective action should be performed.

Therefore, the results achieved after the corrective actions on the labeling process of the samples show large benefits in the analysis of performance indicators. Certainly, this will have a large impact on decision-making related to software processes.

Table 1 Accuracy rate per class before the application of information visualization techniques.

		PREDICTION		
		C1	C2	C3
ACTUAL	C1	**105**	0	0
	C2	0	**79**	14
	C3	0	3	**459**

Table 2 Accuracy rate per class after the application of information visualization techniques.

		PREDICTION		
		C1	C2	C3
ACTUAL	C1	**105**	0	0
	C2	0	**90**	3
	C3	0	0	**462**

7 Conclusion

This paper considered using semi-supervised machine learning techniques for analysis of indicators in software processes. The algorithm Particle Competition and Cooperation (PCC) was used as reference of this type of technique in the presented analysis. In order to improve the results of these techniques, this paper proposes the use of information visualization techniques to support specialists in the labeling process of samples. The recommended techniques were: Least Square Projection, Classical Multidimensional Scaling and Parallel Coordinates.

This paper presented experiments using historical bases of actual indicators, which were provided by a software development company, certified on the level G of the MPS-SW quality model. The information visualization techniques were applied over the labeled samples of a five indicators group. This enabled the project managers to identify probable errors in the labeling process. Is worth emphasizing that the use of the three techniques was important to evidence the probable errors. Thus, the identified samples were relabeled by the project managers. The application of the PCC algorithm before and after the labeling correction of the samples was important, showing significant gains in the accuracy rate of the classes of the indicators after the corrections.

Other versions of PCC algorithm are available including active learning, which can help indicate the degree of confidence in each label. Those versions should be addressed in future work in order to investigate the impact in the selection of samples more representative for the labeling process. In addition, the PCC algorithm will also be analyzed to classify samples which labels have been assigned mistakenly.

From this work, efforts are being invested in the systematization of the indicators analysis process using the PCC algorithm and the recommended visualization techniques. Thus, the core of a support system for monitoring and control of KPIs can be developed as well as graphical user interfaces with dashboards to better assist the specialists in the processes.

Acknowledgements The authors acknowledge the financial support of the São Paulo Research Foundation (FAPESP) (grants #2013/03452-0 and #2011/17396-9) and the National Counsel of Technological and Scientific Development (CNPq) (grant #475717/2013-9).

References

1. Rezende, A.D.: Engenharia de software e sistemas de informação, 3rd edn. Brasport, Rio de Janeiro (2005)
2. Boyd, L.H., Cox, J.F.A.: Cause-and-effect approach to analyzing performance measures. Production and Inventory Management Journal **38**(3), 25–32 (1997)
3. Bodo, L., Oliveira, H.C., Breve, F.A., Eler, D.M.: Semi-supervised learning applied to performance indicators in software engineering processes. In: International Conference on Software Engineering Research and Practice (SERP 2015), Las Vegas, EUA, pp. 255–261 (2015)
4. SOFTEX - Associação para Promoção da Excelência do Software Brasileiro. MPS.BR Melhoria de processo do software brasileiro: Guia de Implementação – Parte 9: Implementação do MR-MPS em organizações do tipo Fábrica de Software (2012). http://www.softex.br/wp-content/uploads/2013/07/MPS.BR_Guia_de_Implementacao_Parte_9_20111.pdf. Acesso em: January 20, 2013
5. Tejada, E., Minghim, R., Nonato, L.G.: On improved projection techniques to support visual exploration of multidimensional Data Sets. Information Visualizarion **2**(4), 218–231 (2003)
6. Breve, F.A.: Aprendizado de máquina utilizando dinâmica espaço-temporal em redes complexas. 2010. 165 f. Tese (Doutorado) - Curso de Ciências de Computação e Matemática Computacional, Departamento de ICMC-USP, Universidade de São Paulo, São Carlos (2010)
7. Breve, F.A., Zhao, L., Quiles, M.G., Pedrycz, W., Liu, J.: Particle competition and cooperation in networks for semi-supervised learning. IEEE Transactions on Knowledge and Data Engineering, [s.i] (2009)
8. Melo, A.C., Silva, A.L., Marte, C.L., Ferreira, M.R., Sassi, R.J., Ferreira, R.P.: Aplicação de técnicas da inteligência artificial para obtenção de indicadores de desempenho como medida de qualidade em rodovias cocessionadas. In: 7º Congresso Brasileiro de Rodovias e Concessões (CBR&C), 2011, Foz do Iguaçu. 3º Salão de Inovação ABCR (Associação Brasileira de Concessionárias de Rodovias) (2011)
9. Neto, S.B., Nagano, M.S., Moraes, M.B.C.: Utilização de redes neurais artificiais para avaliação socioeconômica: uma aplicação em cooperativas. R. Adm, São Paulo **41**(1), 59–68 (2006)
10. Paulovich, F.V., Minghim, R.: Mapeamento de dados multi-dimensionais - integrando mineração e visualização. In: Congresso da Sociedade Brasileira de Computação (CSBC), Bento Gonçalves, pp. 9–10 (2009)
11. Paiva, J.G.S., Cruz, L.F., Pedrini, H., Telles, G.P., Minghim, R.: Improved Similarity Trees and their Application to Visual Data Classification. IEEE Transactions on Visualization and Computer Graphics **17**(12), 2459–2468 (2011)
12. Nascimento, H.A.D., Ferreira, C.B.R.: Visualização de informações - Uma abordagem prática. XXV Congresso da Socuedade Brasileira de Computação, pp. 1262–1312 (2005)
13. Eler, D.M., Nakazaki, M.Y., Paulovich, F.V., Santos, D.P., Andery, G.F., Oliveira, M.C.F., Neto, J.B., Minghim, R.: Visual analysis of image collections. The Visual Computer **25**(10), 923–937 (2009)
14. Oliveira, R.A.P., Silva, L.F., Eler, D.M.: Hybrid visualization: a new approach to display instances relationship and attributes behaviour in a single view. In: 19th International Conference on Information Visualization (2015) (to appear)

15. Paulovich, F.V., Nonato, L.G., Minghim, R., Levkowitz, H.: Least Square Projection: a fast high precision multidimensional projection technique and its application to document mapping. IEEE Transactions on Visualization and Computer Graphics **14**(3), 564–575 (2008)
16. Warren, T.: Multidimensional scaling of similarity. Psychometrika **30**(4), 379–393 (1965)
17. Birchfield, S.T., Subramanya, A.: Microphone Array Position Calibration by Basis-Point Classical Multidimensional Scaling. IEEE Transactions on Speech and Audio Processing **13**(5), 1025–1034 (2005)
18. Kalinowski, M., Weber, K., Franco, N., Barroso, E., Duarte, V., Zanetti, D., Santos, G.: Results of 10 years of software process improvement in Brazil based on the MPS-SW model. In: 9th International Conference on the Quality of Information and Communications Technology. IEEE (2014)
19. Kaplan, R., Norton, D.P.: The Balanced Scorecard – Measures that Drive Performance. Harvard Business Review, pp. 71–79, January–February 1992

Techniques for the Identification of Crosscutting Concerns: A Systematic Literature Review

Ingrid Marçal, Rogério Eduardo Garcia, Danilo Medeiros Eler, Celso Olivete Junior and Ronaldo C.M. Correia

Abstract Modularization is a goal difficult to achieve in software development. Some requirements, named crosscutting concerns, cannot be clearly mapped into isolated source code units, and their implementations tend to cut across multiple units. Although several researches propose new approaches to identify crosscutting concerns, few works aim to provide analysis, synthesis and documentation of the aspect mining literature. To address this research gap, we conducted a systematic literature review to provide researchers with a state-of-the-art of the existing aspect mining techniques. We point out challenges and open issues on most of the techniques analyzed that could be improved in further researches.

Keywords Crosscuttingness · Review · Aspects · Crosscutting-concerns · Concerns

1 Introduction

Crosscutting concerns are pieces of functionality that have their implementation tangled and/or scattered among various modules in a software system. They are difficult to identify and hinder program comprehension. The aspect mining research area is concerned with the development of tools and techniques to support the identification of crosscutting concerns in legacy software. Many techniques have been proposed from source code analysis [6] [5] [8] to the use of information available in software repositories [7] [3].

Whereas researchers demonstrate empirically [15][16] [17] [18] the negative impact of crosscutting concerns in software internal and external quality metrics,

I. Marçal(✉) · R.E. Garcia · D.M. Eler · C.O. Junior · R.C.M. Correia
College of Science and Technology, São Paulo State University – UNESP,
Campus at Presidente Prudente, São Paulo, Brazil
e-mail: in.marcal@gmail.com, {rogerio,daniloeler,olivete,ronaldo}@fct.unesp.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_50

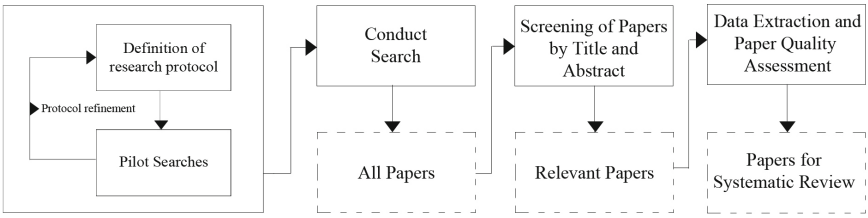


Fig. 1 Systematic Literature Review Process

increases the number of new techniques and tools for identifying them. Few works, however, aim to provide the analysis, synthesis and documentation of aspect mining literature [30] [22]. To address this research gap, our systematic literature review (SLR) focuses on analyzing and synthesizing empirical evidence on aspect mining, aiming to provide researchers with a state-of-the-art of the existing techniques.

Our SLR helped to spot issues beyond weel known aspect mining problems as the subjectivity of most analysis, imprecise definitions and poor discussion of results. Our in-depth analysis of aspect mining techniques can serve as a road map for researchers to improve their techniques and achieve better results. Based on our results and observations, we also propose new research directions to improve the existing aspect mining techniques.

To present our SLR and results obtained, the remainder of this paper is organized as follows. Section 2 describes the adopted research method. Section 3 gives a brief overview of the studies selected for this SLR and Section 4 describes the related work. Finally, Section 6 draws conclusions and presents the observed opportunities for future work in aspect mining.

2 Research Method

This research has been carried out as a Systematic Literature Review, which involves research methodologies developed to collect and assess documented evidence related to a specific topic [2]. SLRs are formal, systematic and follow a rigorous methodology predefined under a research protocol [2], which helps selecting relevant studies, improving rigor and making the reviewing a repeatable process. A protocol contains the definition of inclusion/exclusion criteria, questions for quality assessment, methods for data extraction and the research questions. The protocol for this SLR was elaborated based on rules, policies and procedures determined by the Campbell Collaboration[1].

We conducted a series of pilot searches to refine the definitions in the protocol. After the refinement, we carried out a primary search to obtain the set of papers.

[1] www.campbellcollaboration.org

Then, we conducted a screening of papers by title and abstract to select only those relevant to our SLR. Finally, we evaluated each relevant paper according to our quality assessment criteria and extracted important information. Our systematic literature process is depicted in Figure 1. The following sections describe our research protocol in terms of databases and search strategy, studies selection criteria, quality assessment and data extraction. We, also, state our research question.

2.1 Databases and Search Strategy

A search string was executed in four different databases (see Table 1) considering papers published from 2000 to 2014. We defined the following standard search string:

((crosscutting OR crosscut OR cross-cutting) AND (concern OR concerns)) AND (("aspect mining") OR ("separation of concerns") OR ("code mining") OR "coding mining")

However, different databases employ a particular search syntax. Thus, we used our search string to create an equivalent for each database. Our search resulted in 1289 papers, and 1274 were excluded for non-compliance with our selection criteria or do not achieve the minimum score defined in our quality assessment (See Sections 2.2 and 2.3). Our quality assessment resulted in 53 studies relevant to this SLR. Due to space issues this paper we only discuss details about 14 selected studies and draw conclusions based on all 53 selected studies. The number of papers retrieved from each database is shown in Table 1.

Table 1 Selected Databases

Database	Number of Studies
ACM Digital Library	214
IEEE Xplore	928
SpringLink	97
Willey Inter Science Journal Finder	50

2.2 Studies Selection Criteria

From all papers found by our search strategy, we selected to this SLR only those that showed empirical evidence of their results. Furthermore, we exclude all papers related to: Early aspects, editorials, tutorials, summaries and panels. We also exclude studies that proposed non-automatic or semi-automatic techniques.

2.3 Quality Assessment and Data Extraction

We elaborated 18 questions (Table 2) which corresponds to four analysis criteria considered important to ensure quality: (1) Initial Filter; (2) Rigor; (3) Credibility; (4) Relevance. Each question was answered according to the following. 0 point: Not acceptable; 0.5 point: Poorly Acceptable; 1 point: Acceptable. In Table 2, Questions 1-3 corresponds to the initial filter and helps to decide whether or not a paper must be considered for further analysis. Questions 4-12 assess the rigor with which the paper was elaborated. Questions 13-17 refer to the study credibility and question 18 evaluates the relevance of a paper. Final classification of a paper is obtained by adding the points obtained in each question (The classification is shown in Table 4).

Only papers classified with *Good Quality* and *Very Good Quality* were included in this SLR and read fully to collect relevant data . We used a set of questions (see Table 3) to help extracting all relevant information from each paper.

Table 2 Quality Assessment Questions

ID	Question
Q1	Is the study based on research or is it just an experience report?
Q2	Is the study main focus on crosscutting concerns identification/comprehension?
Q3	Does the study present experimental data or real application of some aspect mining technique?
Q4	Is there a satisfactory description of the context in which the study is inserted?
Q5	Is it clear how the results were obtained?
Q6	Were the results extraction methods well justified?
Q7	Were the results extraction methods well described?
Q8	Is there a clear description of the process analysis of the results?
Q9	Were the results well justified?
Q10	Were contradictory information considered when assessing the results?
Q11	Were appropriate metrics used to assess the results?
Q12	Did researchers critically examine their own influence on the results obtained during the collection and analysis of data?
Q13	Are the results clearly described?
Q14	Did the researchers discuss the credibility of the search results?
Q15	Were the approach limitations/disadvantages discussed?
Q16	Is there an adequate discussion of the results both negatives and positives?
Q17	Is the researchers conclusion justified by their results?
Q18	Is there a discussion about the research contributions to the Software Engineering area and for the identification of crosscutting concerns?

3 Selected Studies: A Brief Discussion

Analysing existing aspect mining techniques we notice different characteristics related to: approach, analysis type, granularity level, symptoms of crosscuttingness, manual effort involved and metrics for results evaluation. Also, each technique has its own limitations which impacts on results. This section presents an overview of

Table 3 Data Extraction Questions (Adapted from [14])

Info.1	Study Identifier (ID)
Info.2	Citation
Info.3	Abstract
Info.4	Study type (Qualitative, quantitative)
Info.5	Study goals
Info.6	Study Hypothesis
Info.7	Study scenario (Which software system was used to mine for crosscutting concerns)
Info.8	Results assessment method (Manual/automatic)
Info.9	Technique application (Combined/isolated)
Inf.10	Results evaluation
Inf.11	Case study applied (if applicable)
Inf.12	Analysis method (static, dynamic, structural, behavioral)
Inf.13	Results granularity (methods, classes, code fragments, etc)
Inf.13	Hypotheses for the existence of crosscutting concerns (clone detection, high number of methods calls, etc)
Inf.14	Observed crosscutting concerns symptoms (Scattering, tangling)
Inf.15	User involvement in the identification process of crosscutting concerns (Input data analysis, manual validation of final results)
Inf.16	Findings and conclusions
Inf.17	Approach limitations, vantages and disadvantages
Inf.18	Study relevance
Inf.19	Other important observations

Table 4 Studies Final Classification Based on Punctuation

Punctuation	Meaning
0-10	Very Poor Quality
10.5-13	Poor Quality
13.5-15.5	Good Quality
16-18	Very Good Quality

Table 5 Selected Studies

Approach	Selected Studies
Fan-in Analysis	[26] [28]
Graphs	[23] [34] [24] [33]
Clone detection	[6] [5]
Clustering	[35] [12] [25] [29]
Software Development History	[7] [3]

14 of the 53 selected studies according to their approach: Clone detection (Section 3.1), Fan-in Analysis (Section 3.2), Graph-based techniques (Section 3.3), Clustering Analysis (Section 3.4) and Development-history based techniques (Section 3.5). We show the selected studies in Table 5.

3.1 Clone Detection Techniques

Clones are generally defined as a code fragment identical or similar to another code fragment. They occur most due to copying existing code and modifying it. Bruntink et al. [5] [6] use three different clone detection tools to find the maximal possible number of code fragments implementing crosscutting concerns. Generally, clone detection tools return tuples of code clones. Bruntink et al. however, studied *clone classes* instead of clone tuples. A clone class is a set of code fragments with an equivalence relation between any pair of fragments in the set [21]. The approach aims to evaluate the number of clone classes needed to cover all the extent of a crosscutting concern (i.e., all fragments of code implementing a crosscutting concern). To validate their results, Bruntink et al. depend upon a given set of confirmed crosscutting concerns annotated manually by an expert on the system that the clone detection tools were executed. Although the use of well established coding conventions and idioms required, Bruntink et al. show that clone detection techniques can be used to find most code fragments implementing crosscutting concerns.

3.2 Fan-in Analysis

Marin et al. [26] [28] manually analyze methods with high fan-in value to classify them as crosscutting concerns or not. The fan-in value is computed as the number of calls to a method. Methods with high fan-in value have higher changes to be classified as crosscutting concerns.

Fan-in analysis is a generative approach based on the scattering symptom of crosscuttingness. According to Marin et al. if the same functionality is scattered throughout the code it is reasonable to assume that this functionality is implemented in a helper method that is called by several other methods. Each call contributes to the fan-in value of the helper method. Fan-in analysis generates crosscutting concerns candidates automatically (*seeds*) which diminish the manual effort made to analyze the concerns and crosscutting concerns in a system.

Fan-in analysis restricts the definition of crosscuttingness to methods frequently invoked from other methods. Some crosscutting concerns, however, do not exhibit such behaviour and consequently are pruned by the technique. Also, fan-in analysis potentially returns a high number of false-positives which are filtered manually or by setting the right configurations of automatic tools.

3.3 Graph-Based Techniques

Graphs allow to represent elements of code as a set of nodes and edges which show the relationship between nodes. Several studies selected for our SLR use graphs

to represent code elements and the relationships among them [23] [34] [24] [33]. Zhang and Jacobsen [34] [33] use graphs to propose a model to automate the manual process followed to identify crosscutting concerns. Zhang and Jacobsen consider each node in a graph a different element in a software system (e.g., components, packages, classes and methods) and each edge the relationship among them. Krinke [23] [24], however, generates a different graph for each method. Each node on the graph represents a line of code that implements the method and edges show the execution flow of them.

We notice that, regardless the common use of graphs and the similarity in the approaches to identify crosscutting concerns, each author uses graphs to model different types of data (methods, fragments of code, etc.).

3.4 Clustering Analysis

Clustering Analysis groups data with similar characteristics. From the aspect mining point of view, clustering techniques aim to generate groups, called clusters, each containing all fragments of code implementing a specific crosscutting concern in the system and one extra cluster containing all elements of code that do not implement any crosscutting concern [29].

The generated clusters are analyzed to determine which are part of a crosscutting concern. Cojocar and Czibula [12] mine for crosscutting concerns searching for symptoms of scattering in methods. The authors apply several clustering techniques (*k*-means [19], Fuzzy Clustering [20], Genetic Clustering [1], Hierarchical Agglomerative Clustering [20], *Kam* [32] and *HAM* [31]) to validate their efficiency in grouping all methods implementing crosscutting concerns in a unique cluster. Furthermore, they investigate the number of crosscutting concerns found by each clustering technique. It was observed that most crosscutting concerns are also found by Fan-in Analysis in [26] [28].

3.5 Development-history Based Techniques

Canfora and Cerulo [7] and Breu and Zimmermann [3] observed that crosscutting concerns are inserted in a software system and extended by a series of changes made in later commits. Software version control systems provide a wealth of historical data related to changes in software artifacts, which can be used to identify crosscutting concerns.

In [7] each transaction is a set of files committed by the same author, with the same change history, comments and with time span between commits less than 200 seconds. From these transactions Canfora and Cerulo study the lines of code that were added, changed or deleted to observe the evolution of crosscutting concerns.

Breu and Zimmermann [3] propose a technique that consider only changes that add code related to method calls. The technique is based on dynamic approaches as in [4], although Breu and Zummermann use information from the development history instead of execution traces.

4 Related Works

Kellens et al. [22] provide an in-breadth survey and comparison of techniques and tools for aspect mining. Their work focus on automated techniques for static or dynamic aspect mining. Ceccato et al. [9] provide a qualitative comparison of three aspect mining techniques (Fan-in Analysis, Identifier Analysis and Dynamic Analysis) applying them to JHotDraw. In [10], Ceccato et al. combine Fan-in Analysis, Identifier Analysis and Dynamic Analysis.

Cojocar et al. [11] noticed the importance of considering all the characteristics of aspect mining techniques for comparing and evaluating them. The authors propose a set of criteria to compare existing aspect mining techniques and a set of new evaluation measures to compare mining results. Marin et al. [27] proposes a common framework based on crosscutting concern sorts which allows easier comparison and combination of aspect mining techniques.

Most authors focus on comparing or combining a set of techniques or proposing methods and frameworks to support the comparison and combination of them. Regarding all cited related work, only Durelli et al. [13] provides a systematic review closely related to ours: the systematic review conducted in [13] aims to find out which techniques perform better on identifying crosscutting concerns and how to combine them for improving precision and recall. This work focus on the analysis and synthesis of challenges and issues regarding existing aspect mining techniques to support improvements.

5 Threats to Validity

Our search for primary studies was conducted in four different search engines, however it is possible we have missed relevant studies. The questions to select studies were defined focusing on our search goal and are broad enough to cover different issues. Some questions defined in our data extraction form were not obvious to answer, and most answer required interpretation while reading the papers. To ensure the validity of our interpretations we analyzed several auxiliary sources as technical reports and papers used as reference by the selected studies. Also, we are aware of selecting only classified papers might be restrictive, but they were enough to draw important conclusions.

6 Concluding Remarks

This systematic review aims to expose some existing aspect mining techniques and detect the issues that impact on their efficiency. We analyzed several studies on aspect mining techniques (total of 53) and observed differences related to the type of analysis they execute (static, dynamic, etc.), their granularity level, hypothesis, symptoms of crosscuttingness, manual interference and metrics used to evaluate results.

Most techniques provides a poor evaluation of their results. Those that in fact discuss them show unsatisfactory results. Low precision is the worst problem encountered on the majority of the analyzed techniques which indicates a high number of false-positives. Furthermore, the results analysis is often done manually, which makes the results susceptible to subjectivity.

Poor evaluation of results is accompanied by lack of documentation. most researchers do not affer a detailed documentation of their results but fragments of code as examples of crosscutting concerns and, in general, a weak discussion of the motifs to classify them as crosscutting. Furthermore, some studies perform experiments in proprietary software which difficult repetition by other researchers.

Lack of documentation and experiments in proprietary software prevents establishing common benchmarks for aspect mining. Without benchmarks it is hard to compare or evaluate results forming a vicious cycle that hampers the evolution of the aspect mining research.

The lack of a common definition for what characterizes crosscuttingness is a well known problem in aspect mining. Most researches, however, is concerned in proposing new tools, approaches and techniques. Only a few studies offer valuable information about the main characteristics and evolution of crosscutting concerns in software systems.

Also, it was observed that techniques which strongly depends upon syntactic information hardly identify symptoms of entanglement. An alternative is using semantic information, however, it is difficult to capture and relate the meaning of syntactically different fragments of code.

The directions for aspect mining research that emerged from our SLR are the need for exploratory studies about the nature of crosscuttingness. Knowing why, how, where and when crosscutting concerns emerge in software systems helps to formulate better definitions for crosscuttingness and consequently improve current techniques. Capturing design intent for example, helps to explore how developers make design decisions in the first place. This can be used to evaluate in what circumstances crosscutting concerns are incorporated in code.

The use of semantic information and combined techniques is still faulty. Indeed, our observations revel that every approach relies on different types of analysis to search for crosscutting concerns. Most, however, relies on syntactic/dynamic information and do not try to combine different approaches to suppress possible deficiencies.

Acknowledgement The authors acknowledge the financial support of the Brazilian financial agency São Paulo Research Foundation (FAPESP) – grant (2013/03452-0).

References

1. Babu, G.P., Murty, M.N.: A near-optimal initial seed value selection in k-means algorithm using a genetic algorithm. Pattern Recogn. Lett. **14**(10), 763–769 (1993)
2. Biolchini, J., et al.: Systematic review in software engineering. System Eng. and Comp. Sci. Dept. COPPE/UFRJ, Tech. Rep. ES 679(05), 45 (2005)
3. Breu, S., Zimmermann, T.: Mining aspects from version history. In: Proc. 21st IEEE/ACM Int. Conf. on Automated Softw. Eng, pp. 221–230 (Sept 2006)
4. Breu, S., Krinke, J.: Aspect mining using event traces. In: Proc. 19th IEEE Int. Conf. on Automated Softw. Eng., ASE 2004, pp. 310–315. IEEE Comput. Soc., Washington, DC, USA (2004). http://dx.doi.org/10.1109/ASE.2004.12
5. Bruntink, M., et al.: An evaluation of clone detection techniques for crosscutting concerns. In: Proc. 20th IEEE Int. Conf. on Soft. Maint, pp. 200–209 (Sept 2004)
6. Bruntink, M., et al.: On the use of clone detection for identifying crosscutting concern code. IEEE Trans. on Softw. Eng. **31**(10), 804–818 (2005)
7. Canfora, G., Cerulo, L.: How crosscutting concerns evolve in jhotdraw. In: 13th IEEE Int. Workshop on Softw. Tech. and Eng. Practice, pp. 65–73 (2005)
8. Canfora, G., et al.: On the use of line co-change for identifying crosscutting concern code. In: 22nd IEEE Int. Conf. on Soft. Maint, pp. 213–222 (Sept 2006)
9. Ceccato, M., et al.: A qualitative comparison of three aspect mining techniques. In: Proc. 13th Int. Workshop on Program Comprehension, pp. 13–22 (May 2005)
10. Ceccato, M., et al.: Applying and combining three different aspect mining techniques. Softw. Quality J. **14** (2006)
11. Cojocar, G.S., Şerban, G.: On some criteria for comparing aspect mining techniques. In: Proc. 3rd Workshop on Linking Aspect Tech. and Evol
12. Cojocar, G., Czibula, G.: On clustering based aspect mining. In: Proc. 4th Int. Conf. on Intelligent Comp. Commun. and Processing, pp. 129–136 (Aug 2008)
13. Durelli, R.S., et al.: A systematic review on mining techniques for crosscutting concerns. In: Proc. 28th Annu. ACM Symp. on Appl. Comp. ACM, New York (2013)
14. Dyb, T., Dingsyr, T.: Empirical studies of agile software development: A systematic review. Inf. and Softw. Tech. **50**, 833–859 (2008)
15. Eaddy, M., Zimmermann, T., Sherwood, K., Garg, V., Murphy, G., Nagappan, N., Aho, A.: Do crosscutting concerns cause defects? IEEE Transactions on Software Engineering **34**(4), 497–515 (2008)
16. Filho, F.C., Cacho, N., Figueiredo, E., Maranhão, R., Garcia, A., Rubira, C.M.F.: Exceptions and aspects: The devil is in the details. In: Proceedings of the 14th ACM SIGSOFT International Symposium on Foundations of Software Engineering, SIGSOFT 2006/FSE-14, pp. 152–162. ACM, New York (2006)
17. Garcia, A., Sant'Anna, C., Figueiredo, E., Kulesza, U., Lucena, C., von Staa, A.: Modularizing design patterns with aspects: A quantitative study. In: Proceedings of the 4th International Conference on Aspect-oriented Software Development, AOSD 2005, pp. 3–14. ACM, New York (2005)
18. Greenwood, P., Bartolomei, T., Figueiredo, E., Dosea, M., Garcia, A., Cacho, N., Sant'Anna, C., Soares, S., Borba, P., Kulesza, U., Rashid, A.: On the impact of aspectual decompositions on design stability: An empirical study. In: Proceedings of the 21st European Conference on Object-Oriented Programming, ECOOP 2007, pp. 176–200. Springer (2007)
19. Jain, A.K., et al.: Data clustering: A review. ACM Comput. Surv. **31**(3), 264–323 (1999)
20. Jain, A.K., Dubes, R.C.: Algorithms for Clustering Data. Prentice-Hall Inc, Upper Saddle River (1988)
21. Kamiya, T., et al.: Ccfinder: a multilinguistic token-based code clone detection system for large scale source code. IEEE Trans. on Softw. Eng. **28**(7), 654–670 (2002)
22. Kellens, A., et al.: A survey of automated code-level aspect mining techniques. In: Trans. on Aspect-Oriented Softw. Develop. IV, pp. 143–162. Springer (2007)

23. Krinke, J.: Mining control flow graphs for crosscutting concerns. In: Proc. 13th Work. Conf. on Reverse Eng, pp. 334–342 (Oct 2006)
24. Krinke, J.: Mining execution relations for crosscutting concerns. Softw. IET **2**(2), 65–78 (2008)
25. Maisikeli, S., Mitropoulos, F.: Aspect mining using self-organizing maps with method level dynamic software metrics as input vectors. In: Int. Conf. on Soft. Tech. and Eng., vol. 1, pp. V1–212–V1–217 (Oct 2010)
26. Marin, M., et al.: Identifying aspects using fan-in analysis. In: Proc. 11th Work. Conf. on Reverse Eng., pp. 132–141 (Nov 2004)
27. Marin, M., et al.: A common framework for aspect mining based on crosscutting concern sorts. In: Proc. 13th Work. Conf. on Reverse Eng
28. Marin, M., et al.: Identifying crosscutting concerns using fan-in analysis. ACM Trans. Softw. Eng. Methodol. **17**(1), 3:1–3:37 (2007)
29. McFadden, R., Mitropoulos, F.: Aspect mining using model-based clustering. In: Southeastcon, 2012 Proc. IEEE, pp. 1–8 (March 2012)
30. McFadden, R., Mitropoulos, F.: Survey and analysis of quality measures used in aspect mining. In: Southeastcon, 2013 Proc. IEEE, pp. 1–8 (April 2013)
31. Serban, G., Cojocar, G.S.: A new hierarchical agglomerative clustering algorithm in aspect mining. In: 3rd Balkan Conf. in Informatics, pp. 143–152 (Sept 2007)
32. Serban, G., Moldovan, G.: A new k-means based clustering algorithm in aspect mining. In: Proc. 8th Int. Symp. on Symbolic and Numeric Algorithms for Sci. Compu., pp. 69–74 (Sept 2006)
33. Zhang, C., Jacobsen, H.A.: Mining crosscutting concerns through random walks. IEEE Trans. on Softw. Eng. **38**(5), 1123–1137 (2012)
34. Zhang, C., Jacobsen, H.A.: Efficiently mining crosscutting concerns through random walks. In: Proc. 6th Int. Conf. on Aspect-oriented Softw. Develop., AOSD 2007, pp. 226–238. ACM, New York (2007)
35. Zhang, D., et al.: Automated aspect recommendation through clustering-based fan-in analysis. In: 23rd IEEE/ACM Int. Conf. on Automated Softw. Eng., pp. 278–287 (Sept 2008)

ValiPar Service: Structural Testing of Concurrent Programs as a Web Service Composition

Rafael R. Prado, Paulo S.L. Souza, Simone R.S. Souza, George G.M. Dourado and Raphael N. Batista

Abstract The testing of concurrent programs is essential to ensure the quality of such programs. One of the main challenges of such testing activity is to provide tools that enable it at a reasonable operational cost. The service orientation paradigm provides guidelines for the development of tools as services, addressing such requirement. The division of the structural testing of concurrent programs into services faces fundamental challenges. The objective of this paper is to provide structural testing of concurrent programs as a Web service composition. We divided this monolithic structure by defining the concepts, relations and parameters of the structural testing of concurrent programs to support this activity using a service composition. Our main contributions are in the definition of Web services for the structural testing of concurrent software, focusing on the service contracts and capabilities. The developed services present flexible contracts and capabilities that can be reused in different contexts.

Keywords Web services · Structural testing tools · Concurrent software

1 Introduction

Software testing is essential for software development. Sequential programs are already assisted with a broad range of approaches to ensure quality. Concurrent programs, on the other hand, lack resources for this same purpose. The TestPar project supports the structural testing of concurrent programs by providing testing models, criteria and tools. Structural testing of concurrent programs is an approach based on

R.R. Prado(✉) · P.S.L. Souza · S.R.S. Souza · G.G.M. Dourado · R.N. Batista
Department of Computer Systems, University of Sao Paulo, ICMC, Sao Carlos, Brazil
e-mail: {rafaelrp,pssouza,srocio,georgemd,rbatista}@icmc.usp.br

The authors acknowledge Brazilian funding agencies FAPESP, CAPES, CNPq for their financial support (processes 2013/05750-8 and 2013/01818-7).

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_51

the knowledge of the software's internal implementation, i.e., it considers control, data and synchronization flows of the program. This approach helps to reveal concurrent bugs by requiring the coverage of all possible synchronization flow among processes and threads.

One of the main challenges of the testing, in general and in particular for concurrent programs, is to provide tools that enable the testing activity at a reasonable operational cost. This challenge involves aspects such as maintainability, reusability and performance. Indeed, the testing tool should be easy to maintain and execute the test in a reasonable response time. Also, the tool's components should be reusable to reduce development costs.

The service orientation paradigm provides guidelines for the development of tools as services addressing such requirements. Services can handle different parts of the testing activity such as testing analysis, evaluation and testing execution. Every resulting service exposes a distinct contract designed to solve specific tasks effectively. This division encourages the service reuse and makes it easier to maintain the overall infrastructure. The division of the structural testing of concurrent program into services, however, faces fundamental challenges, because it follows a monolithic vision that unifies testing models and criteria with programming languages and communication and synchronization paradigms into a testing tool that offers the required support.

The objective of this paper is to provide structural testing of concurrent program as a Web service composition. The services must aim for service orientation aspects to reduce the overall cost of the testing, making it easier to develop new testing tools and maintain existing ones. This paper is different from our recent publications, since it is the first one presenting results of the use of the testing of concurrent programs as a service on the Web. We divided this monolithic structure by defining the concepts, relations and parameters of the structural testing of concurrent programs to support this activity using a service composition. ValiPar Service supports the structural testing of concurrent programs using six services. While the main service handles the entire process, the other ones are specialized on testing activities and can be reused.

The main contributions of this work are in the definition of Web services for the structural testing of concurrent software. The development of this composition emphasizes service orientation principles and the remote execution of the whole structural testing. In this way, the developed services present flexible service contracts and capabilities, that can be reused in different contexts.

2 Structural Testing of Concurrent Software

Testing of concurrent software is not trivial due to the communication and synchronization among threads and processes. These features must be considered during the testing. The TestPar project [9, 10, 11] supports the verification and validation to ensure a quality level of concurrent programs by providing testing models, criteria and tools for the structural testing of concurrent programs. Testing model is represented on a Parallel Control Flow Graph (PCFG) [9], which summarizes the control, data and

synchronization flows of the tested program. Required elements are based on this testing model and address important aspects such as the definition and use of variables, control flow paths, and communications and synchronizations. These required elements are later evaluated using structural testing criteria based on test cases.

In a typical scenario, the structural testing of concurrent programs is divided into a set of testing activities. The source code is imported into the testing tool, instrumented to produce trace files that describe the program execution and fully analyzed for the PCFG generation. Next, the required elements are generated based on the testing model information (PCFG). Test cases are generated and executed, producing trace files. Finally, the coverage analysis verifies the required elements coverage of every execution and evaluates the coverage for selected testing criteria. New test cases are generated, executed and evaluated interactively until a bug is revealed, or testing criteria target coverage is reached. ValiPar testing tool implements these testing activities for message passing and shared memory paradigms using different languages and libraries as desktop tools. ValiMPI [4] allows the structural testing of message passing paradigm using MPI. ValiPThread [8] supports the testing of shared memory programs using C language and PThreads library. ValiPar Java [7] considers both synchronization paradigms using Java and its standard library.

The development of these desktop testing tools has raised important issues, also observed in the literature, which must be solved to improve the testing process. The current tools have a low level of reusability, i.e., they cannot be effectively used in other contexts. Reusability is a desirable feature because several testing activities are present in more than one process. Another observed issue is the maintainability. It is hard to maintain all testing tools with bug fixes and testing model updates because every testing tool has particular features that must be handled differently. Testing tools can also depend on specific resources such as compilers, libraries and virtual machines. Also, a bug fix can take a long time to reach all the testing tool users. All these factors slow down the tool's development and the test conduction. Performance issues must also be solved to provide a reliable testing process in a reasonable response time. This requirement represents a significant challenge because the testing of concurrent programs often requires multiple executions of test cases to cover determined required elements (e.g. a synchronization edge). The use of desktop testing tools in a single machine does not scale in scenarios with bigger concurrent programs that must be executed thousands of times. Finally, the availability of the testing environment is also important. This aspect includes the reliable access to test infrastructure, which helps testers to keep track of the testing execution and trigger new test case executions. Desktop testing tools are limited in this aspect because the test is executed locally on a machine that is a single point of failure. Moreover, the access and operation of test artifacts is, in general, possible with direct access to the machine.

These issues can be addressed with the use of service orientation. In this scenario, the testing tool's features are organized as services remotely accessed by clients that can be testers or other services. The organization allows the definition of interfaces that are reusable in a variety of compositions. It can also reduce the maintenance cost because of the centralized processing and the performance can be improved independently of the client.

3 The ValiPar Service

ValiPar Service brings service orientation paradigm benefits to the structural testing of concurrent program context. The service was developed based on a study of all structural testing of concurrent programs domain. We have adopted the OntoTest [1] ontology, i.e., a formal description of the testing domain, to this task because of its domain coverage and compliance with existing standards.

OntoTest describes the concept of testing steps and testing resources. A testing resource is used by testing steps to support testing strategies and processes. It can be a human, hardware, supporting system and primary, organizational and support modules. A testing step is a set of testing primary, organizational and support activities that consume and generate artifacts and interact with other steps to help the conduction testing strategies and processes. The structural testing of concurrent programs focuses on primary activities and resources.

The development of services structures the testing domain as a set of capabilities and contracts. Every capability represents a testing activity and contracts give access to the provided capabilities. ValiPar Service implements the following capabilities: **Source code acquisition:** The source code is imported into the testing tool and stored; **PCFG generation:** Source code is fully analyzed and described as testing model elements (i.e., the PCFG); **Source code instrumentation:** The source code is submitted for a transformation process that inserts instructions to produce trace files reporting control, data and synchronization flow events. It also inserts instructions to allow controlled execution; **Required element generation:** Required elements are based on the testing model information; **Test case generation:** Test cases are based on manual or automated approach. This work focuses on the manual approach but does not prevent the support of the automated one in future extensions; **Test case nondeterministic execution:** Test cases are executed nondeterministically, producing trace files; **Test case deterministic execution:** Test cases are executed systematically in a deterministic way to cover a selected number of execution variants, i.e., alternative executions for a given test case and input data; and **Structural criteria coverage analysis:** Required elements are evaluated based on the execution of trace files and structural testing criteria coverage is computed;

Figure 1 shows the ValiPar Service implementation. It is a composition of 6 services. **ValiPar Service** supports the whole structural testing step. It acts as a proxy service receiving client requests and delegating tasks to specialized services, so the client sees the service composition as a single service. It uses a set of five services: ValiInst Service, ValiElem Service, ValiEval Service, ValiExec Service and ValiSync Service. Every specialized service represents specific testing resources and is used in different testing activities.

ValiInst Service manages the source code instrumentation and PCFG generation. Initially, the client uploads the program's source code and additional information such as what are the processes and their *main()* functions. It is also possible to specify the used synchronization primitives regarding testing model information. This information is essential to the static analysis of concurrent programs. Next, the client can request instrumented source code and PCFG.

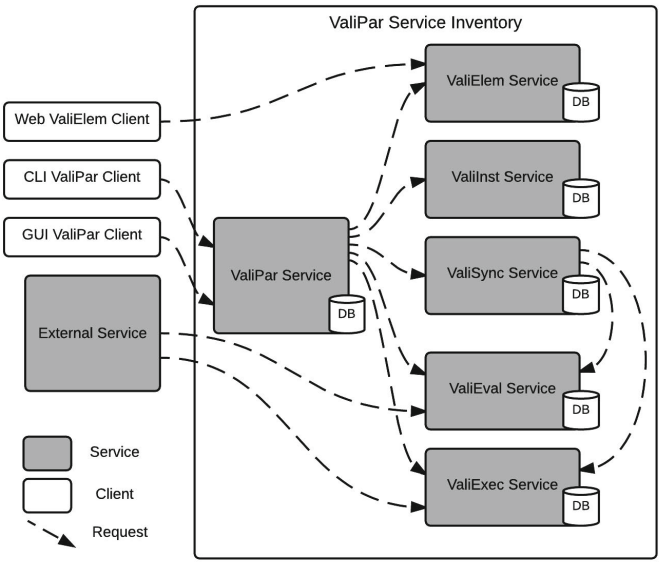


Fig. 1 ValiPar Service inventory with interactions among services and clients.

ValiElem Service generates the required elements. The client uploads the PCFG and requests the generation of determined required elements. **ValiEval Service** evaluates the coverage of structural testing criteria based on execution trace files. Firstly, the client uploads the required elements. Next the client can provide execution trace files to be evaluated. In this process, the service marks all covered required elements with *covered* status. After the evaluation of a trace file, the client can request the structural testing criteria coverage. In this case, the service verifies the coverage of specific required elements and returns the total coverage for every criterion.

ValiExec Service executes test cases. The client uploads the instrumented source code and creates test cases specifying the command-line arguments, input and expected output for each program's process. Then executions are created based on a test case and additional parameters such as the mode of execution (deterministic and nondeterministic) and the time limit. The service executes the test case and makes the result available. The service contract also supports the creation of variant execution. In this case, the client must upload a variant specification and create an associated deterministic execution.

ValiSync Service generates and executes variants based on selected test cases. The client informs the location of the program's evaluation. Next, it creates test cases in the same way as in ValiExec Service. Finally, the client can trigger the generation and execution of all test cases. In this process, the test cases are executed nondeterministically. The service then generates new execution variants and executes them in a deterministic way iteratively. Optionally, it uses the current testing coverage

to guide the variant generation. This service reuses ValiExec Service and ValiEval Service to complete its tasks. In this sense, ValiExec Service executes all test cases and ValiEval Service evaluate the executions to provide the current coverage to the related required elements.

The implemented services group capabilities according to its coupling. For instance, PCFG generation and source code instrumentation are clustered in the same service (ValiInst Service) due to their high dependency. Indeed, both capabilities must follow the same strategy to produce compatible artifacts, i.e., testing model elements such as the use of variables, control flow nodes and synchronization edges must be present and correspondent in both representations. The required element generation, on the other hand, can be separate from criteria coverage analysis. Although the coverage analysis depends on required element, its generation can be performed by another tool, as long as the contract remains the same. In similar fashion, the required element generation can be useful for other tools and testing activities.

We used the REST architectural style to implement the defined services, which allows the organization of the testing tool as a collection of resources accessible via a uniform interface and URIs. Thus, it is possible to map the testing artifacts and capabilities as resources and their relations. The architecture has two layers to handle service contracts and capabilities. The contract layer handles which resources must be available to the client and how the client can interact with them. Currently, the client can interact with the resource in four levels: “no-interaction”, “read-only”, “upload-and-read”, “generation-and-read”. In this sense, a required element generation service would enable the PCFG and required element resources. Because it does not generate the PCFG locally, this resource would be enabled with “upload-and-read” interface. The required element resource would be enabled with “generation-and-read” interface.

The service capability layer manages which operations are available and how they are performed. Every operation can be internal and external. The execution of an internal operation can be sequential or parallel. While the sequential mode uses only local computational resources, the parallel mode uses a distributed architecture to execute tasks. An external operation, on the other hand, delegates the task to specialized services.

This architecture allows the creation of services to meet specific needs. It is possible to deploy the services in a different environment with or without distributed architecture, making it easier to try a service without installing a complex infrastructure. This feature can also be used to execute experiments to evaluate the performance of different parallel execution setups if compared to the sequential mode. The external mode is used to build custom services that depend on other specialized services. Based on this, it is possible to create composite services that handle specific tasks and reuse existing service implementations.

Additionally, the layered and modular organization supports the creation of alternative services that group different capabilities to offer custom inventories. Depending on the context of the structural testing of concurrent programs, the PCFG generation can be offered as a separate service if no instrumented source code execution is not required. In the same way, if the related testing process only reuses

required element generation and structural criteria coverage analysis, this capabilities can be offered in a single customized service.

The definition of a service is done statically using external configuration files. Every contract and capability can be enabled and further defined individually. This organization encourages the tool reuse and minimize the interaction of testers with service implementation.

As shown in Figure 1, the service inventory can be used by Web clients (from a browser, for example), CLI (command line interface) clients and GUI (graphical user interface) clients. Moreover external services can make use of any of the implemented services in the inventory. The use of HTTP as the application protocol makes the service available for any programming language with HTTP library in a variety of environments (e.g. Web browsers and operating systems). ValiPar Service makes resources available through URIs. The client uses HTTP verbs (GET, POST, PUT, DELETE) to access or modify them. Using them, the tester has the control over the testing process, making it possible to customize the operations sequence. The alternative would be to specify all the artifacts in a single command, but this restricts the interaction between the client and the service. ValiPar Service contract level enables, for example, the creation of new test cases or executions in the middle of the testing process.

4 Experimental Evaluation

We have evaluated the performance of the ValiPar Service in relation to the response time to execute the testing activities. Our objective was to verify the impact of the organization of testing functionalities as services on the Web.

We have deployed the six services in six different virtual machines (VMs) and the distributed infrastructure for the parallel execution in eight VMs. The VMs are part of a Gigabit Ethernet network and each physical machine is an Intel(R) Core(TM) i7-4790 3.60GHz processor with 32 GB of RAM. Each VM has 8GB RAM and uses Ubuntu 14.04.1 LTS and OpenJDK Java 1.8.

The evaluation includes the desktop testing tool (Desktop), as the parameter and two versions of the service composition: Sequential and Parallel. The sequential version uses the sequential mode for all internal processing on specialized services. The parallel one uses the parallel mode for internal processing related to testing case execution, variant generation and coverage evaluation on ValiEval Service, ValiExec Service and ValiSync Service.

The experiment consists of a general workflow execution using each tool for the testing of a set of concurrent programs. The selected workflow includes the program upload, followed by the source code instrumentation, PCFG generation, required elements generation and test case creation. In a second step, the variant generation triggers test case execution (deterministic and nondeterministic) and criteria evaluation, iteratively.

Table 1 describes the selected programs. It is a total of 9 concurrent programs implemented in Java that are semantically representative. Their implementation are available at http://testpar.icmc.usp.br/benchmarks. The programs contain synchronization primitives with different semantics, have a varied number of processes and threads and exercise various synchronization and communication patterns. They also require different quantities of variant executions to achieve the target criteria coverage. Such variety is necessary to show scalability and capacity of the tool on the testing of different types of concurrent programs.

Table 1 Concurrent programs used in experiments and the response time (in seconds) of the test activity using the desktop (D), sequential (S) and parallel (P) versions. Programs contain a variety of synchronization primitives and number of processes/threads for distinct interleavings.

Programs	Paradigm	LOC	Processes	Threads	Syncs	Variants	D (s)	S (s)	P (s)
NB	MP	71	3	3	6	12	11.1	12.7	13.6
PG		180	4	4	24	54	35.7	40.4	31.5
GT		201	3	3	12	58	56.6	64.2	42.1
MC	SM	43	1	3	20	4	64.5	68.9	91.4
RW		145	1	4	52	606	333.3	403.2	390.4
HO		113	1	4	27	6	6.9	10.3	15.2
MS	MP/SM	102	2	3	4	0	5.7	7.61	7.9
TR		246	3	12	96	12	1716.6	1931.5	1351.7
MA		214	5	13	68	34	919.5	1214.1	562.4

The programs NB (non_blocking_bsend), PG (parallel_gcd) and GT (gcd_two_slaves) use message passing. NB uses non-blocking primitives, PG calculates GCD (Greatest commom divisor) using blocking primitives, and GT calculates GCD, but uses other blocking primitives inside a loop. The programs MC (micro_sm), HO (h2o) and RW (readers_writers) use the shared memory. MC synchronizes the memory access between two threads using a binary semaphore, HO simulates the water molecule formation using semaphores, and RW implements the Readers-Writers problem using binary semaphores. Finally, the programs MS (master_slave), TR (token_ring_file) and MA (matrix) use both paradigms. MS implements the Master-Slave problem, with semaphores and blocking communication, TR simulates the token ring topology using both blocking and non-blocking primitives, and MA implements a matrix multiplication.

The columns “D (s)”, “S (s)”, “P (s)” show the response time for the execution of benchmarks on Desktop, Sequential and Parallel versions, respectively. As expected, the results show that the sequential ValiPar Service always has a greater response time if compared to the Desktop version. The network latency and other processes such as message serialization and deserialization, which are necessary to the service implementation, add an overhead to the execution.

The parallel ValiPar service, despite the network and message processing, presents different behaviors depending on the tested program. We have observed that the response time is inversely proportional to the number of necessary variant executions

(Table 1). The execution of other programs presents this same behavior. Indeed, the distributed infrastructure is only efficient with a big load of concurrent tasks. A small load may not compensate the cost of tasks distribution, increasing the resulting response time.

The results show that the use of implemented service adds an overhead to the test activity. Indeed, the cost of network transfer, serialization and task distribution adds a response time overhead that varies according to the nature of the tested program. However, the use of a distributed architecture can minimize this effect and reduce the total response time for the testing.

5 Related Works

A wide range of testing tools has used service orientation to deliver testing functionalities. None of them, however, approaches the structural testing of concurrent programs. We have also observed that, in general, there is no focus on the service contracts for reusability purposes.

Cloud9 is a symbolic execution engine that runs as a Web service [2] deployed in public cloud infrastructure. The development of architecture for parallel symbolic execution has showed improvements with relation to the sequential equivalent. This fact reinforces the web service role in delivering computational infrastructure, also verified for ValiPar Service.

JaBUTiService is a Web service tool for the structural testing of sequential Java programs [3]. It presents an interface for the testing that allows the execution of different operations separately, which can be useful for reusability. However, it does not deliver the test case execution, so the client must perform this particular activity locally and upload the result. The authors have also pointed out the overhead caused by the network latency and message parsing. The tool is implemented as a stateful service i.e., keeps the client's current state loaded on the server side. This fact can negatively affect its scalability and imposes a pre-defined operation sequence, limiting the custom use of the service. ValiPar Service uses the stateless approach, following the REST architectural style.

Other testing services approach the same issues pointed out by these services. They handle aspects such as performance, scalability and confidentiality. RTaaS [5] implements a platform to provide regression test selection and Path Crawler Online [6] provides structural unit testing, both as a service.

6 Conclusion

We have developed the structural testing of concurrent programs as a service composition to take advantage of the benefits offered by the service orientation paradigm. Our service's internal architecture is flexible and reusable. The separation in layers allows the creation of all implemented services in the ValiPar Service inventory. It

also supports the creation of services for specific purposes and the deploy of the services in custom infrastructures.

Our experiments show the overhead imposed by network latency, message serialization and deserialization and task distribution. They evince the feasibility of the testing of concurrent programs as a service composition despite the identified overhead. The distributed architecture makes ValiPar Service more scalable, resulting in smaller response times when compared to the desktop tool.

ValiPar Service inventory improves reusability aspects of the testing activity. Indeed, it not only gives a complete set of capabilities for the structural testing of concurrent programs but also provides contracts that allow further customizations and the use of the services in custom compositions. It also improves maintainability, performance and availability aspects. The testing service can make use of an infrastructure (such as a cluster or a cloud) to distribute testing tasks, potentially reducing the overall testing cost (i.e., response time). The centralized nature of the services allows the tool to update with reduced cost to the clients (testers). Finally, the services are available through the use of clients in environments such as browsers or command line interfaces.

References

1. Barbosa, E.F., Nakagawa, E.Y., Riekstin, A.C., Maldonado, J.C.: Ontology-based development of testing related tools. In: SEKE 2008, pp. 697–702 (2008)
2. Ciortea, L., Zamfir, C., Bucur, S., Chipounov, V., Candea, G.: Cloud9: a software testing service. SIGOPS Oper. Syst. Rev. **43**(4), 5–10 (2010)
3. Eler, M., Endo, A., Masiero, P., Delamaro, M., Maldonado, J., Vincenzi, A., Chaim, M., Beder, D.: Jabutiservice: A web service for structural testing of java programs. In: 33rd Annual IEEE SEW, pp. 69–76 (2009)
4. Hausen, A.C., Vergilio, S.R., Souza, S.R., Souza, P.S., Simao, A.S.: A tool for structural testing of mpi programs. In: 8th IEEE LATW (March 2007)
5. Huang, S., jie Li, Z., Liu, Y., Zhu, J., Xiao, Y.H., Wang, W.: Regression testing as a service. In: IEEE Int. Conf. on Web Services, pp. 484–491 (July 2011)
6. Kosmatov, N., Williams, N., Botella, B., Roger, M.: Structural unit testing as a service with pathcrawler-online.com. In: IEEE 7th SOSE, pp. 435–440 (2013)
7. Prado, R.R., de Souza, P.S.L.L., Dourado, G.G.M., Souza, S.R.S., Estrella, J.C., Bruschi, S.M., Lourenco, J.: Extracting static and dynamic structural information from java concurrent programs for coverage testing. In: XLI Latin American Computing Conference (CLEI), pp. 659–666. Arequipa (2015)
8. Sarmanho, F.S., Souza, P.S.L., Souza, S.R.S., Simao, A.S.: Structural testing for semaphore-based multithread programs. In: ICCS. LNCS 5101, pp. 337–346 (2008)
9. Souza, P.S., Souza, S.S., Rocha, M.G., Prado, R.R., Batista, R.N.: Data flow testing in concurrent programs with message passing and shared memory paradigms. Procedia Computer Science **18**, 149–158 (2013). ICCS 2013
10. Souza, S.R.S., Souza, P.S.L., Zaluska, E.: Structural testing for message-passing concurrent programs: an extended test model. Concurrency and Computation: Practice and Experience **26**, 21–50 (2014)
11. Souza, S.R.S., Vergilio, S.R., Souza, P.S.L., Simao, A.S., Hausen, A.C.: Structural testing criteria for message-passing parallel programs. Concurrency and Computation: Practice and Experience **20**(16), 1893–1916 (2008)

A Model-Driven Solution for Automatic Software Deployment in the Cloud

Franklin Magalhães Ribeiro Jr., Tarcísio da Rocha, Joanna C. S. Santos and Edward David Moreno

Abstract Through virtualization, cloud computing offers resources that reduce the costs in the institutions that use hardware and software resources. In this paper, we present a model-based approach to automatically deploy software in the cloud. To evaluate our approach, we conducted an experiment in an IT company in which their software developers used our solution instead of manually deploying software in the cloud. After that, they answered a survey, so we could investigate the following metrics: maintainability, learnability and reduction of developer's workload to deploy software services. The results showed that our solution presented a positive impact of at least of 25% percent for all the metrics. Moreover, since our approach relies upon UML models, it requires less effort to deploy the services as well as it can be used by any professionals that have basic skills about UML.

Keywords Cloud computing · Automatic software deployment · Model-driven deployment · UML

1 Introduction

The resources offered by cloud computing decrease the costs associated with high data processing demands [13]. Cloud resources have been applied in many fields of knowledge (such as as astronomy, genetics and chemistry) that use exhaustive

F.M. Ribeiro Jr.(✉)
Universidade Do Tocantins (Unitins), Palmas, TO, Brazil
e-mail: franklin.mr3@gmail.com

T. da Rocha · E.D. Moreno
Departamento de Computação, Universidade Federal de Sergipe, São Cristóvão, SE, Brazil

J.C.S. Santos
Department of Software Engineering, Rochester Institute of Technology, Rochester, NY, USA

J.C.S. Santos—The author is sponsored by CAPES Brazil to pursue a MS at RIT

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_52

search algorithms to decode certain structure types [4]. Through virtualization, which abstracts the cloud physical layer in order to offer hardware resources according to demands, cloud computing also has an approach based on Software as a Service (SaaS) [18]. In this approach, clients can request a software service in the cloud provider, which was deployed by the provider itself, or deploy their own service in the cloud [16].

Since cloud environments have their own software architecture, the deployment of a software system in the cloud requires the reconstruction of many existing requirements [4]. Moreover, in environments in which the software deployment in the cloud is costly, introducing the automatic deployment of services and applications would reduce the developer's workload and, as such, it brings advantages.

Thus, in this work we developed and analyzed a model-driven solution to automatically deploy software in the cloud. This solution allows not only software developers, but anyone who has basic skills in UML (Unified Modeling Language) [15], to deploy software in the cloud with a high-level abstraction.

This paper is organized as follows: in Section 2 we explain the requirements needed to be addressed when deploying software in the cloud; in Section 3 we discuss the related work; in Section 4 we elaborate on our proposed solution; in Section 5 we explain the solution's architecture; in Section 6 we show the results we obtained by applying the solution into an IT company; lastly, in Section 7, we make our final considerations about this work.

2 Deploying Software in the Cloud

For a software to properly execute in a cloud architecture, it is required to deploy not only the software itself but also its dependencies (i. e., other services that the software depends on when it is running) [8, 18]. To fulfill this requirement, each developer, who needs to deploy an application in the cloud, has to acquire an access key to the cloud provider, select a machine instance in the provider that supports the application as well as configure and install the virtual machine. After that, the developer needs to deploy the application and its dependencies (in a proper order) in the selected virtual machine [2, 4].

A more efficient way to deploy software in the cloud environment is through automating that process, since the configuration details could be abstracted at a higher level thereby decreasing the human efforts to perform it. Another advantage of the automatic deployment of software in the cloud is the elimination of manual tasks which are prone to error such as: software stack configuration, authentication between virtual machines, specification of dependencies between service components and between services, definition of temporal and spatial dependencies and, lastly, analysis of the resources in the cloud environment in order to identify whether the node fulfills the minimum requirements of the service [11].

3 Related Work

After an extensive investigation of the current state-of-art of approaches for automatic software deployment in the cloud, we selected some of them to describe and analyze. For example, the solutions presented in Vega [6], Wrangler [10] and Disnix [3] used manual codification through scripts, which required more time to deploy the software. Besides that, according to [19] and Fazziki et. al. [9] the disadvantage of such deployment approach is that it increases the costs associated with the codification and human efforts.

We also noted that D&C Based [4], SDO [12] and User-Level Deployment [20] automated the deployment of services in the cloud with the use of programming languages. Comparing to the approaches presented in Vega [6], Wrangler [10] and Disnix [3], those approaches are more advantageous in terms of the time effort in the deployment process.

In addition, Virtual Models [11] and Disnix [3] presented semi-automated mechanisms for deployment. In other words, these approaches still require that certain steps are performed manually during the deployment process.

The solution proposed by Ardagna et. al. [1] presented an approach partially models oriented and semi-automated for software deployment. It requires some level of comprehension from the final user about the details of the cloud structure and a heavy information load in the models for deployment demanded to the user.

The solution presented in this work is based on models for deploying software in the cloud automatically. Our solution only requires the developer to have knowledge about the access key and service provider name, since the specific details are abstracted. The goal is to deploy the software at a higher level of abstraction in order to reduce the human efforts and the time spent in performing the deployment tasks, since the model-based approach is a better way to increase the developer's productivity [9].

4 Proposed Solution

This section explains how the deployment models were defined and adapted to our solution as well as the solution's architecture.

4.1 Solution Workflow

This solution proposes that the developers have the services they need to deploy as well as to create the software deployment UML models. These models define all the information required for the deployment (virtual machines, services, applications, dependencies, operating systems of the virtual machines, services repositories and

virtual machines, databases, services provider and access key) as parameters without the need of coding of the configuration of any virtual machine.

After the creation and definition of the models with their respective input parameters, the developer provides them as an input of the system. This triggers an automatic deployment, which encompasses four steps: two for interpreting the UML models, one for creating the software stack and another for automatically generating code. Regarding the steps related to UML models interpretation, the first one is about interpreting a specific model (which contains the elements related to the cloud provider as well as the repository) and the second one is the interpretation of the general model (that has the elements for specifying virtual machines, operating system, services, databases, dependencies and applications). The software stack is defined as the stack of dependencies between services, which were discovered in the interpreting phase of the general model.

Hence, firstly, the system interprets the general model by collecting data within a file which has a format similar to XML and contains the information specified by the developer in the model. During this collecting process, the system instantiates a list of objects of virtual machines, dependencies, operating systems and services. Subsequently, during the software stack creation step, the system binds the services to the corresponding virtual machines. After that, it individually associates the dependencies with the services in each virtual machine. When these two steps are completed and all data about virtual machines and services were collected, then the step of interpreting of specific models is started. In this step, an XML-like file (.uml format) is read in order to collect data and, finally, automatically generate the code to deploy the software in the cloud.

4.2 Deployment Models Definition

The solution has two UML deployment diagrams as inputs: the *general model* (which is independent of a cloud service provider) and the *specific model* (which has the particular aspects related to the cloud infrastructure). Figures 1a and 1b shows examples of the general model and specific model, respectively. In the example model shown in Fig. 1a there are all the necessary features (services, dependencies between services, operating system, virtual machine, application, databases and optional fields for each service, such as service version, the service directory, etc.) for the software deployment independent of the cloud provider. The features exposed in these general models are used by the specific model (Fig. 1b) to collect the necessary information for the effective implementation of the software in the cloud, such as: access key, virtual machine service repository and the VM instance in the cloud.

Therefore, for an effective modeling of the specific model, developers need to create the general model first, since the specific model needs the information of the virtual machine. Such relationship between the specific and the general model can be noticed in the example shown in Fig. 1b in which the Virtual Machine node corresponds to the same node represented in the example shown in Fig. 1a.

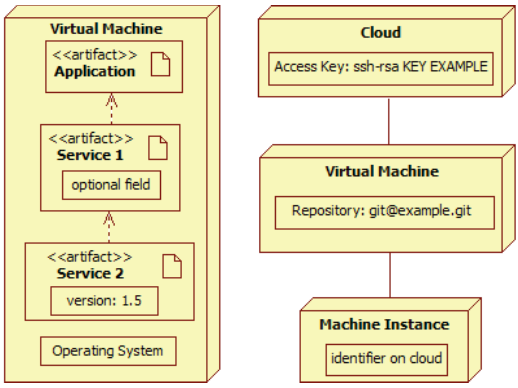


Fig. 1 (a) General model. (b) Specific model

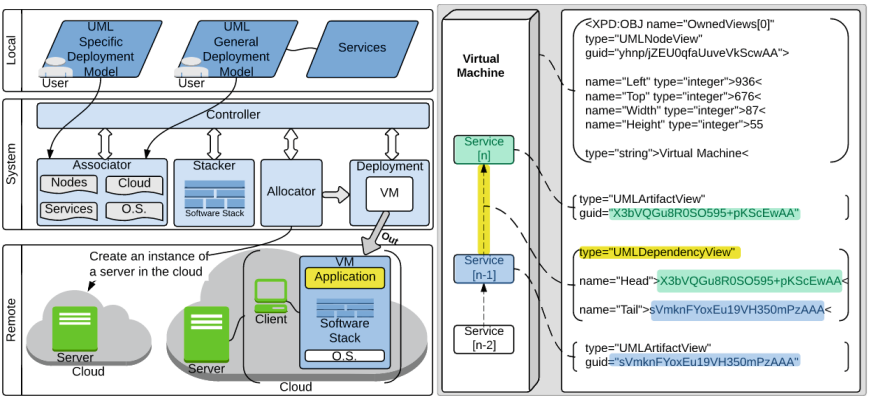
When designing a practical solution to be applied on a large scale environment (with multiple VMs), considering that the developer needs the deployment of multiple applications, the use of models can decrease the work effort to deploy software by reducing coding tasks and increasing the deployment productivity. Hence, our solution allows the creation of multiple VM nodes in the solution model (providing scalability).

5 Solution Architecture

The implementation of the proposed architecture is divided into three views: system, local and remote. This solution is composed of five modules: *Controller*, *Associator*, *Stacker*, *Allocator* and *Deployment*. The Fig. 2a shows the three views of the proposed system architecture for model-based software deployment in the cloud (i) System: showing the five modules of the system, (ii) Location: corresponding to the user's view (which uses the UML deployment diagrams as inputs) and (iii) Remote: it includes the creation of a Chef [5] server instance in the cloud and the final stage of deployment, corresponding to the software installation in the cloud.

5.1 Controller Module

The controller module manages the four other modules, acting as an intermediary in the communication between them. It is responsible for (i) sending a list of services obtained from the Associator to the Stacker and a list of dependency



(a) (b)

Fig. 2 (a) Software deployment solution architecture. (b) Example of identifying the attributes of a model

relationships between services (software stack) from the Stacker to the Deployment Module, (ii) enabling communication of the Allocator with the cloud, (iii) notifying the Allocator upon the server instance creation in the cloud, (iv) informing to the Deployment Module which operating system must be installed on the VM, (v) reporting to the Deployment Module which software stack should be allocated to the respective VM and (vi) managing the order of information transmission between the modules and the tasks performed.

5.2 Associator Module

The Associator Module is responsible for interpreting the data within UML models. Its execution routine is divided into two stages: (i) the General Model interpretation and (ii) interpretation of the Specific Model, in which the input files (metadata from UML models) are read line by line at each stage. For each line read in the file, the algorithm within this module searches for the substring "OwnedViews", which can refer to either a UML component with the "artifact" type (service), or a node (VM), or an association or dependency. When the searched substring is found, the method search for another substring in the same line of the file which can be: (i) "UMLNodeView" (represents a node), (ii) "UMLArtifactView" (representing an artifact), (iii) "UMLAssociationView" (indicating an association and used in the interpretation of the specific model) and (iv) "UMLDependencyView" (representing a dependency). Fig. 2b shows an example of this metadata. Once the attributes related to the UML components are collected, the algorithm instantiates objects with

their attributes (GUID, types, and coordinates) and creates lists of these objects for each type of UML component (such as list of node objects, virtual machines, cloud providers, instance machine, services, operating systems and attributes of services).

5.3 Stacker Module

To define the software stack and specify the services belonging to their respective virtual machines, we developed the Stacker module. This Stacker is divided into three steps. In the first step, we establish the dependencies to the services through the two lists of objects obtained from the Associator module. The algorithm of this module then scans the entire list of objects for each dependency artifact (service). For example, when the "guid" identifier (see Fig. 2b) of the service object with an *x* index from the list is the same as the *head* attribute identifier of an object (dependence) in the dependency list, an array of attributes of the service *x* will receive the attribute *tail* of the object dependency. With that, the *x* service will get the service it depends upon to run. The reverse steps of this process are performed for obtaining the identifier, e.g., a parent service which depends on a child service, which stores the GUID (id) of the parent service in a variable named *previous* (used later to set up the software stack).

In the second step we establish the relationship between the services and virtual machines. To do so, we implemented a method that scans a list of virtual machines (nodes) for each service (device). For each service, it is verified if the service is part of the VM node envelope (if the X and Y coordinates of the service are contained in the envelope of the node).

After the second step, each virtual machine will have a list of services with their dependencies, but unordered. Thus, in the third step we create each software stack for each virtual machine. For that, we developed an algorithm based on a topological sorting algorithm that sorts a graph [7]. Our algorithm creates a graph whose nodes of this data structure are the services objects, in which there is a variable named *previous* for each service that contains the value of the identifier (ID) of the parent service. With these variables that indicate the respective *parent* and *child* services, it is possible to assemble the software stack for each VM (when the graph is mounted, the software stack is ready) by doing a topological sort.

5.4 Allocator Module

The Allocator module performs the allocation of cloud provider(s) into virtual machines and the establishment of machine instances (machine image) in the cloud with the virtual machines that should be deployed. This allocation will make it possible to perform deployment later.

5.5 Deployment Module

In this module, two output files are automatically generated for each virtual machine modeled by the developer in the General and Specific Model. The first file corresponds to a deployment script containing the executable code for deployment. The second one has the application rules and services. The code generated for the deployment was written by the Deployment Module in Ruby language and follow the standards of the Chef platform [5].

We implemented the generation of the deployment executable file in five parts: (i) the creation of the deployment file, (ii) the file header generation, containing the path of the services' settings ("cookbooks"), (iii) the creation of names and definitions of services and the application, (iv) a call to the rules of services and application (which are defined in another file that will be generated by the module) and (v) the executable code deployment, which is comprised of the cloud specification, image VM, operating system, software stack and application.

The generation of the application rules was also implemented in five parts: (i) the creation of the related to the rules of the services, (ii) the file header, consisting of the application name and an indicator that it represents the rules, (iii) the file body with the application name, the repository of the application and services, access key to the cloud provider and the application database 'url', (iv) optionally, the services' rules containing, for example, the version, port, host, user and password of each service and (v) the classification of each service and the applications that will run on the virtual machine, in the order they are in the software stack.

6 Experiment

We conducted an experiment to evaluate our solution with respect to usability, learnability and the efforts to deploy software in the cloud, according to the point view of developers of software services from industry. To do so, we asked developers from an IT company to deploy an e-commerce software. Thus, these developers created models (General and Specific ones) containing all the necessary requirements for the application deployment.

6.1 Metrics

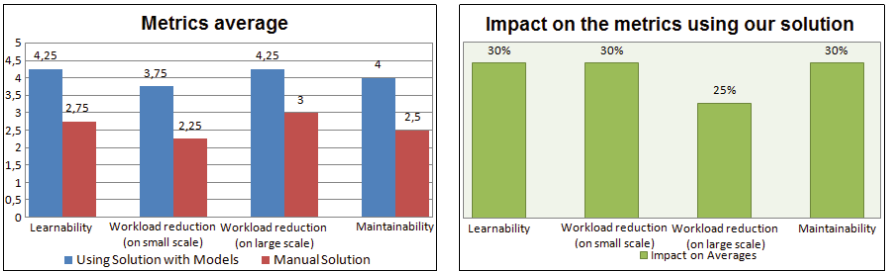
In this experiment, we defined objective assessment metrics, since evaluating software in this scope is subjective. Among the software evaluation metrics found in Nielsen [14] and Santos [17], we selected for this study: (i) Learnability (ease of learning), (ii) Workload and (iii) Maintainability.

6.2 Preparation and Experiment Execution

We asked the participants to deploy a software service in the Amazon Web Services (AWS) cloud provider using the software solution developed in this research. The participants were trained on how to use the model-based solution on performing the assigned task. Moreover, they were told that they would use StarUML as an auxiliary tool to design the General and Specific models. After completing the deployment assigned task, participants filled a survey. This survey aimed to collect data related to the metrics previously mentioned with answers in a Likert scale ranging from 1 (totally agree) to 5 (totally disagree).

6.3 Results and Analysis

After performing the experiment, we collected and analyzed the data. For each metric, we obtained the values shown in Fig. 3a. As shown in this figure, we observed that all the average metric values of learnability, workload reduction on large and small deployment scales and maintainability capacity were higher for those that used the solution. It means that there was a positive impact with the use of the solution proposed by this research. Fig. 3b shows a graph that contains the percentage difference of the weights assigned by the average of each participant, regarding the use and non-use of the model-based solution for software deployment in the cloud. We observed that there was a positive impact in the participants of 30% with respect to each metric of maintainability, learnability and reduced workload in small scale of deployment. Regarding the workload reduction for large-scale implementation, the perceived positive effect was 25%.



(a) (b)

Fig. 3 (a) Metrics average with/without using our solution. (b) Impact on metrics related to the usage of the solution

7 Conclusion

This research presented a model-based approach for automating software deployment in the cloud by using UML models. Our main goal was to reduce the workload for developers and researchers who need to use cloud resources. To evaluate our solution, we conducted an experiment involving a real usage in an IT company. As a result of the experiment, we noticed that there was a reduction in the average workload of up to 25% for developers, apart from the positive impact of 30% increase in learnability metrics and maintainability of the software deployment solution. As a future work, we aim to develop a custom graphical user interface (GUI) for the design of specific UML models to the cloud in order to make modeling even easier for the user and to overcome the results obtained in this research regarding learnability. We also expect to investigate a way to allow the solution to automatically deploy in the same virtual machine in multiple cloud providers using the same model, including hybrid clouds.

References

1. Ardagna, D., Di Nitto, E., Casale, G., Petcu, D., Mohagheghi, P., Mosser, S., Matthews, P., Gericke, A., Ballagny, C., D'Andria, F., et al.: Modaclouds: a model-driven approach for the design and execution of applications on multiple clouds. In: Proceedings of the 4th International Workshop on Modeling in Software Engineering, pp. 50–56. IEEE Press (2012)
2. Armstrong, D., Djemame, K., Nair, S., Tordsson, J., Ziegler, W.: Towards a contextualization solution for cloud platform services. In: 2011 IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom), pp. 328–331. IEEE (2011)
3. Van der Burg, S., De Jonge, M., Dolstra, E., Visser, E.: Software deployment in a dynamic cloud: from device to service orientation in a hospital environment. In: Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing, pp. 61–66. IEEE Computer Society (2009)
4. Cała, J., Watson, P.: Automatic software deployment in the azure cloud. In: Distributed Applications and Interoperable Systems, pp. 155–168. Springer (2010)
5. Chef Software, I.: Chef - code can. https://www.chef.io/ (Visited on June 01, 2016)
6. Chieu, T., Karve, A., Mohindra, A., Segal, A.: Simplifying solution deployment on a cloud through composite appliances. In: 2010 IEEE International Symposium on Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), pp. 1–5. IEEE (2010)
7. Cormen, T.H.: Introduction to algorithms. MIT press (2009)
8. Dudin, E., Smetanin, Y.G.: A review of cloud computing. Scientific and Technical Information Processing **38**(4), 280–284 (2011)
9. Fazziki, A.E., Lakhrissi, H., Yetognon, K., Sadgal, M.: A service oriented information system: a model driven approach. In: 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS), pp. 466–473. IEEE (2012)
10. Juve, G., Deelman, E.: Automating application deployment in infrastructure clouds. In: 2011 IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom), pp. 658–665. IEEE (2011)
11. Konstantinou, A.V., Eilam, T., Kalantar, M., Totok, A.A., Arnold, W., Snible, E.: An architecture for virtual solution composition and deployment in infrastructure clouds. In: Proceedings of the 3rd International Workshop on Virtualization Technologies in Distributed Computing, pp. 9–18. ACM (2009)

12. Li, W., Svard, P., Tordsson, J., Elmroth, E.: A general approach to service deployment in cloud environments. In: 2012 Second International Conference on Cloud and Green Computing (CGC), pp. 17–24. IEEE (2012)
13. Muthunagai, S., Karthic, C., Sujatha, S.: Efficient access of cloud resources through virtualization techniques. In: 2012 International Conference on Recent Trends In Information Technology (ICRTIT), pp. 174–178. IEEE (2012)
14. Nielsen, J.: Usability engineering. Elsevier (1994)
15. OMG: Uml 2.4.1. http://www.omg.org/spec/UML/2.4.1/ (Visited on June 01, 2016)
16. Salapura, V.: Cloud computing: virtualization and resiliency for data center computing. In: 2012 IEEE 30th International Conference on Computer Design (ICCD), pp. 1–2. IEEE (2012)
17. Santos, R.C.: Revisão das métricas para avaliação de usabilidade de sistemas (review of the metrics for evaluating the usability of systems). In: Congresso Internacional GBATA (2008)
18. Savu, L.: Cloud computing: deployment models, delivery models, risks and research challenges. In: 2011 International Conference on Computer and Management (CAMAN) (2011)
19. Talwar, V., Milojicic, D., Wu, Q., Pu, C., Yan, W., Jung, G.P.: Approaches for service deployment. IEEE Internet Computing **9**(2), 70–80 (2005)
20. Zhang, Y., Li, Y., Zheng, W.: Automatic software deployment using user-level virtualization for cloud-computing. Future Generation Computer Systems **29**(1), 323–329 (2013)

Software Process Improvement in Small and Medium Enterprises: A Systematic Literature Review

Gledston Carneiro da Silva and Glauco de Figueiredo Carneiro

Abstract The knowledge of characteristics and profile of a company is the key to plan its software process improvement. It helps focusing efforts to promote alignment with organizational culture and to support the consolidation of best practices already implemented. This paper presents a systematic literature review to identify evidences in the literature related to the challenges and opportunities of the adoption of software process improvement in small and medium enterprises. The results from the study indicate that there are relevant issues that can be considered in the effective adoption of software engineering best practices in small and medium enterprises.

Keywords Software quality · Software process improvement · Small and medium enterprises · Systematic literature review

1 Introduction

Small and medium enterprises (SMEs) are fundamental for the economy of the majority of countries. According to [1], in countries such as United States, Brazil, Canada, China and India, companies with this profile have significant representativeness in the economy. For this reason, it should take into account what are the difficulties faced by these companies towards the Software Process Improvement (SPI). This paper presents the results of a systematic literature review based on papers selected between January 2004 and June 2015 addressing challenges and difficulties faced by SMEs endeavoring software process improvement adoption. The results indicated that the alignment to the software engineering best practices is a key to success for those companies, although being a hard work with pitfalls and representative direct and indirect costs.

G.C. da Silva · G. de Figueiredo Carneiro(✉)
Computer Science Graduate Program Salvador University (UNIFACS),
Alameda das Espatódias 915, Salvador, Bahia 41820-460, Brazil
e-mail: gledston.silva@pro.unifacs.br, glauco.carneiro@unifacs.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_53

The next sections are organized as follows. Section 2 presents related works. In Section 3 we describe our research method for the systematic literature review. Section 4 presents the results and Section 5 discusses them. Section 6 presents final remarks and future work.

2 Background

Pino and colleagues [2] conducted a systematic literature review to identify reports and studies focusing on efforts of SMEs in the software process improvement journey. The main goal of the paper was to analyze approaches of SPI related to these organizations. According to the authors, the following items can influence on the success of SPI adoption in SMEs: (a) hiring expert advice on software process improvement; (b) financial support to fund the software process improvement; (c) cooperation between organizations interested in software process improvement so that they can share resources employed; (d) strategies to gain support from the organization teams; (e) perform a gap analysis; (f) establish and institutionalize a communication plan considering all stakeholders; (g) senior management sponsorship and strong commitment of all stakeholders.

Lavallée e Robillard [3] conducted a systematic literature review to identify papers that focused on the impact of SPI on developers. Among the positive impacts the authors identified the reduction in the number of crises, and an increase in team communications and morale, as well as better requirements and documentation. On the other hand, as negative impacts they mentioned the increased overhead on developers through the need to collect data and compile documentation, an undue focus on technical approaches, and the fact that SPI is oriented toward management and process quality, and not towards developers and product quality.

3 Research Methodology

In contrast to non-structured review processes, Systematic Literature Reviews (SLRs) [4][5] reduce bias and follow a precise and rigorous sequence of methodological steps to research literature. SLRs rely on well-defined and evaluated review protocols to extract, analyze, and document results [4][5] as the steps conveyed in Figure 1. This section describes the methodology applied for the phases of planning, conducting and reporting the review.

3.1 Planning the Review

Identify the Needs for a Systematic Literature Review. We searched the literature for evidences related to challenges and opportunities faced by small and medium enterprises while implementing software process improvement.

Specifying the Research Question. We aimed to answers the following research question (RQ) with this review: *What are the challenges and difficulties faced by SMEs in the adoption of software process improvement?* The previous knowledge of these challenges and difficulties can support SMEs to plan and perform effective SPI alignment. **Developing a Review Protocol.** We conducted a SLR focusing on papers published in journals and conferences. We extracted 33 peer-reviewed literature papers published from January 2004 to June 2015 (inclusive).

3.2 Conducting the Review

This phase is responsible for executing the review protocol. **Identification of research.** From the research question, we extracted keywords and used them to search the primary study sources with the following search string based on the same strategy cited in [6]: (*"challenges" OR "difficulties"*) AND (*"small and medium enterprises" OR "sme"*) AND (*"SPI" OR "software process improvement"*)

Selection of Primary Studies. The following steps guided the selection of primary studies. **Step 1 - Search string results automatically obtained from the engines -** Submission of the search string to the following repositories: Digital Library ACM, IEEE Xplore, Science Direct and Google Scholar. The motivation for the selection of these libraries is their relevance as sources in software engineering [7]. We performed the search using the specific syntax of each database, considering only the title, keywords, and abstract. We configured the search in each repository to select only papers carried out within the prescribed period. We complemented the automatic search by a manual search to obtain a list of studies from journals and conferences. At the end, we discarded the duplicates. **Stage 2 - Read titles & abstracts to identify potentially relevant studies -** Identification of potentially relevant studies, based on the analysis of title and abstract, discarding studies that are clearly irrelevant to the search. If there was any doubt, about whether a study should be included or not, it was included for consideration at a later stage. **Stage 3 - Apply inclusion and exclusion criteria on reading the introduction, methods and conclusion -** Selected studies in previous stages were reviewed, by reading the introduction, methodology section and conclusion. Afterwards, we applied the inclusion and exclusion criteria as follows. *Inclusion criteria*: (*IC1*) The publications should be "journal" or "conference" and written in English; (*IC2*) Works involving an empirical study or have "lessons learned" (experience report); (*IC3*) If several journal articles reporting the same study the latest article will be included; (*IC4*) The articles that address at least one of the research questions; (*IC5*) The download of the articles should be made available directly in the search source. At this stage, in case of doubt preventing a conclusion, we read the study in its entirety. *Exclusion criteria*: (*EC1*) Studies not focused on challenges and difficulties of SPI in SMEs; (*EC2*) Studies merely based on expert opinion without locating a specific experience, as well as editorials, prefaces, summaries of articles,

interviews, news, analysis/reviews, readers' letters, summaries of tutorials, workshops, panels, and poster sessions; (*EC3*) Publications that are earlier versions of last published work; (*EC4*) Publications that were published out of the period January 1st, 2005 to June 2015; (*EC5*) Studies that do not address the discussion of challenges and difficulties faced by SMEs in SPI. **Stage 4 - Obtain primary studies and make a critical assessment of them -** A list of primary studies was obtained and later subjected to critical examination using the 11 quality criteria [8] as presented in the following items: (*QC1*) Is the paper based on research (or is it merely a ''lessons learned'' report based on expert opinion)? (*QC2*) Is there a clear statement of the aims of the research? (*QC3*) Is there an adequate description of the context in which the research was carried out? (*QC4*) Was the research design appropriate to address the aims of the research? (*QC5*) Was the recruitment strategy appropriate to the aims of the research? (*QC6*) Was there a control group with which to compare treatments? (*QC7*) Was the data collected in a way that addressed the research issue? (*QC8*) Was the data analysis sufficiently rigorous? (*QC9*) Has the relationship between researcher and participants been considered to an adequate degree? (*QC10*) Is there a clear statement of findings? (*QC11*) Is the study of value for research or practice?

Data Extraction. All relevant information on each study was recorded on a spreadsheet. This information was helpful to summarize the data and map them with its source. The following data were extracted from the studies: (i) name and authors; (ii) type of article (journal, conference, workshop); (iii) aim of the study; (iv) research question. The following data were extracted on study results: (i) type of design of the study (empirical, experience report, theoretical); (ii) research method (case study, experiment, action research, survey); (iii) methodology of the analysis (qualitative, quantitative); (vii) benefits; (viii) limitations and challenges; (ix) validity; (x) relevance. Once the list of primary studies is decided, we extracted the data from the tools cited by the papers. The phase of data extraction aims to summarize the data from the selected studies for further analysis.

Data Synthesis. This synthesis aimed at grouping findings from the studies in order to: identify the main concepts (organized in spreadsheet form), conduct a comparative analysis on the characteristics of the study, as well as issues related to the research question from each study. We synthesized other information when necessary. We used the meta-ethnography method [9] as a reference for the process of data synthesis.

Conducting the Review. We started the review with an automatic search followed by a manual search to identify potentially relevant studies and afterwards apply the inclusion/exclusion criteria. The first tests using automatic search began in January 2015. We had to adapt the search string in some engines without losing its primary meaning and scope. The manual search consisted in studies published in conference proceedings and journals and suggested by the authors while searching the theme in different repositories. We analyzed these suggested studies considering both their

titles and abstracts. Figure 1 conveys these results of 36 studies. We tabulated everything on a spreadsheet to facilitate the subsequent phase of identifying potentially relevant studies. Figure 1 presents the results of 644 papers in the first phase obtained from each electronic database used in the search.

Potentially Relevant Studies. Results obtained from the automatic search and manual search were included on a single spreadsheet: an overall of 56 papers, namely 54 from the automated search plus 02 from the separate manual search. We sorted the studies were sorted by title in order to eliminate redundancies. We considered redundant studies for which the title, author(s), year and abstract were identical. After removing redundant items, remained 46 papers.

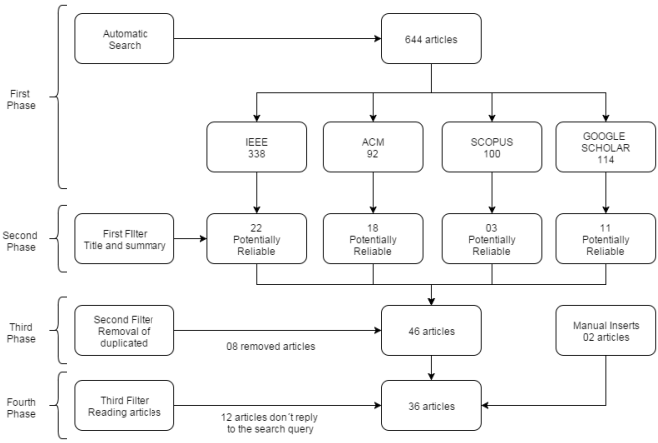


Fig. 1 Systematic Literature Review Phases in Numbers

Figure 1 conveys the selection phases with their corresponding number of papers. The first author performed the data extraction and discussed the result with the second author. He registered all the extracted data in a spreadsheet. We selected 644 papers from the search string. We grouped the papers in the first phase according to their respective source (repository or manual inclusion). At the end of the first phase, we obtained 54 papers according to their title and abstract information. The distribution of the papers was ACM (18 papers), IEEE (22 papers), Scopus (03 papers) and GOOGLE SCHOLAR (11 papers). The authors identified eight duplicated from the 46 papers in the third phase. The authors also included two papers from manual search. These new papers met the quality criteria presented before. From this point, we had 48 papers potentially relevant for the selected research questions. In the fourth phase, we applied the third filter that consisted in the complete reading of each of the 46 papers from the previous phase. In this phase, we discarded 12 papers. At the end, we considered 36 papers as relevant to answer the research question (RQ) of this systematic literature review. Table 1 presents an overview of the selection process per public data source.

Table 1 Selection Process Overview per Public Data Source.

Public Data Source	Search Result	Relevant Studies	Search Effectiveness
IEEE	338	19	5,6%
ACM	92	02	2,1%
SCOPUS	100	03	3%
GOOGLE	114	11	9,6%

4 Results

The selection process illustrated in Figure 1 ended up with 36 papers[1] from which we can obtain evidences to answer the research question. Figure 2 depicts the temporal distribution of the selected studies. As can be observed, the year of publication of more than 97% of the papers is after 2006, including this year. This is an evidence that the research on software process improvement in small and medium enterprises has been gaining increasing interest from the international software engineering community.

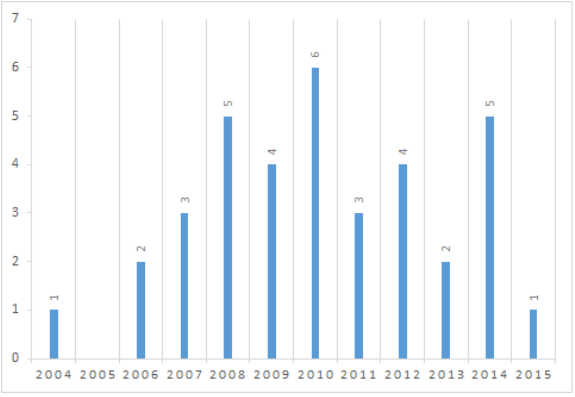


Fig. 2 Systematic Literature Review Phases in Numbers

The list of the top ten cited papers according to Google Scholar[2] shows relevant publications regarding challenges and opportunities in the adoption of software process improvement practices in small and medium enterprises. The list contains selected papers from our systematic literature review sorted in descending order according to their respective citation number in Google Scholar. The most cited paper is a systematic literature review of empirical studies of knowledge management in software engineering. Among the selected primary studies, five publications discuss the use of knowledge management approaches to support software process improvement in SMEs. Despite the selected papers of this SLR do

[1] The list is available in Appendix A.
[2] The list is available at http://www.sourceminer.org/slrsme.html.

not focus only on these types of companies; these studies are evidences of its importance in the overall context of SPI.

Characteristics, Challenges and Difficulties faced by SMEs towards SPI Adoption Relevant Studies. The selected studies show evidences of peculiarities regarding SMEs, indicating the necessity of specific approaches and solutions for an effective adoption of software process improvement. This is especially evident when comparing this scenario with large organizations. In this subsection, we present the top ten cited issues related to characteristics, challenges and difficulties faced by SMEs towards SPI adoption. In the following paragraph we present the factors along with the corresponding percentage in the selected papers: (1) Financial restrictions and human resource constraints (67%); (2) Projects with short time frame (39%); (3) Small teams with work overload (25%); (4) Low focus on process (22%); (5) Few clients and high dependency on them (19%); (6) Agility to deal with requirements volatility (17%); (7) Absence of training focusing on process (17%); (8) Protection of intellectual property (17%); (9) Difficulty to include best practices (14%); (10) High cost of SPI qualified professionals (14%).

Factors related to *financial restrictions and human resource constraints (Challenge 1)* were cited in 24 papers of the 36 of our SLR and presented at the Appendix at the end of this paper (M01-M07, M10, M12-M17, M19, M22-M24, M26, M28, M29, M31, M32 and M36), representing 67% of the selected papers. It is worth to mention that this does not only refers to direct costs (external consulting and training, for example), but also to indirect costs such as effort required by the team to implement SPI, time required by the teams to understand the SPI rationale and the promote the several required adjustments during the SPI adoption. Due to the financial restrictions, many companies have applied for funding for such end. Typical SMEs project characteristics was the second most cited issue in 14 papers (M01, M03, M04, M07, M11, M12, M14, M15, M17, M19, M20, M27, M28 e M33), i.e. 39% of the selected primary studies pointed out this fact. Due to the ever-changing business and customer demands, *projects with short time frame (Challenge 2)* are also a typical characteristic of SMEs, especially for the ones that focus on the evolution of software products of a given domain.

The third issue highlighted by the selected studies is small teams with work overload - (Challenge 3). This issue was cited by nine papers (M01-M03, M08, M12, M14, M15, M20 and M31), corresponding to 25% of the Appendix list. Small teams in SMEs should be taken into account when planning the software process improvement. In this case, participants of the teams can take different roles depending on the demand. There are roles such as quality auditor that cannot be assumed by the team members due to conflict of interest. In this case, an external professional can provide an on-demand service for this purpose.

Low focus on process - (Challenge 4) is an evidence provided by eight papers (M01, M03, M08, M12, M14, M15, M19 and M36), corresponding to 22% of the Appendix list. It reveals how SMEs deal with daily development practices. This means that organizations can be subjected to conduct software projects in a non-coordinated way and the teams involved in the software projects are not always

concerned with software process improvement. *Limited number of customers - (Challenge 5)* is another issue pointed out by seven (M01, M03, M06, M07, M12, M15 and M22) (19%) of the selected studies. These type of companies need to maintain the relationship with their clients, otherwise the company can suffer the consequences of the competitors. *Agility to deal with requirements volatility - (Challenge 6)* was an issue identified in six papers (M01, M03, M07, M15, M18 and M34) corresponding to 17% of the selected studies. Even after validating the requirements with the client, change in requirements are inevitable. Responding to these demands is crucial to keep clients satisfied. *Absence of training focusing on process - (Challenge 7)* was also identified in six papers (M03, M08, M12, M14, M19 and M35) corresponding to 17% of the selected studies. *Protection of intellectual property - (Challenge 8)* also corresponds to 17% of the selected studies (M06, M24, M25, M30, M33 and M35) and has the potentiality to encourage companies and their teams to innovate in new products, new techniques or approaches that can lead to market differential. The next issues are *difficulty to include best practices- (Challenge 9)* was mentioned in five studies (M01, M12, M14, M19 and M36) and *high cost of SPI qualified professionals- (Challenge 10)* (M02, M05, M11, M13 and M34).

Despite the difficulties and the effort needed to overcome them, it was founded that most of these companies recognize the importance of software process improvement and its adoption, especially when guided by a well-known model reference such as CMMI-DEV. However, they also recognized the need of external support for this purpose.

Limitations of This Review. The possibility of bias is the main limitation of this review. To deal with this issue, the second author supervised all the work performed in all phases of this SLR. The two authors assessed each criterion in quality assessment. When they disagreed about the assessment of a particular study they discussed to establish a consensus. Another strategy to minimize the possibility of bias was the definition of a protocol and the constant review of the spreadsheet containing the data at each stage of the review (study selection, quality assessment, data extraction and synthesis). It is also possible that relevant studies were not included. However, to reduce this threat, we performed manual and automatic searches in the selected public repositories.

5 Conclusion and Future Work

This systematic literature review initially identified 644 studies, of which 54 were considered potentially relevant and 36 studies were selected to answer the research question of this paper. The selected studies were evaluated according to quality criteria. The next step was to extract data and present the characterization of results and their respective evidences. The number of studies included and the fact that the number of published studies is growing indicate that discussion and research on this theme is relevant and current. The results present evidences of characteristics,

challenges and difficulties faced by SMEs towards the software process improvement. Despite the reported difficulties faced by SMEs, there are evidences of initiatives taken by them to overcome these difficulties and to enable the success of SPI initiatives. We plan to conduct future empirical research together with small and medium enterprises that manifest interest in adopting software process improvement. The goal is to identify and map the main characteristics and challenges faced by these companies and to which extent they coincide with the evidences collected from the selected studies from this systematic literature review.

References

1. Dyba, T.: An empirical investigation of the key factors for success in software process improvement. IEEE Transactions on Software Engineering **31**(5), 410–424 (2005)
2. Pino, F.J., Garca, F., Piattini, M.: Software process improvement in small and medium software enterprises: a systematic review. Software Quality Journal **16**(2), 237261 (2008)
3. Lavalle, M., Robillard, P.N.: The impacts of software process improvement on developers: A systematic review. In: Proceedings of the 34th International Conference on Software Engineering, pp. 113−122. IEEE Press (2012)
4. Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M., Khalil, M.: Lessons from applying the systematic literature review process within the software engineering domain. Journal of Systems and Software **80**(4), 571–583 (2007)
5. Keele, S.: Guidelines for performing systematic literature reviews in software engineering. In: Technical report, Ver. 2.3 EBSE Technical Report. EBSE (2007)
6. Chen, L., Babar, M.A.: A systematic review of evaluation of variability management approaches in software product lines. Information and Software Technology **53**(4), 344–362 (2011)
7. Zhang, H., Babar, M.A., Tell, P.: Identifying relevant studies in software engineering. Information and Software Technology **53**(6), 625–637 (2011)
8. Dyba, T., Dingsoyr, T.: Empirical studies of agile software development: A systematic review. Information and software technology (2008)
9. Noblit, G.W., Hare, R.D.: Meta-ethnography: Synthesizing qualitative studies (2011)

Appendix

ID	List of Selected Papers from the Systematic Literature Review
M01	MUNOZ, M. et al., **Expected Requirements in Support Tools for Software Process Improvement in SMEs**. Electronics, Robotics and Automotive Mechanics Conference (CERMA): 135-140 (2012).
M02	HABIB, M. et al., **Blending Six Sigma and CMMI – an approach to accelerate process improvement in SMEs**. Multitopic Conference (INMIC): 286-391 (2008).
M03	JEZREEL, M. et al., **Identifying Findings for Software Process Improvement in SMEs: An Experience**. Electronics, Robotics and Automotive Mechanics Conference (CERMA): 141-146 (2012).
M04	ALLISON, I. **Organizational Factors Shaping Software Process Improvement in Small-Medium Sized Software Teams: A Multi-Case Analysis**. Quality of Information and Communications Technology (QUATIC): 418-423 (2010).
M05	SIVASHANKAR, M., KALPANA, A.M., JEYAKUMAR, A.E.: **A framework approach using CMMI for SPI to Indian SME'S**. Innovative Computing Technologies (ICICT): 1-5 (2010).

M06	QYSER, A.A.M., RAMACHADRAM, S., FARHAN, A.: **A Staged Capability Improvement Model for Software Small and Medium Enterprises (SME)**. Asia-Pacific Services Computing Conference (APSCC): 724-729 (2008).
M07	HAMED, A.M.M., ABUSHAMA, H.: **Popular agile approaches in software development: Review and analysis**. Computing, Electrical and Electronics Engineering (ICCEEE): 160-166 (2013).
M08	SCHOEFFEL, P., BENITTI, F.B.V.: **Influential factors in software process improvement: a survey comparing "Micro and Small Enterprises" (MSE) and "Medium and Large Enterprises" (MLE)**. IEEE Latin America Transactions. 10(2): 1634-1643 (2012).
M09	MONTONI, M. et al., **MPS Model and TABA Workstation: Implementing Software Process Improvement Initiatives in Small Settings**. Software Quality (ICSE): 4 (2007).
M10	ADAM, S., DOERR, J., EISENBARTH, M.: **Lessons Learned from Best Practice-Oriented Process Improvement in Requirements Engineering: A Glance into Current Industrial RE Application**. Requirements Engineering Education and Training (REET): 1-5 (2009).
M11	SANTOS, G. et al., **Implementing Software Process Improvement Initiatives in Small and Medium-Size Enterprises in Brazil**. Quality of Information and Communications Technology (QUATIC): 187-198 (2007).
M12	ALEXANDRE, S., RENAULT, A., HABRA, N.: **OWPL: A Gradual Approach for Software Process Improvement In SMEs**. Software Engineering and Advanced Applications (SEAA): 328-335 (2006).
M13	IBRAHIM, S., ALI, R.Z.R.M.: **Study on acceptance of customised Software Process Improvement (SPI) model for Malaysia's SME**. 5th Malaysian Conference in Software Engineering: 25-30 (2011).
M14	BIN BASRI, S., O'CONNOR, R.V.: **Organizational commitment towards software process improvement an irish software vses case study**. Information Technology (ITSim): 1456-1461 (2010).
M15	SULAYMAN, M., MENDES, E.: **An extended systematic review of software process improvement in small and medium Web companies**. Evaluation & Assessment in Software Engineering (EASE): 134-143 (2011).
M16	DA ROCHA, A.R.C. et al., **A Nationwide Program for Software Process Improvement in Brazil**. Quality of Information and Communications Technology (QUATIC): 167-176 (2007).
M17	PINO, F., GARCIA, F., PIATTINI, M.: **Software process improvement in small and medium software enterprises: a systematic review**. Software Quality Journal 16(2): 237-261 (2008).
M18	DAMIAN, D., LAKSHMINARAYANAN, D.: **An Industrial Case Study of Immediate Benefits of Requirements Engineering Process Improvement at the Australian Center for Unisys Software**. Empirical Software Engineering 9(2): 45-75 (2004).
M19	HABRA, N. et al., **Initiating software process improvement in very small enterprises: Experience with a light assessment tool**. Information and Software Technology 50(7-8): 763-771 (2008).
M20	HURTADO, J. et al., **MDE software process lines in small companies**. Journal of Systems and Software 86(5): 1153-1171 (2013).
M21	MONTONI, M., DA ROCHA, A.R.C: **Applying grounded theory to understand software process improvement implementation: a study of Brazilian software organizations**. Innovations in Systems and Software Engineering. 10(1): 33-40. (2014).
M22	VON WANGENHEIM, C. et al., **Experiences on establishing software processes in small companies**. Information & Software Technology 48(9): 890-900 (2006).
M23	MUÑOZ, M.; GASCA, G.; VALTIERRA, C. **Caracterizando las Necesidades de las Pymes para Implementar Mejoras de Procesos Software: Una Comparativa entre la Teoría y la Realidad**. Iberian Journal of Information Systems and Technologies E1, 1-15 (2014).
M24	CASAÑOLA, Y. et al**., La gestión de información y los factores críticos de éxito en la mejora de procesos**. Ciencias de la información, 44(3): 27-33 (2013).
M25	CASAÑOLA, Y. et al., **Indicadores para valorar una organización al iniciar la mejora de proceso de software**. Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology, 14-16 (2014).
M26	MONTONI, M. **Uma investigação sobre os fatores críticos de sucesso em iniciativas de melhoria de processos de software**. Tese de Doutorado. Universidade Federal do Rio de Janeiro. (2010).
M27	UNTERKALMSTEINER, M. et al., **Evaluation and measurement of software process improvement—a systematic literature review**. Software Engineering, IEEE Transactions on. 38(2): 398-424 (2012).
M28	MISHRA, D.; MISHRA, A. **Software process improvement in SMEs: A comparative view**. Computer Science and Information Systems 6(1): 111-140 (2009).
M29	O'CONNOR, R.; BASRI, S.; COLEMAN, G. **Exploring managerial commitment towards SPI in small and very small enterprises**. Systems, Software and Services Process Improvement. Springer Berlin Heidelberg, 268-279 (2010).
M30	BASRI, S.; O'CONNOR, R. **Towards an understanding of software development process knowledge in very small companies**. Informatics Engineering and Information Science. Springer Berlin Heidelberg, 62-71 (2011).
M31	MÜLLER, S.; MATHIASSEN, L.; BALSHØJ, H. **Software Process Improvement as organizational change: A metaphorical analysis of the literature**. Journal of Systems and Software, 83(11): 2128-2146 (2010).

M32	TRIENEKENS, J. et al., **Entropy based software processes improvement**. Software Quality Journal, 17(3): 231-243 (2009).
M33	BJØRNSON, F. O.; DINGSØYR, T. **Knowledge management in software engineering: A systematic review of studied concepts, findings and research methods used**. Information and Software Technology. 50(11): 1055-1068 (2008).
M34	KABAALE, E.; KITUYI, M.; MBARIKA, I. **Requirements Engineering Process Improvement Challenges faced by Software SMEs in Uganda**. International Journal of Computer Applications, 88(5): 2014.
M35	KHANKAEW, S.; RIDDLE, S. **A review of practice and problems in requirements engineering in small and medium software enterprises in Thailand**. IEEE Fourth International Workshop on. 1-8 (2014).
M36	ALMOMANI, M. A. T. et al, **Software Development Practices and Problems in Malaysian Small and Medium Software Enterprises: A Pilot Study**. IT Convergence and Security (ICITCS), 2015 5th International Conference on. 1-5 (2015).

Investigating Reputation in Collaborative Software Maintenance: A Study Based on Systematic Mapping

Cláudio Augusto S. Lélis, Marco Antônio P. Araújo, José Maria N. David and Glauco de F. Carneiro

Abstract [Background] Reputation systems have attracted the attention of researchers when it comes to collaborative systems. In the context of collaborative software maintenance, systems of this type are employed to facilitate the collection, aggregation and distribution of reputation information about a participant. GiveMe Infra is an infrastructure that supports collaborative software maintenance performed both by co-located and geographically distributed teams. In this last case, reputation is one of the factors that influence collaboration. Despite this recognized relevance, to the best of our knowledge, there is a shortage of tools providing reputation functionalities in the context of collaborative software maintenance. [Objective] However, GiveMe Infra needs to identify and correlate metrics, measures, criteria and factors (called parameters) that are used in defining the value of reputation of an entity. These parameters, used to determine the degree of reputation, can provide evidence of parameters to be used in the context of software maintenance and evolution. [Method] In order to achieve this goal, a systematic mapping was performed. Both the established protocol and the process

C.A.S. Lélis(✉) · M.A.P. Araújo · J.M.N. David
Post Graduation Program in Computer Science -
Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil
e-mail: lelis@ice.ufjf.br, {marco.araujo,jose.david}@ufjf.edu.br

M.A.P. Araújo
Federal Institute of Education, Science and Technology of Southeast of
Minas Gerais –Campus Juiz de Fora (IF Sudeste MG), Juiz de Fora, Brazil

G. de F. Carneiro
Post Graduation Program in Systems and Computing
University of Salvador (UNIFACS), Salvador, Brazil
e-mail: glauco.carneiro@unifacs.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_54

adopted during the mapping are shown in this article. The parameters identified, as well as how to apply them in the context of software maintenance are demonstrated through an analysis scenario. [Results] Our goal has been achieved since the systematic mapping allowed the identification of the parameters used in defining reputation and even parameters that would not allow a correlation with the context of collaborative software maintenance. The contribution of this research is threefold, in its investigation carried out, the list of identified parameters and also the application in the collaborative software maintenance context.

Keywords Reputation · Collaborative software maintenance · Software evolution · Metrics

1 Introduction

Reputation systems have attracted the attention of researchers in the field of Collaborative Systems. The need for information about reputation is becoming greater as increasing numbers of people and services interact across different systems. During this interaction it can often be observed that the parties involved do not know one another. Therefore, they make use of reputation systems, so as to obtain an opinion on the entity with which they wish to interact[1]. Reputation, according to [2], can be understood as the perception that an agent (entity) creates through previous actions, based on their intentions and rules.

Facilitating the collection, aggregation and distribution of reputation information about an entity are the main functions for which reputation systems are employed [1]. The calculation of reputation is an important step in the execution of reputation systems. So, it is necessary to analyze and establish the parameters that will be considered in the aggregation of information, since they influence the calculation of the value of reputation. In [3], different systems were analyzed and the researchers realized that, in them, the entities evaluate each other. For example, upon completion of a transaction, you can use the aggregate ratings on a given entity, seller or buyer, for example, to get a reputation score. Such information may assist other parties in deciding whether or not to conduct transactions in the future and with which entity. Similarly, this mutual assessment may occur in collaborative systems. In the context of software maintenance and evolution, reputation systems could be applied to support managers in the formation of groups or allocation of developers to maintenance tasks.

This research arose in the context of GiveMe Infra [4], an infrastructure to support the collaborative software maintenance and evolution activities performed by co-located or geographically dispersed teams.

Considering geographically distributed teams, reputation is one of the key issues that affect collaboration between members. Reputation is one of the ways of establishing trust in the team, and it can be said that trust is demonstrated through the use of reputation [1]. In essence, reputation information may serve as a basis for users to decide who they will trust, and to what degree. Thus, reputation,

through the trust established, becomes an important element for coordination and collaboration activities, and it can lead the team to fail completely [5] or to become more cohesive [6], favoring collaboration. For some time, the concepts of reputation and trust were applied without distinction. In [2], these concepts are different, since reputation is seen as a bridge to establish trust. Therefore, the influence reputation has in the perception of trust among members of distributed teams that adopt the global software development is known. However, the factors influencing reputation in those teams are still unknown.

Previous studies conducted surveys to analyze the existing reputation systems and proposed taxonomies in order to classify these systems under different dimensions [1][3][7]. As a result, there is a list of reputation systems supporting online systems, including peer-to-peer networks. However, none of the tools focused on collaborative software maintenance activities. This is indeed an opportunity for the use of GiveMe Infra, in the sense that it can support the formation and maintenance of team members' reputation information. This scenario also reveals the opportunity to conduct a systematic mapping to collect evidences from the literature related to these issues, especially the identification and relationship among metrics, measures, criteria and factors to determine the value of the reputation. According to [3], there is a lack of coherence in research on reputation and its different systems, indicated by the fact that often, authors propose new systems from scratch, without trying to extend or improve previous proposals.

Due to the lack of proposals in the context of software maintenance, the analysis was performed assuming the use of reputation data in several contexts. In addition, the study that considered metrics, measurements, factors and criteria used to determine the degree of reputation can provide indications of use in the context of maintenance and evolution.

This study is divided into four sections in addition to this introduction. In Section 2, the planning of the systematic mapping is described. Section 3 presents the performance of mapping and then, in Section 4, the parameters are identified from the established issue. Section 5 shows an analysis scenario for these parameters. Section 6 deals with the threats to the validity of this research. Finally, Section 7 presents the final considerations.

2 Planning the Systematic Mapping

The activity of planning the systematic mapping includes the identification of objectives and the definition of a research protocol. It is of foremost importance to highlight that the protocol was prepared using the guidelines defined by [8].

Studies on systematic mapping are designed to provide an overview of a field of research. The goal of this study was defined according to the Goal/Question/Metric approach (GQM) [10]. The objectives, according to this approach were formulated as: "Analyzing the metrics, measures, criteria and factors used in the aggregation and calculation of the degree of reputation, in order to characterize them with regard

to effectiveness from the point of view of the managers or team leaders in relation to their collaborators in the context of software maintenance and evolution.

Formulating research questions is a key activity during the definition of the protocol [8]. This systematic study is designed to answer the main research question.

What are the metrics, measurements, criteria, factors that can be used in the aggregation and calculation of the reputation of those involved in the maintenance and evolution of software?

A research question leads to the definition of a search to be performed in publication databases. To perform this query, keywords are defined. The search terms are built in three stages: the structuring of research questions in order to identify keywords, the identification of synonyms for each of the keywords, and the construction of the search string based on the combination of the terms and their synonyms, using the AND and OR operators. The keywords used in the research were defined based on the PICOC [9] strategy. In this project, a comparison is not relevant since this mapping aims to conceive an overview of the subject through an exploratory study. The result of the process is shown in Table 1.

Table 1 PICOC

PICOC	**KEYWORDS**
Population	entity, user, developer, leader, manager, stakeholder
Intervention	reputation, trust, calculation aggregation
Control	Not defined
Outcome	criteria, criterion, metrics, measure, factor
Context	software maintenance, software evolution

The early studies selected must contribute directly to the answer to the research questions regarding mapping, so that these articles are not ignored by the search strategy [8]. Thus, articles, such as, [11] and [12] were defined as control studies because they belong to the investigated domain and were the ones most aligned with the research, with regard to clarity in presenting the parameter used, opinion [11] and relationship [12], as well as by specifying the context of application.

In order to select the databases to search for publications, certain criteria were taken into consideration. The bases adopted in mapping fulfill the requirements set out in [13]:

- they are able to interpret logical and similar expressions;
- they allow search in the whole text or in specific fields (title, abstract);
- they cover the area of research of interest to this mapping: computer science;
- they have no word limitation and combinations in the string to be searched.

In accordance with these requirements, the databases known as Scopus, ScienceDirect, IEEExplore, El Compendex and Web of Science, available at the CAPES periodical portal, accessible at UFJF and recommended by [14] were used.

The search string (Table 2) was defined as a combination of keywords according to the PICOC, besides being aligned to what is shown in the control articles. This was verified by a specialist in the field, and also used to generate specific strings, considering the syntactic peculiarities of each database chosen.

Table 2 Search String

```
("entity" OR "user" OR "developer" OR "leader" OR "manager"
OR "stakeholder") AND ( "criteria" OR "criterion" OR "metric"
OR "measure" OR "factor") AND ("reputation" OR "trust") AND
("calculation" OR "aggregation")
```

For the selection of results, criteria for the inclusion and exclusion of articles were defined. All documents obtained by the implementation of the search string in each selected source were analyzed to verify their relevance to this research. Only articles in English were included when searching with the string in the databases, and the search domain was "Computer Science". The process used to include or exclude articles was based on [8]. The criteria adopted are shown in Table 3.

Table 3 Inclusion and Exclusion Criteria

Type	Description
Inclusion	Articles containing experimental studies, case studies or reviews of any metrics, criteria, measurements or factors to determine reputation.
	Studies containing a taxonomy or survey of metrics, criteria, measurements or factors to determine reputation.
Exclusion	Book chapters, calls to congresses and secondary education material.
	Studies that cannot be completely accessed.
	Studies that do not address the topic in the field of Computer Science.
	Studies related to reputation, but which do not have any metrics, criteria, measurement or factor.
	Studies related to a metric, criterion, measurement or factor, but not to determine reputation.

3 Conducting the Systematic Mapping

Initially, the research conducted generated a list of all of the documents found. This process was completed with the help of a tool called Parsifal[1]. This tool imports BibTex standards and allows you to select the item and mark it as accepted, rejected or duplicated for each research database, and supports the entire development of the mapping protocol.

[1] http://parsif.al/

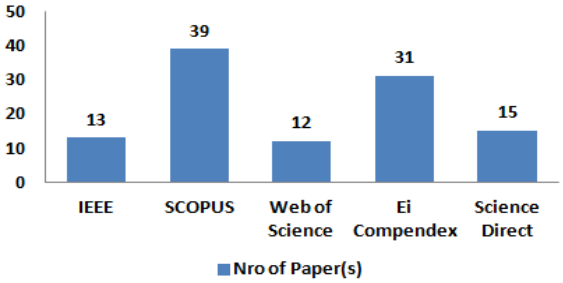


Fig. 1 Articles Returned by Database

Upon completion of the protocol defined in the previous section, it was run as described in this section. The survey covered all articles published until 2015. The search string was calibrated and applied in each digital database. The string in each database was applied to the fields 'title of the article', 'abstract' and 'keywords'. The distribution of the articles returned between the databases is shown in Fig. 1.

The screening process had three stages of analysis. Fig. 2 shows the process flow from the implementation of the string, showing the number of items excluded at each step.

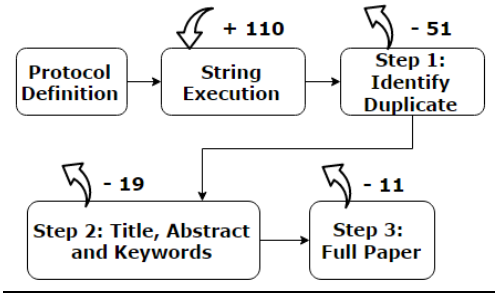


Fig. 2 Screening Process

The first step of the screening process was supported by the Parsifal tool, which identified automatically a total of 51 duplicate items, with the duplicates being excluded.

The second step was the application of the inclusion and exclusion criteria established in the protocol in the articles resulting from the previous step. At this stage, the fields analyzed were the 'title', 'abstract' (summary) and 'keywords'. A total of 19 articles were excluded.

The third step considered a full analysis of the remaining articles, and again, the inclusion and exclusion criteria were applied, seeking to identify the suitability of the article to the criteria established. In this process 11 articles were excluded. At the end, 27 were accepted. For space restrictions, these articles are not presented. The complete list of accepted papers can be accessed online[2].

[2] http://givemeinfra.com.br/files/accepted_articles.pdf

4 Presentation of Results

The parameters identified were cataloged and summarized, aiming to answer the research question defined in the protocol in search of *which metrics, measurements, criteria, and factors can be used in the aggregation and calculation of the reputation of the entities involved in software maintenance and evolution*. It was even possible to identify parameters that would not allow a correlation with the context of collaborative software maintenance since they are specific in their domains. An example of this is the parameter 'network efficiency' [15], considered in the context of *grid* computing. Similarly, the parameter 'value factor' [16], in the context of online auctions, was identified. It was associated with the category of the product being sold and is used to calculate the reputation of a seller. Due to space limitations, Table 4 shows the most commonly used and mentioned parameters in the articles accepted in the mapping, and their frequency of use. A brief discussion of the parameters is presented below.

In general, feedback is the parameter considered as the basis for calculating reputation, and it is present in [12] and [11]. It is treated as the views of other entities regarding a specific one. When considered, this parameter is used to calculate reputation as a value that represents the weighted average of the opinions collected. The feedback can be seen and analyzed from three points of view. The first one considers the point of view from an entity, in which an entity gives its opinion about another entity. However, these entities form groups with the passing of time, even unintentionally, since they have similar characteristics [17] or opinions [18]. This raises the point of view of team, through which a team offers its opinion about another one, or its members evaluate the whole [19]. These teams may a have similar origin or be in different contexts. Thus, the domain interferes with its reputation [20][21], which leads to the third point of view of the domain. It has a reputation in relation to another domain, and this is used as a parameter if entities from different domains want to have a relationship.

Table 4 Most widely used parameters

Parameter	**References**	**Frequency**
Opinion	[12] [11], [17], [18], [19],[20],[21]	25.93%
Time	[16], [22], [23], [24]	14.81%
Relationship	[11], [19], [25], [26]	14.81%
Behavior	[11], [21], [27]	11.11%
Subjectivity	[17], [18], [28]	11.11%
Reliability	[21], [22]	7.41%

An established relationship may be considered as a parameter, e.g. [11] distinguishes the user's trust from the trust in the user's relationship as two separate parameters, and both influence reputation. The goal is to determine the quality of that relationship from the point of view of the entities involved. The relationship between social groups is the measurement considered in [25] to determine the reliability of a group and to define the concept of trust between groups based on previous relationships. A relationship is important for generating a subsequent opinion. This can occur between a group and an organization, as in [19], which deals with the relationship of trust between schools and students in the e-learning environment. The sharing of classified information can also be seen as a measurement of the strength of the relationship [26]. The sharing of private information for a person is an expression of trust, so the more private information two people exchange, the greater the trust between them. This sharing is measured by the ratio between the level of privacy of the shared item and the number of people with access to it.

User behavior addressed by [11] is a factor that influences reputation in that it changes the view of other users. Despite the connection with feedback, behavior projects a sense of evolution, based on a history. This is dealt with by [27] and [21]. That is, the history is centered on a target entity and considers its previous reputation values, while the feedback considers the point of view of other entities regarding the target entity. In [21], this history is used as a sub-criterion for calculating another metric identified, the reliability of the information. In [22], in turn, it is applied to third-party recommendations, which are employed when entity "A" has never had a direct relationship with entity "B".

Time is identified as a factor influencing the calculation of reputation. It was called efficiency time in [22], where it is used as a way of indicating the depreciation of the information over time, where newer information is more valued than older information. In that case, it was defined as a function that gradually depreciates a piece of information. Time was also called 'time passed' [23] and it was defined as a configurable value which represents a flag, from which the views or feedback from the other entities regarding a given entity are no longer considered. This was established due to data storage efficiency reasons. This idea that time depreciates an opinion is also seen in [16] and [24]. In the latter, it was known as volatility. In general, it should be noted that a piece of information is only valid for a period of time.

When the members offer their opinion, the fact that each one has a distinct personality brings about a new factor named **subjectivity** by [17]. This factor adversely affects the value of reputation and should be minimized, as it generates distortion in the evaluations. One way to minimize its effect is to consider only those ratings shared by users with similar characteristics in the calculation. These characteristics are called **trust attitudes** by [17], and they are implicitly derived from their past ratings. Something similar is done by [28] when considering a measurement of similarity with certain ratings from evaluators, being described as "**the degree of coherence of criteria**" and used as a weight for the aggregation of

reputation. In [18], it is also called **similarity factor**. In order to illustrate this mechanism, suppose you want to analyze the reputation of a user (U1) by utilizing 5 previous ratings, each from a different user (U2, U3, U4, U5 and U6). Also suppose that on a rating scale of Bad, Average and Good, the ratings were, respectively: Good, Average, Good, Bad and Bad. At first glance, without analyzing the characteristics of each user, a weight equal to the ratings is distributed and one may come to the conclusion that the reputation of U1 is Average, however, since U2 and U4 have characteristics similar to those of U1 (programming language knowledge, practice time, or opinion, for example), their ratings have greater weight. Similarly, U5 and U6 have characteristics that differentiate them from U1, so their ratings will have less weight. As a result, the reputation of U1 tends to differ, and may reach Good.

As these parameters were obtained and used in other domains, the next section presents a usage scenario aiming to show the way in which they could be used in the context of collaborative software maintenance.

5 Reputation Parameters Analysis

In this section we illustrate how the identified parameters can be used to support reputation analyses. The scenario was set in the context of the GiveMe Infra infrastructure and its visualization components. The analysis of these parameters together generates a multi-criteria value of reputation, where the parameters interfere positively or negatively to calculate the final value. To represent the reputation of value, a T-Viz visualization displayed in [29] was used. T-Viz combines radar charts, which are two-dimensional previews of multivariate data with pie charts, resulting in a view of the scores in the helix style. The KGB color pattern is used on each helix representing the value of each criterion. Thus, it is possible to have a representation of a value of reputation based on multi-criteria.

Scenario. The scenario describes the use of the parameters by the team leader when assigning tasks to developers. This starts when a new case of software maintenance (C6) is registered in GiveMe Infra and the team leader, represented here by the project manager, needs to assign a developer to handle the case. The manager recovers, via GiveMe Infra, the data from all developers on a table, and another table lists maintenance cases which have already been completed, and the like. This similarity is between the change requests. The manager then selects a case that has been dealt with previously and triggers a preview offered by GiveMe Infra. The preview shows the relationship between previous cases, for example, and by hovering the mouse, the manager has access to the reputation information of the developer that dealt with the case. Fig. 3 shows this situation. From the analysis of previous cases and the current reputation of the developer, the manager can make his decision about the best option to handle the case in question.

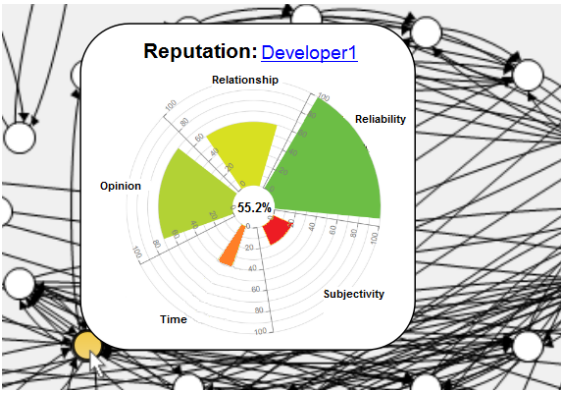


Fig. 3 Developer's Reputation Representation

Parameters in Action. In Fig. 3, the five parameters (Reliability, Subjectivity, Time, Opinion and Relationship) represented are the ones discussed in the previous section, and they have been adapted to GiveMe Infra. In order to compose the final value of reputation, the weight of each parameter is configurable by the manager to suit them to what the organization values most. This value is displayed in the center of the radar.

The parameter 'Subjectivity' follows the idea of [17], giving priority to similar ratings. In that way, suppose Developer1 (D1) can be assigned to handle case C6. Then, the analysis of his reputation is made using ratings from the 5 cases (C1, C2, C3, C4 and C5), dealt with by D1. Also suppose that on a numerical scale of 1 to 10 for the rating, these ratings were as follows: 9, 6, 10, 3 and 4. Considering the reputation as a weighted average of the values, and without analyzing the characteristics of each case, an equal weight is distributed (= 0.75) to the ratings, and one may think that the reputation of D1 is 4. However, since C1, C3, C5 and C6 have features similar to those of C6 (type of maintenance, request origin, device affected, for example), their ratings have greater weight (= 1). Similarly, C2 and C4 have characteristics that differentiate them from C6, thus their ratings have lower weight (= 0.5). As a result, the reputation of D1 tends to be different. In this case, equal to 5. This is shown in Fig.3 in the red portion of the radar.

In this scenario, the parameter 'Opinion', presented in [12], [11] and [18], underwent an adaptation, being considered the value of reputation associated with each case dealt with previously. For Developer 1 (D1), this parameter has a score of 70% and is the weighted average of the previous cases.

The factor 'Time' was introduced following the definition of [22] and [24] to depreciate the reputation of value over time. Thus, a case that has been resolved for more time has less influence than a more recent case. Considering the 5 ratings already mentioned for cases C1 through C5 and their respective values, it is assumed that the parameter 'Time' represents a maximum weight (= 1) and, according to a definition made by the manager, is depreciated at 0.1 every two weeks past the date that the case was dealt with. Considering reputation as a weighted

average of the values and that C1 was dealt with a week ago, C2 3 weeks ago, and C3, C4 and C5 5 weeks ago, the weight for each case is, respectively, 1, 0.9, 0.8, 0.8 and 0.8. Thus, the older cases have less influence on the final value of reputation, while the newer ones (C1, for example) have increased participation in the calculated value. As a result, the reputation of D1 may vary with time.

The parameter 'Relationship' was adapted and measures the quality of the relationship with team members [11] as well as the strength of this relationship [26]. Reliability, in turn, adapted from [22], represents the degree to which the information passed by D1 is reliable.

This simple scenario shows: (1) how the parameters can be used to determine the reputation of a team member in the context of software maintenance and evolution, and (2) how visualization techniques can be used to support the representation and analysis of reputation values.

6 Threats to Validity

This project has limitations regarding its validity. Even attempting to obtain a comprehensive search string that would meet the research needs, it is possible that other parameters exist and that have not been returned. To reduce this risk, control items were defined, and then returned. This provides evidence that the string obtained good coverage. Thus, it was possible to highlight parameters that can be considered in future studies.

Regarding the query conducted, the search processes in the digital databases are not exactly the same, and how they work internally is not shown. In order to mitigate this threat, the search strings were adapted to each digital database, taking care so that the logical expressions consistently become equivalent across all of them.

7 Final Considerations

The initial idea that parameters used to determine the degree of reputation can provide evidence of parameters to be used in the context of collaborative software maintenance and evolution was verified. Systematic mapping allowed the identification of the parameters used in defining reputation through the calculation and aggregation of these parameters.

The contribution of this work consists of the investigation performed, the list of parameters identified and also the way in which these parameters can be migrated to the context of collaborative software maintenance. In order to do that, a usage scenario of the parameters in the context of an infrastructure for the maintenance and evolution of software was presented, and it is called GiveMe Infra. Once reputation information has been integrated into the infrastructure, it can contribute to estimate the effort spent by the developer in the next cases to be treated, based on reputation and previous cases. It has also contributed to the formation of groups

for handling a case of maintenance in a collaborative manner, and also in the assigning of tasks to the team, as handled by the usage scenario.

As future work, we intend to propose a configurable model supported by ontology for the calculation of reputation based on different domains. This model will be implemented on a reputation service in GiveMe Infra. We also intend to conduct another study in the form of an interview and questionnaire, from the point of view of managers and developers to confirm parameters previously identified, as well as analyzing other factors and metrics that influence the calculation and aggregation of the value of reputation.

Acknowledgements To CAPES, FAPEMIG and CNPq for the financial support.

References

1. Hendrikx, F., Bubendorfer, K., Chard, R.: Reputation systems: A survey and taxonomy. J. Parallel Distrib. Comput. **75**, 184–197 (2015)
2. Mui, L., Mohtashemi, M., Halberstadt, A.: A Computational Model of Trust and Reputation. In: 35th Hawaii Int. Conf. Syst. Sci., vol. 00, no. c, pp. 1–9 (2002)
3. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decis. Support Syst. **43**(2), 618–644 (2007)
4. Tavares, J., David, J.: Uma Infraestrutura baseada em Múltiplas Visões Interativas para Apoiar Evolução de Software. iSys-Revista Bras (2015)
5. Audy, J.L.N., Prikladnicki, R.: Desenvolvimento Distribuído de Software (2007)
6. Jarvenpaa, S.: Toward contextualized theories of trust: The role of trust in global virtual teams. Inf. Syst. (2004)
7. Liu, L., Munro, M.: Systematic analysis of centralized online reputation systems. Decis. Support Syst. **52**(2), 438–449 (2012)
8. Kitcheham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Tech. report, Ver. 2.3 EBSE Tech. Report. EBSE (2007)
9. Petticrew, M., Roberts, H.: Systematic reviews in the social sciences: A practical guide. John Wiley & Sons (2008)
10. Basili, V.R., Weiss, D.M.: A Methodology for Collecting Valid Software Engineering Data. IEEE Trans. Softw. Eng. **SE-10**(6), 728–738 (1984)
11. Caverlee, J., Liu, L., Webb, S.: The SocialTrust framework for trusted social information management: Architecture and algorithms. Inf. Sci. (Ny) **180**(1), 95–112 (2010)
12. Rosaci, D., Sarné, G.M.L., Garruzzo, S.: Integrating trust measures in multiagent systems. Int. J. Intell. Syst. **27**(1), 1–15 (2012)
13. Costa, C., Murta, L.: Version control in distributed software development: A systematic mapping study. In: 2013 IEEE 8th Int. Conf. (2013)
14. Kitchenham, B.: What's up with software metrics?–A preliminary mapping study. J. Syst. Softw. (2010)
15. Selvi, K., Wahida Banu, R.S.D.: A Hybrid Model for Load Aware Trust Management in Grid. J. Comput. Sci. **7**(8), 1237–1243 (2011)

16. Chang, J.-S., Wong, H.-J.: Selecting appropriate sellers in online auctions through a multi-attribute reputation calculation method. Electron. Commer. Res. Appl. **10**(2), 144–154 (2011)
17. Zupancic, E., Juric, M.B.: TACO: a novel method for trust rating subjectivity elimination based on Trust Attitudes COmparison. Electron. Commer. Res. **15**(2), 207–241 (2015)
18. Xu, J., Jiang, D., Wang, B., Yang, D., Reiff-Marganiec, S.: Local Reputation Management in Cloud Computing. In: 2015 IEEE World Congress on Services, pp. 261–267 (2015)
19. Jayashree, R., Christy, A.: Improving the enhanced recommended system using Bayesian approximation method and normalized discounted cumulative gain. Procedia Comput. Sci. **50**, 216–222 (2015)
20. AlNemr, R., Meinel, C.: Getting More from Reputation Systems: A Context Aware Reputation Framework Based on Trust Centers and Agent Lists. In: 2008 Third Int. Multi-Conference Comput. Glob. Inf. Technol. (ICCGI 2008), pp. 137–142 (2008)
21. Khiabani, H., Sidek, Z.M., Manan, J.-L.A.: Towards a Unified Trust Model in Pervasive Systems. In: 2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops, pp. 831–835 (2010)
22. Jiang, L., Xu, J., Zhang, K., Zhang, H.: A new evidential trust model for open distributed systems. Expert Syst. Appl. **39**(3), 3772–3782 (2012)
23. Cao, Z., Li, Q., Lim, H.W., Zhang, J.: A multi-hop reputation announcement scheme for VANETs. In: Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics, pp. 238–243 (2014)
24. Leberknight, C.S., Sen, S., Chiang, M.: On the volatility of online ratings: An empirical study. In: Lect. Notes Bus. Inf. Process., LNBIP, vol. 108, pp. 77–86 (2012)
25. Memarmoshrefi, P., Alfandi, O., Kellner, A., Hogrefe, D.: Autonomous Group-Based Authentication Mechanism in Mobile Ad Hoc Networks. In: 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 1097–1102 (2012)
26. Wrobel, S., Heupel, M., Thiel, S.: Evaluation of the di.me trust metric in CRM settings. In: Second International Conference on Future Generation Communication Technologies (FGCT 2013), pp. 132–136 (2013)
27. Yin, G., Wang, Y., Dong, Y., Dong, H.: Wright–Fisher multi-strategy trust evolution model with white noise for Internetware. Expert Syst. Appl. **40**(18), 7367–7380 (2013)
28. Xiangli, Q., Xuejun, Y., Jingwei, Z.: Towards reliable trust establishment in grid: a pre-evaluating set based reputation evaluation approach. In: Sixth IEEE International Symposium on Cluster Computing and the Grid (CCGRID 2006), pages 5, p. 385 (2006)
29. Volk, F., Hauke, S., Dieth, D., Muhlhauser, M.: Communicating and visualising multicriterial trustworthiness under uncertainty. In: 2014 Twelfth Annual International Conference on Privacy, Security and Trust, pp. 391–397 (2014)

A 2-Layer Component-Based Architecture for Heterogeneous CPU-GPU Embedded Systems

Gabriel Campeanu and Mehrdad Saadatmand

Abstract Traditional embedded systems are evolving into heterogeneous systems in order to address new and more demanding software requirements. Modern embedded systems are constructed by combining different computation units, such as traditional CPUs with Graphics Processing Units (GPUs). Adding GPUs to conventional CPU-based embedded systems enhances the computation power but also increases the complexity in developing software applications. A method that can help to tackle and address the software complexity issue of heterogeneous systems is component-based development.

The allocation of the software application onto the appropriate computation node is greatly influenced by the system information load. The allocation process is increased in difficulty when we use, instead of common CPU-based systems, complex CPU-GPU systems.

This paper presents a 2-layer component-based architecture for heterogeneous embedded systems, which has the purpose to ease the software-to-hardware allocation process. The solution abstracts the CPU-GPU detailed component-based design into single software components in order to decrease the amount of information delivered to the allocator. The last part of the paper describes the activities of the allocation process while using our proposed solution, when applied on a real system demonstrator.

1 Introduction

Due to the advances in micro and nano technology fabrication, traditional embedded systems are becoming more and more complex. The traditional homogeneous

G. Campeanu(✉) · M. Saadatmand
Mälardalen Real-Time Research Center (MRTC), Mälardalen University, Västerås, Sweden
e-mail: {gabriel.campeanu,mehrdad.saadatmand}@mdh.se

M. Saadatmand
SICS Swedish ICT, Västerås, Sweden
e-mail: mehrdad@sics.se

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_55

unicore CPU-based systems have emerged as heterogeneous systems which combine various processing units, such as multi-core CPUs and GPUs. The GPU computation power brings considerable speed-ups compared to the traditional CPU, for various software applications such as n-body simulations [11] or 3D reconstruction medical systems [13]. Also, the large GPU parallel computation power made possible the appearance of new and more complex system applications such as vehicle vision systems [7] and autonomous vision-based robots [9]. The combination of CPU and GPU increases the system computation power, but comes with a cost: being a different computation unit with its own memory system, the GPU increases the complexity of the software system.

A way to tackle the newly increased software complexity is through component-based development (CBD). CBD increases developers productivity by constructing complex software applications out of existing software blocks know as components. Several advantages arrive by following this development approach; among them, we can mention an increased productivity and a faster time-to-market.

Developing the component-based design of a system which requires to use the heterogeneous CPU-GPU hardware is demanding due to the variety of component candidates. From a repository of CPU and GPU-based components, different combinations of components with their respective properties can construct different alternatives that have the same functionality. In an embedded system with multiple heterogeneous hardware nodes, the software-to-hardware allocation process needs to consider all component-based design alternatives[1] in order to choose the right one w.r.t. the system properties and constraints (e.g., performance optimization). For instance, having a system design composed of only GPU-based components may result in an infeasible allocation scheme due to constraints such as the GPU hardware resource limitation.

Determining the software-to-hardware allocation is an NP-hard problem [5]. This challenge is increased even more when instead of using common CPU-based systems, we use complex CPU-GPU based platforms. In our previous work [6], we developed a software component allocation model for heterogeneous CPU-GPU systems. We used the CBD approach to model the software system, and characterize it with extra-functional properties (e.g., CPU and GPU memory usage). Using the allocation model, we constructed a semi-automatic allocator which balances the hardware resource usage and optimizes the system performance. The work does not consider multiple alternatives in its allocation process. This negatively influences the outcome of the allocation process because it excludes feasible solutions.

Having multiple alternatives increases the information load of the allocator. For example, the allocator needs to take in consideration each component properties (e.g., memory, CPU and GPU usage) and the information regarding the communication links between the connected components from each alternative. Due to the tightly-connected nature of the GPU to the CPU, the allocator also requires to fulfill the

[1] For the rest of the paper, the “component-based design alternative" term will be simply referred as “alternative"

constraint of deploying the entire variant onto a single heterogeneous processing node in order to not negatively influence the overall system performance.

In this paper, we propose a 2-layer architecture to ease the software-to-hardware allocation. Both of the layers describe the same system that is using GPU computation power to fulfill its functionality. The first layer contains the alternatives and their detailed properties such as communication links or memory usage. We propose a second layer that encapsulates the alternatives into a software component with multiple variants. Each variant is characterized by a distinctive set of properties that reflects the attributes of all of the components contained by its corresponding alternative. For computing the allocation scheme, we use the properties of the abstracted second layer. After a suitable component variant is selected by the allocator, we return to the first layer in order to describe the full details of the selected alternative which was initially abstracted away. For future adjustments of the system e.g., to be used in other contexts, the first layer provides the needed component detailed view to the developer. Using two distinct levels of granularity, decreases the information load and constraints on the allocator, making the allocation process easier. Another advantage that follows from our approach is that it can improve the allocation scalability. More complex systems (e.g., number of components or properties) may be handled now by the allocator using the 2-layer approach, when the information load of all possible allocation scenarios is decreased.

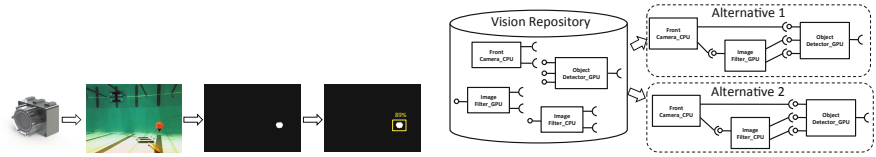
The rest of the paper is organized as follows. Section 2 describes, using a running example, the software-to-hardware allocation challenges when dealing with many alternatives. The details of our solution are presented in Section 3. A case study is implemented in Section 4 in order to describe the feasibility and benefits of our solution. Related work is described by Section 5 while Section 6 presents the paper conclusion and future work.

2 A CPU-GPU Component-Based Design

A component-based application with heterogeneous computing uses a hardware platform which contains several computation nodes with different architectures, such as multi-core CPU and GPU. Due to its parallel processing power, GPUs bring as benefits an increased computation diversity and power. The disadvantage of using, adjacent to CPU, a different processing unit with its own architecture and memory system, reflects in an increased complexity of the software application. Another drawback is that the GPU can not be used independently of the CPU. Considered as the brain of the system, the CPU is the one that triggers all the GPU specific operations, such as data transfer activities between the main RAM memory and GPU memory system. Hence, there is a high communication between the two processing units. In a multi-node system, the connected CPU and GPU-based components are desired to be placed on the same heterogeneous computation node to not negatively affect the total system performance.

We consider that a repository contains two types of components; one type requires only the CPU for its functionality, and another uses also GPU to fulfill its functionality. For a specific function, we may have both types of components in the repository as different version implementations of the respective function. For example, an image processing component may have three component versions in the repository, one that uses only the CPU and the other two that uses also the GPU. The components that use the GPU, may have a different usage of the GPU resources, hence, their properties are different, e.g., in GPU memory and computation threads usage. From a repository populated with CPU and GPU software components, by having different component combinations (e.g., CPU-GPU, CPU-GPU-GPU), several potential alternatives with the same functionality can be constructed. In a complex system with multiple heterogeneous hardware nodes, an alternative composed of only GPU-based components may not be a feasible allocation solution due to the hardware resource limitations. Hence, the allocation process should consider all the alternatives in its allocation activity.

In the following example, we describe a component-based design for a demonstrator with heterogeneous embedded CPU-GPU hardware. The demonstrator, an underwater robot, is developed at Mälardalen University, Sweden [2]. The purpose of the demonstrator is to autonomously navigate under water in e.g., tracking various objects, using its vision system that contains two (front and bottom) cameras.



(a) The vision frame activities (b) Two vision alternatives

Fig. 1 Data activities and alternatives of the front vision system

To develop the demonstrator's front vision system, CPU and GPU software components are used from the system repository, as follows. In our example, a simplified vision system is composed of three components. The *FrontCamera* component communicates with the physical camera and receives frames from the underwater environment. The frames are forwarded to the *ImageFilter* component which processes and converts them into black-and-white frames. The *ObjectDetector* component analyzes the filtered frames to detect various objects (e.g., red buoy). Figure 1a describes the data flow activity of the front vision system, using an existing underwater frame from the demonstrator front camera.

Only the *ImageFilter* and *ObjectDetector* benefit from utilizing a GPU due to nature of their functionality, i.e., image processing. For *ImageFilter* component, the repository contains two versions, i.e., a CPU and GPU-based component, each with different properties. Using CPU-GPU combinations of the repository's components results in different vision alternatives. Figure 1b depicts the alternatives

of the front vision system based on the repository content. In our case, for image processing, GPU provides a better performance than the CPU due to its massive parallel processing power. The alternative that contains two GPU-based components (i.e., *ImageFilter_GPU* and *ObjectDetector_GPU*), has a high usage of the GPU (e.g, GPU memory and computation threads) and a good performance. The alternative which contains only one GPU-based component, uses less the GPU resources and has a lower performance compared to the previous alternative.

For our small example, all the front vision system alternatives have the same simple design, i.e., three connected components. In a more advanced application, the vision system may have different design for different alternatives. For example, an alternative may contain three GPU-based components, while another may have two GPU-based components and two CPU-based components.

In general, considering all possible alternatives and selecting one best fit to be allocated on a heterogeneous hardware node, brings an information load that may hurt the efficiency of the allocator. The allocator needs to take in consideration each component properties from each variant and to have a component-to-variant mapping. Also, because of the CPU-GPU dependency, the GPU-based components are desired to be deployed with their connected CPU-based components, otherwise e.g., a high communication between them may negatively influence the system overall performance. The complexity of the allocation process is increased by the number of the alternatives and their content components, which may affect the allocator performance (e.g., scalability, allocation time).

3 Solution Overview

To ease the software-to-hardware allocation process, we propose a 2-layer architecture view for the component-based design of the system. The first layer contains the details of the system alternatives. We propose to abstract away the complexity of these system alternatives, by using an abstracted second layer. This second layer compacts a system alternatives into a single component with many variants. A component variant is characterized by properties that reflect the components' properties of the corresponding alternative.

Figure 2 presents a component with two variants of the front vision system. The figure illustrates only one alternative which is composed of three components. The elements of the alternative, i.e., components and communication links, are characterized by different properties such as RAM memory usage or bandwidth. Each variant hides away the full details of its corresponding alternative, and exposes the overall properties of the abstracted structure.

The allocator, using the abstracted layer properties, chooses the fittest component variant w.r.t. the rest of the system properties and other constraints. Once a component variant is selected, the system is described by the first layer where the alternative that corresponds to the selected component variant is displayed. Figure 3 presents the steps of our approach. From the components contained by the repository, several

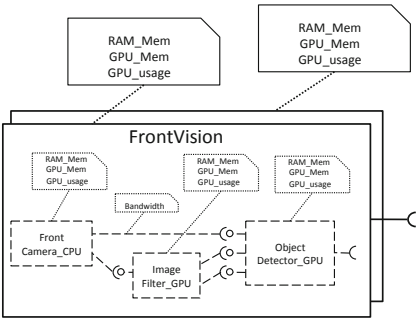


Fig. 2 Two front vision alternatives abstracted to a component with two variants

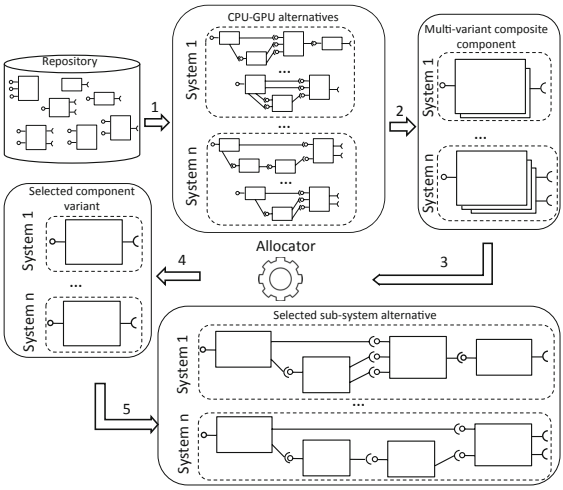


Fig. 3 The steps of the allocation process using our approach

alternatives for different systems are constructed in step 1. These may be constructed either manually by the system developer, or automatically. In step 2, all alternatives with the same functionality are compacted into a multi-variant component. Although not described in the figure, at this step, the alternatives properties are synthesized into the properties of the component variants. The allocator, which is an automatic system, receives in step 3, the properties of the multi-variant components. Using also other information (e.g., software and hardware models, constraints), the allocator selects, during step 4, the fittest component variants for the system allocation scheme. Once the variants are selected, the last step exposes the detailed alternatives abstracted by the selected variants.

In general, using a high-level layer containing simplified multi-variant components to abstract the details of the alternative, eases the information load of the allocation process. Instead of considering each component properties and other information (e.g., the communication links between connected components) from each

alternative, the allocator considers only the abstracted variant properties, where the rest of the information is hidden away. Also, the concern of distributing connected CPU-GPU components to the same computation node is now implicit considered in the allocation process.

4 Running Example

To illustrate the usage of our solution, we describe the allocation process of the underwater robot system, used during the paper. For the front and bottom vision systems, we developed several GPU and CPU-based components as follows. The repository of the vision systems contains three CPU-based components, i.e., *FrontCamera_CPU*, *BottomCamera_CPU* and *ImageFilter_CPU*, and two GPU-based component, i.e., *ImageFilter_GPU* and *ObjectDetector_GPU*. In our simple example, we characterized each CPU-based component with RAM memory usage and performance. The GPU-base components are defined by similar properties, such as the GPU memory usage and performance (e.g., execution time on GPU).

Dealing with a robot where performance and real-time responses may be crucial, a performance-driven component-based design of the front or bottom vision system would be composed mostly of GPU-based components. Not knowing beforehand various criteria such as the available hardware resources, the software-to-hardware allocation may be infeasible. For example, an alternative of the front vision system composed of only GPU-based components may demand more GPU computation threads than the hardware limitations, without even considering the resource requirements of the bottom vision system.

In the upper part of Figure 4, the hardware and the software architectures of the robot are presented. Figure 4a describes the hardware platform that contains two processing nodes, where only one has GPU capabilities. The nodes communicate over a CAN bus. To simplify our example, each hardware element is characterized by a minimal set of properties such as the available RAM and GPU memory and the GPU processing power. We use GPU threads capacity as a simple metric for specification of GPU computation power. In a more complex example, this property may be extended to include other metrics such as registers per thread.

Figure 4b describes the component-based design of the robot. The main component, *DecisionCenter*, controls the system settings (e.g., water pressure, color calibration specifications) and the robot missions. The robot propellers are controlled by the *MovementNavigation* component that maneuver the underwater robot based on the commands received from the *DecisionCenter* component. There are two vision systems, one for the front camera and another for the bottom camera. The *VisionManager* component takes decisions based on the information received from the vision systems. Both of the vision systems are constructed using components from the repository content. The front vision system is displayed as a component with two variants and the bottom vision system is described by a three variant component.

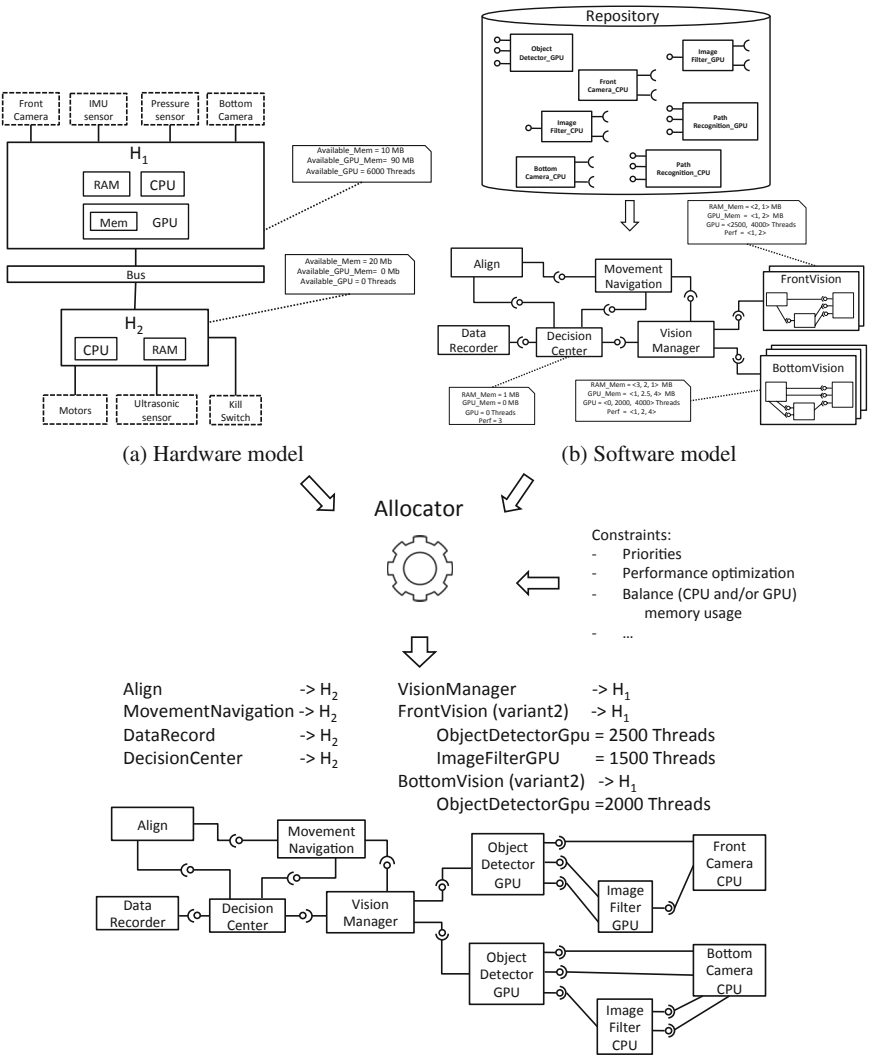


Fig. 4 The allocation process of the underwater robot demonstrator

Each software component, such as *DecisionCenter* component, is characterized by extra-functional properties and performance, as seen in the Figure 4b. For the multi-variant components, we specify their properties in the following way. The *CPU_mem*, *GPU_mem* and *Perf* attributes are described as a sequence of values, where each value represents the resource usage of the corresponding variant. The *Perf* property is a sequence of variant values; the higher the value is, the better the variant performance is. For example, Table 1 presents the properties of the front

Table 1 The properties of the 2-variant front vision component

Component variant	RAM_Mem (Mb)	GPU_Mem (Mb)	GPU (Threads)
1	3	1	2500
2	1	2	4000

vision multi-variant component. The component has two variants, where the first one is using 3Mb of RAM memory, 1 Mb of GPU memory and 2500 of GPU threads. The second variant, having two GPU-based components, uses more of the GPU resources (e.g., 2 Mb of GPU memory and 4000 Threads) than the previous variant.

Table 2 presents a detailed description of the properties of the two alternatives which corresponds to the front vision bi-variant component. Each alternative is composed of three components as follows. The first alternative has two CPU-based components and one GPU-based component, while the second alternative has two GPU-based components and one CPU-based component. The amount of information is much larger; each alternative is characterized by nine properties while a component variant has three properties.

Table 2 The properties of front vision alternatives

Alternative	Component	RAM_Mem (Mb)	GPU_Mem (Mb)	GPU (Threads)
1	FrontCameraCPU	2	0	0
	ImageFilterCPU	1	0	0
	ObjectDetectorGPU	0	1	2500
2	FrontCameraCPU	1	0	0
	ImageFilterGPU	0	1	1500
	ObjectDetectorGPU	0	1	2500

For our case study, we use simple properties such as static memory or thread usage; for synthesizing the variant properties, we use a simple addition operation. In general, other properties of a variant may also be derived from the components' properties of the corresponding alternative using different methods or techniques, when feasible.

The robot relies on the front vision system as the main vision system, and uses the bottom vision system as a secondary vision system when e.g., the front vision does not detect anything. Hence, the front vision has higher priority in accessing the GPU hardware resources. After the allocator receives the hardware and software information and other constraints such as priorities of the vision systems, the system allocation scheme is computed, as presented in the upper part of the Figure 4c. On the heterogeneous node H_1, both of the vision systems are allocated, while the rest

of the components are allocated on the H_2 computing node. Having a higher priority, front vision system is allowed to access more of the GPU resources than the bottom vision system. Hence, the allocator selects the front vision variant that contains two GPU-based components: *ImageFilterGPU* and *ObjectDetectorGPU*. Because most of the GPU processing resources are occupied by the front vision system, the allocator selects the bottom vision variant that contains only one GPU-based component.

In our simplified example, we used a small number of simple component properties. We did not consider other information such as the communication link properties between components. Also, the analysis and synthesizes of different properties into a set of variant properties is beyond the subject of this paper.

5 Related Work

There is a lot of research done in the allocation and software-to-hardware optimization domain, described in surveys like [3] and [4]. Several works present the tasks distribution used in the automotive industry [10], where an optimization method allocates task onto ECUs with different memory capacity and processing power. Various criteria are addressed by the allocation process, such as balancing local memory [12] for safety-critical multi-core systems or balancing the CPU processing power [8].

An allocator that covers CPU-GPU component allocation is described in our previous work [6]. The work translates an allocation optimization model into a mixed-integer programming solver (SCIP [1]). The solver, based on the software and hardware inputs, calculates feasible allocation schemes. The allocation model is formally describing the software component-model, the CPU-GPU hardware model and various allocation constraints such as balancing resources (i.e., memory, CPU and GPU computation power) and performance optimization. The work is limited by not considering multiple alternatives in the allocation process. The evaluation section describes the time and scalability limitations due to the amount of components and hardware information. Extending the work with our solution may result in a more accurate allocation with a better scalability property.

6 Conclusion

In this paper, we have proposed a solution to ease the software-to-hardware allocation process. Using a 2-layer component-based design for heterogeneous CPU-GPU embedded systems, we decreased the information load delivered to the allocator which can positively influence the total allocation process efficiency. Another advantage of our solution is that it can improve the allocation scalability by reducing the information load, making possible the allocator to handle more complex systems and component combinations. The disadvantage of our work is that it introduces new

steps in the allocation process but can be worth accepting when the allocator gains benefits in efficiency.

For future activities, we propose to extend our previous work to include the solution presented in this paper and compare the efficiency results between the two work versions. Covering a limited number of simple component properties, a future work extension may include more extra-function properties and GPU specific properties to describe the software model. Another future work continuation may include an adaptation or development of an existing method or technique, to allow us to synthesize the component properties to the variant properties.

Acknowledgements Our research is supported by the RALF3 project - (IIS11- 0060) [2] through the Swedish Foundation for Strategic Research (SSF).

References

1. Achterberg, T.: Constraint Integer Programming. Ph.D. thesis, TU Berlin, Germany (2007)
2. Ahlberg, C., Asplund, L., Campeanu, G., Ciccozzi, F., Ekstrand, F., Ekström, M., Feljan, J., Gustavsson, A., Sentilles, S., Svogor, I., Segerblad, E.: The black pearl: an autonomous underwater vehicle. Technical report, Mälardalen University, June 2013. Published as part of the AUVSI Foundation and ONR's 16th International RoboSub Competition, San Diego, CA
3. Aleti, A., Buhnova, B., Grunske, L., Koziolek, A., Meedeniya, I.: Software architecture optimization methods: A systematic literature review. IEEE Transactions on Software Engineering **39**(5), 658–683 (2013)
4. Balsamo, S., Marco, A.D., Inverardi, P., Simeoni, M.: Model-based performance prediction in software development: A survey. IEEE Transactions on Software Engineering **30**(5), 295–310 (2004)
5. Baruah, S.K.: Task partitioning upon heterogeneous multiprocessor platforms. In: 2004 Proceedings of 10th IEEE Real-Time and Embedded Technology and Applications Symposium, RTAS 2004, pp. 536–543 (2004)
6. Campeanu, G., Carlson, J., Sentilles, S.: Component allocation optimization for heterogeneous CPU-GPU embedded systems. In: The 40th Euromicro Conference on Software Engineering and Advanced Applications SEAA 2014 (2014)
7. Geronimo, D., Lopez, A.M., Sappa, A.D., Graf, T.: Survey of pedestrian detection for advanced driver assistance systems. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(7), 1239–1258 (2010)
8. Ma, P.-Y.R., Lee, E., Tsuchiya, M.: A task allocation model for distributed computing systems. IEEE Transactions on Computers **31**(1), 41–47 (1982)
9. Michel, P., et al.: GPU-accelerated real-time 3D tracking for humanoid locomotion and stair climbing. In: 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007, pp. 463–469. IEEE (2007)
10. Moser, I., Mostaghim, S.: The automotive deployment problem: a practical application for constrained multiobjective evolutionary optimisation. In: IEEE Congress on Evolutionary Computation, pp. 1–8. IEEE (2010)
11. Nguyen, H.: GPU Gems 3, 1st edn. Addison-Wesley Professional (2007)
12. Voss, S., Schätz, B.: Deployment and scheduling synthesis for mixed-critical shared-memory applications. In: Rozenblit, J.W. (ed.) ECBS, pp. 100–109. IEEE (2013)
13. Xu, F., Mueller, K.: Real-time 3d computed tomographic reconstruction using commodity graphics hardware. Physics in medicine and biology **52**(12), 3405 (2007)

[2] http://www.mrtc.mdh.se/projects/ralf3/

Part V
High-Performance Computing Architectures

Block-Based Approach to 2-D Wavelet Transform on GPUs

Michal Kula, David Barina and Pavel Zemcik

Abstract This paper introduces a new approach to computation of 2-D discrete wavelet transform on modern GPUs. The proposed approach involves block-based processing enabling one seamless transform even for high resolution input data. Inside the blocks, two distinct methods can be used – either separable or non-separable 2-D lifting scheme. Furthermore, the paper presents a comparison of the proposed approach under different conditions to the best existing methods, whereas our approach consistently outperforms the other ones. Our methods are implemented using the OpenCL framework and tested on a wide range GPUs.

Keywords Discrete wavelet transform · Image processing · Lifting scheme · Graphics processing units · Memory barrier

1 Introduction

The 2-D discrete wavelet transform (DWT) is the signal-processing transform suitable for decomposition of the analysed 2-D signal into several scales. On each scale, three directional subbands are formed. These are usually referred to as HL, LH, and HH subbands. The 2-D transform is defined as separable product of 1-D transforms performed sequentially on rows and columns (or vice versa). Each of these one-dimensional transforms can be computed through either the convolution or the lifting scheme. Different strategies of 2-D DWT implementation were developed for various computational platforms.

In this paper, we focus on implementation of DWT using modern graphics cards (GPU) capable of a general-purpose computing. In these architectures, the GPU contains thousands of stream processors that are clustered into blocks. All processors

M. Kula · D. Barina(✉) · P. Zemcik
Brno University of Technology, Brno, Czech Republic
e-mail: {ikula,ibarina,zemcik}@fit.vutbr.cz

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_56

in each block execute the same instruction with different operands at one time. The blocks are grouped into multiprocessors which form the basic functional units of the GPUs. The thread scheduler allocates as many work groups to multiprocessors as their resources allow. The work groups are defined as a group of threads that can interoperate with each other using the local memory and memory barriers. Thus, the resources, such as the local memory size, should be minimized. The allocated work groups created by OpenCL framework is then divided into warps (hardware blocks with 32 threads). Execution instructions of these warps on blocks of processors are provided using warp schedulers dynamically. Global memory accesses in warp should be coalesced. Otherwise, additional memory operations are executed. The local memory is organized into banks. Access to the same banks from warp causes serialization. This issue is referred to as a bank conflict. The serialization of local memory operations and uncoalesced global memory access can cause a performance degradation.

This paper is further focused on the OpenCL framework.[1] OpenCL is a framework for general-purpose parallel programming across multiple device types. In this framework, a platform independent executable program is called the kernel. The kernel is executed on required number of threads that identify their data and control flow by their N-dimensional indices. These threads are organized into work groups with identical user-defined number of threads. The threads in such a group can cooperate with each other through local memory and barriers.

Several methods for the 2-D DWT computation using GPU have been published in the last decade. For example, the simplest row–column methods transform the whole image at once. Usually, the transposition is needed between the horizontal and vertical part. Furthermore, the block-based methods transform the image using smaller blocks utilizing the row-column method inside. Unfortunately, such a method results in several independent transforms instead a single seamless one. Finally, the pipelined methods transform the image using column strips while employing the sliding window on them.

In this paper, we propose two novel block-based methods computing the seamless 2-D transform. The first of them employs the separable (row–column) lifting scheme inside the overlapping blocks. The second uses a non-separable lifting scheme recently proposed. Both of the proposed methods consistently outperform the existing methods.

The rest of the paper is organized as follows. The Related Work section summarizes the state of the art, especially existing GPU implementations. The heart of our work is presented in Block-Based Approach section. First, we propose the separable transform. Further in the text, the non-separable method is discussed. Finally, Conclusions section summarizes the paper and outlines the future work.

2 Related Work

This section takes a closer look at the discrete wavelet transform and revises the state of the art of its implementation on contemporary graphics cards.

[1] http://www.fit.vutbr.cz/research/prod/index.php?id=434.

The DWT can be understood as a transform suitable for decomposition of a signal into low-pass and high-pass frequency components. Usually, such a decomposition is performed at several scales resulting in a multi-scale signal representation. At this point, we are considering one-dimensional signals. The transform of 2-D signals is computed through the tensor product of these 1-D transforms. For various requirements, different strategies of 2-D transform computation emerged. Going back to 1-D transform, as the discrete wavelet transform is a linear one, the decomposition into the low-pass and high-pass components can be performed through a convolution scheme with two filters. However, the more efficient computational scheme according to the number of arithmetic operations exists. This scheme is referred to as the lifting scheme. Additionally, using this scheme, the whole signal can be transformed in-place. Specifically, any discrete wavelet transform can be factored into a finite sequence of lifting steps. These steps alternately update odd and even intermediate results using short FIR (finite impulse response) filters. When evaluating this scheme, intermediate results can be appropriately shared between neighbouring coefficients.

The discrete wavelet transform is often used as a basis for sophisticated compression algorithms. This paper focuses on a popular CDF (Cohen-Daubechies-Feauveau) 9/7 wavelet. This wavelet is used, e.g., in JPEG 2000 image compression standard. In [2], Daubechies and Sweldens factored CDF 9/7 wavelet into four successive lifting steps, employing short symmetric two-taps FIR filters. A data-flow diagram of the factorization (without scaling) is depicted in Fig. 1, where, the α, β, γ, δ are real constants specific to CDF 9/7 transform. Formally, the forward transform in Fig. 1 can be expressed by the dual polyphase matrix

$$\tilde{P}(z) = \begin{bmatrix} 1 & \alpha\left(1+z^{-1}\right) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta\,(1+z) & 1 \end{bmatrix} \begin{bmatrix} 1 & \gamma\left(1+z^{-1}\right) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \delta\,(1+z) & 1 \end{bmatrix}. \quad (1)$$

In this paper, we consider the lifting scheme only, as it is usually a better alternative. A detailed comparison of the convolution and lifting schemes on GPUs was addressed, e.g., in [11] and [12].

The implementation of this transform was comprehensively studied on various platforms including the modern GPUs. Considering this scenario, the input image has to be initially transferred from main memory into memory on the graphics card. Similarly, the resulting coefficients could be transferred back. Having the input 2-D image in the GPU global memory, different strategies of 2-D DWT implementation can be used. These strategies can be divided into three groups – row–column, block-based, and pipelined methods.

The row–column method applied on the entire 2-D image was used for instance in [11], [12], [3], [1], [4], [5]. In [3] and [1], data transposition was performed in between the horizontal and vertical series of 1-D transforms. In [11] and [12], Tenllado *et al.* adapted the discrete wavelet transform on GPU fragment shaders. As this paper is focused on the OpenCL framework, we will not discuss their paper in more details. The other cited papers are focused on the CUDA architecture. In [3], the convolution scheme is applied on each row. Then, the image matrix is transposed and the convolutions are applied on each column. Finally, the image is transposed

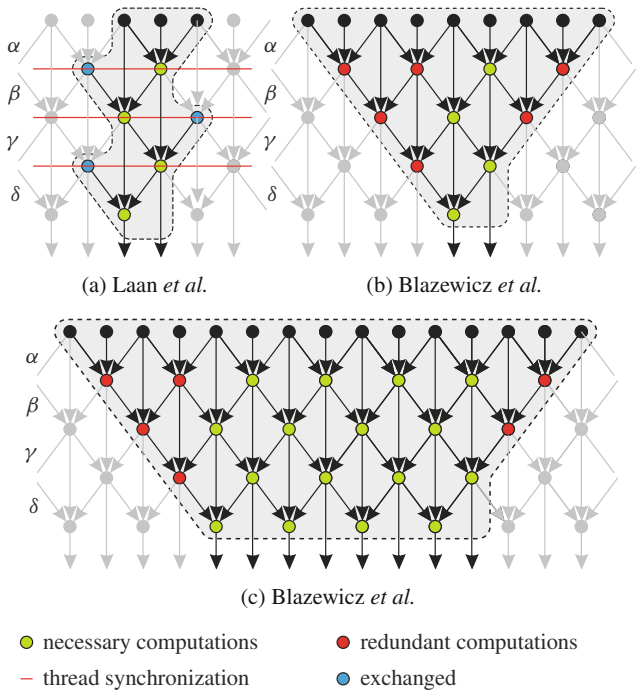


Fig. 1 A portion of the data-flow graph attributable to individual threads. The method of (a) Laan and two methods used by Blazewicz – with (b) one, and (c) four pairs.

back. In [4] and [5], V. Galiano *et al.* compared several CUDA implementations of DWT. They used the CDF 9/7 wavelet and convolution-based algorithm on entire rows/columns. Their fastest implementation uses the coalesced memory access.

In [1], the authors calculate the wavelet transform through 4 kernels. The first kernel performs an image transposition using work groups of size 16×16 threads, where the thread processes one image element. To ensure coalesced global memory access, the transposition in the shared memory is used rather than directly in the global memory. In the second kernel, the vertical wavelet transform is performed as follows. Each thread loads its elements from the global memory and stores them into the shared memory. Then, the adjacent elements, that are required for the computation of the output coefficient, are loaded from the shared memory into registers. The threads compute their output coefficients using 4 steps of the wavelet scheme independently to each other (with no synchronization). When the computation is finished, the output coefficients are written back to the global memory. The third and the fourth kernels calculate the image transposition and the horizontal wavelet transform in the same way as the first two kernels. The calculations that are performed by a single thread using the approach described can be seen in Fig. 1b and Fig. 1c.

The pipelined approach was used in [8] and [9]. In [8], Laan *et al.* accelerated the Dirac video codec using the CUDA platform. In [9], the authors provided a

detailed analysis of the DWT implementation using the lifting scheme on the CUDA platform. They focused on 2-D and 3-D methods of DWT implementation using several wavelets including CDF 9/7. In the horizontal part of their transform, each work group is mapped to a single image row. Each thread computes one coefficient per a single step and shares it with other threads. Because of non-atomic instructions issued in whole group, memory barrier is needed in between each two steps. See Fig. 1a. The vertical part of their transform maps each work group to multiple vertical strips with a width that ensures coalesced global memory accesses and bank-conflict-free shared memory transfers.

Another row-column approach was used in [7]. The horizontal transform is computed in the same way as Blazewicz *et al.* did. The vertical transform is computed using 32 coefficients wide strips per work group like in Laan's implementation. The difference between the Laan's and Kucis's vertical methods lies in processing assigned to a single thread. Kucis *et al.* used rotated Blazewicz's approach with 2 pairs per thread mapping. Moreover, Kucis *et al.* also demonstrated that their approach outperforms Laan's and Blazewicz's ones. The approach of Kucis is used as a reference approach and labeled as *Kucis2014*.

The approaches in [10] as well as [1] are focused on the lifting scheme. Their implementations split the image into small tiles and perform several independent transforms on each of them. Thus, they performed several independent transforms (introducing a block effect) which is different and much easier task comparing to what we are dealing with in this paper.

As it can be seen, the problem of the efficient 2-D discrete wavelet transform implementation on conventional GPUs was fairly well studied. However, we see several gaps which can allow for additional speedups. Specifically, only the separable 2-D schemes were examined so far. These schemes require to pass the results through the global memory, while causing unnecessary memory traffic.

3 Block-Based Approach

The heart of our work is presented in this section. At the beginning, we propose the separable block-based method. Afterwards, we discuss the block-based method utilizing non-separable lifting scheme recently proposed. The performance comparison of the proposed methods is shown in Fig. 4. As it can be seen, the block-based methods perform consistently faster compared to the best of the existing methods. Our implementation is based on the OpenCL framework. All of the algorithms are evaluated using AMD R9 290X and NVIDIA TitanX graphics cards. The main benefit of the block-based methods is the reduction of memory access count as the data is read as well as written only once.

3.1 Separable Method

Except for the sliding window, our separable block-based approach uses the same scheme as the Laan's method. The threads in each work group are responsible for processing of 2×2 input coefficients. At the beginning, the thread loads their coefficients from the global memory and stores them into separate shared memory locations. The computation is briefly illustrated in Fig. 2. In the first step of the horizontal pass, each of the threads computes the LH coefficient using two LL coefficients of the thread itself and the thread on the right. Additionally, the HH coefficient is computed using HL coefficients of the current thread and the thread on the right. In the second step, the computation of LL and HL coefficients is performed in the same way as the computation of the LH and HH coefficients. After that, these two steps are repeated with a substitution of α and β with γ and δ coefficients.

The vertical steps are performed in the same way as the horizontal steps except for a rotation of the scheme by 90 degrees. Unlike the horizontal pass, synchronization using the memory barrier is required between the steps. Horizontal steps are synchronization-free thanks to the atomicity of hardware instructions. Fig. 2 shows individual steps of the underlying data-flow graph.

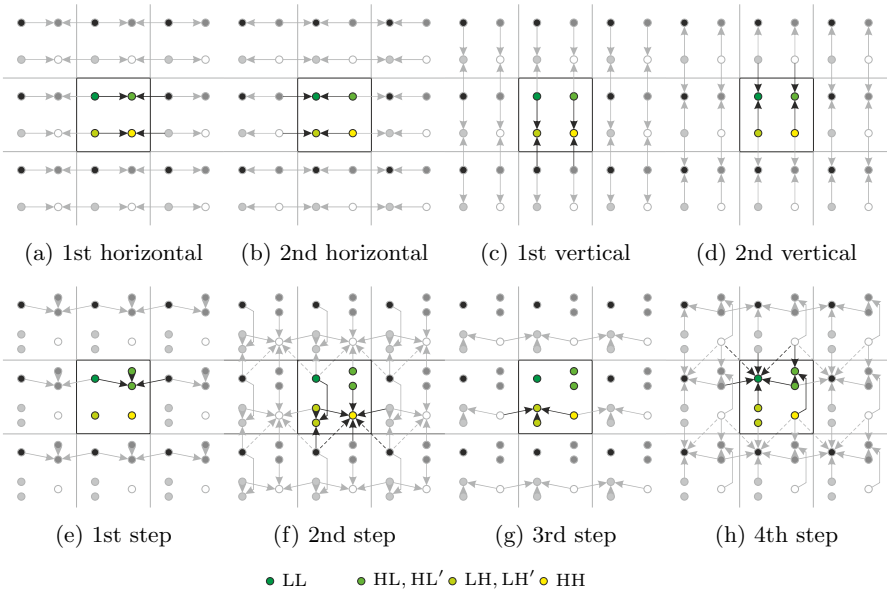


Fig. 2 The separable (top) and non-separable block-based approach (bottom).

3.2 Non-Separable Method

In [6], the authors derived a non-separable 2-D lifting scheme for CDF 5/3 and subsequently CDF 9/7 transforms. As initial step of CDF 5/3 transform, the input signal is split into quadruples (LL, HL, LH, HH). Then, lifting steps leading to the calculation HH coefficients are performed. This is followed by parallel computation of the HL and LH coefficients. In the third step, the LL coefficient is updated. Finally, the coefficients can be scaled. The scheme for CDF 9/7 comprises two these connected transforms.

Motivated by the work of Iwahashi *et al.* [6], we have reorganized the elementary lifting FIR filters in order to obtain a highly parallelizable scheme suitable for the modern GPUs. The main purpose of this modification is to minimize the number of memory barriers that slow down the calculation. As a result, we get several non-separable two-dimensional FIR filters. For their description, we employ the well known z-transform notation. The transfer function of the two-dimensional FIR filter $x(k_m, k_n)$ is defined as

$$X(z_m, z_n) = \sum_{k_m=-\infty}^{\infty} \sum_{k_n=-\infty}^{\infty} x(k_m, k_n)\, z_m^{-k_m} z_n^{-k_n}, \tag{2}$$

where m refers to the horizontal axis and n to the vertical one. Moreover, to keep consistency with [6], the $H^*(z_m, z_n) = H(z_n, z_m)$ denotes a filter transposed to the $H(z_m, z_n)$. Furthermore, the $\overline{H}(z_m, z_n) = H(z_n^{-1}, z_m^{-1})$ denotes a filter reversed along the m- as well as n-axis. Coupled together, the $\overline{H}^*(z_m, z_n)$ denotes a transposed and reversed filter to the original $H(z_m, z_n)$. The scheme we formed is composed of three elementary filters F, G, H given by

$$\begin{bmatrix} \mathrm{F_a} \\ \mathrm{G_a} \\ \mathrm{H_a} \end{bmatrix} = \begin{bmatrix} F_a(z_m, z_n) \\ G_a(z_m, z_n) \\ H_a(z_m, z_n) \end{bmatrix} = a \begin{bmatrix} 1 \\ z_n \\ 1 + z_m \end{bmatrix}, \tag{3}$$

where a denotes a filter parameter. The filters above are assembled into more complex operations. Our scheme consists of two halves between which a memory barrier is placed. The first half of the scheme uses the following filters. Similarly, the second half uses these filters in the reverse orientation. Due to the limited place, we have made a small abuse of notation. Instead of the full notation $H(z_m, z_n)$, we only use a shortened labeling, such as H.

$$\begin{bmatrix} \mathrm{F_a} \\ \mathrm{G_a} \\ \mathrm{H_a} \\ \mathrm{H_a^*} \\ \mathrm{G_a H_a} \end{bmatrix} = \begin{bmatrix} a \\ a\, z_n \\ a\,(1 + z_m) \\ a\,(1 + z_n) \\ a^2\,(z_n + z_m z_n) \end{bmatrix}, \qquad \begin{bmatrix} \overline{\mathrm{F}}_\mathrm{a} \\ \overline{\mathrm{G}}_\mathrm{a} \\ \overline{\mathrm{H}}_\mathrm{a} \\ \overline{\mathrm{H}}_\mathrm{a}^* \\ \overline{\mathrm{G}}_\mathrm{a} \overline{\mathrm{H}}_\mathrm{a} \end{bmatrix} = \begin{bmatrix} a \\ a\, z_n^{-1} \\ a\,(1 + z_m^{-1}) \\ a\,(1 + z_n^{-1}) \\ a^2\,(z_n^{-1} + z_m^{-1} z_n^{-1}) \end{bmatrix} \tag{4}$$

Finally, our scheme is composed of four steps referred to as S^1 to S^4. Between the second S^2 and the third S^3 step, the memory barrier must be inserted in order to properly exchange intermediate results. Additionally, our scheme requires the induction of two auxiliary variables per each quadruple of coefficients LL, HL, LH, and HH. These are denoted as HL′, LH′. This is valid regardless of their initial as well as final values. The scheme

$$\mathbf{y} = S_\beta^4 S_\beta^3 S_\alpha^2 S_\alpha^1 \mathbf{x} \tag{5}$$

describes the relation between input $\mathbf{x}$ and output $\mathbf{y}$ vectors

$$\left[\ \mathrm{LL}\ \mathrm{HL}\ \mathrm{LH}\ \mathrm{HH}\ \mathrm{HL}'\ \mathrm{LH}'\ \right]^T. \tag{6}$$

Each single thread of the work group is responsible of one such a vector.

Regarding this notation, the individual steps are defined as follows. For better understanding, the signal-processing block diagram of this scheme is shown in Fig. 3. In addition, the operations are graphically illustrated in Fig. 2.

$$S_\alpha^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \mathrm{H}_\alpha & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$$S_\alpha^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \mathrm{G}_\alpha\mathrm{H}_\alpha & \mathrm{G}_\alpha & \mathrm{H}_\alpha & 1 & \mathrm{F}_\alpha & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \mathrm{H}_\alpha^* & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

$$S_\beta^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \overline{\mathrm{H}}_\beta & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

$$S_\beta^4 = \begin{bmatrix} 1 & 0 & \overline{\mathrm{F}}_\beta & \overline{\mathrm{G}}_\beta\overline{\mathrm{H}}_\beta & \overline{\mathrm{H}}_\beta & \overline{\mathrm{G}}_\beta \\ 0 & 0 & 0 & \overline{\mathrm{H}}_\beta^* & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

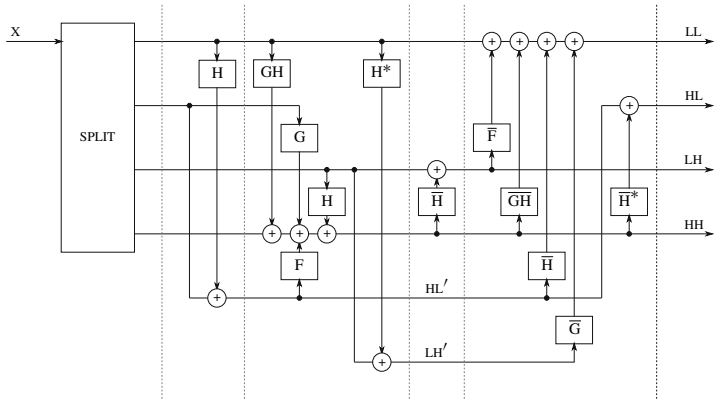
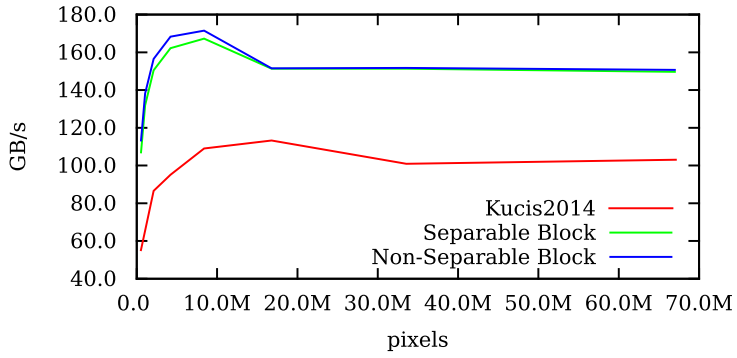
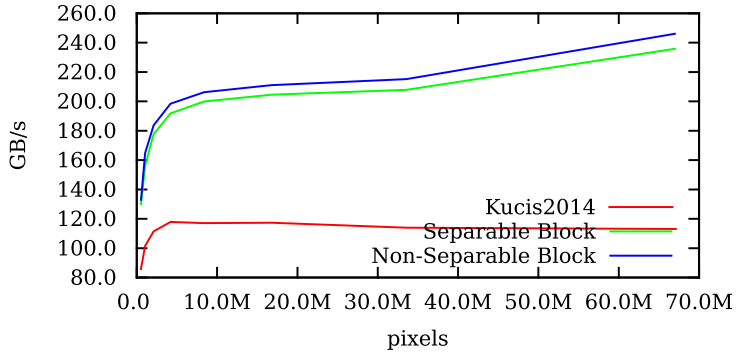


Fig. 3 Block diagram of the proposed non-separable scheme.



(a) AMD R9 290X



(b) NVIDIA TitanX

Fig. 4 Throughput performance. *Kucis2014* is the reference state-of-the-art method.

Compared to [6], the total number of arithmetic operations has been reduced form 24 to 20. The calculation of CDF 9/7 transform comprises two of these connected transforms (the first one with α, β, the second with γ, δ) between them another barrier is placed. In total, the calculation contains 3 memory barriers.

4 Conclusions

We have presented two novel block-based approaches to 2-D wavelet transform using modern GPUs. These approaches can handle high resolution images while producing the seamless transform. Both of the proposed methods consistently outperform the existing methods with all tested GPUs.

The first presented approach utilizes classical separable 2-D lifting scheme. Whereas the second approach employs a novel 2-D non-separable scheme. Considering the second one, we have minimized the number of memory barriers. Moreover, as compared to the existing non-separable scheme, the total number of arithmetic operations has been reduced form 24 to 20.

The future work includes behavior of the proposed methods under a multi-scale decomposition. Another direction of our research may include a connection with some practical application (e.g., JPEG 2000 scheme).

Acknowledgements This work has been supported by the TACR Competence Centres project V3C – Visual Computing Competence Center (no. TE01020415).

References

1. Błażewicz, M., Ciżnicki, M., Kopta, P., Kurowski, K., Lichocki, P.: Two-dimensional discrete wavelet transform on large images for hybrid computing architectures: GPU and CELL. In: Euro-Par 2011: Parallel Processing Workshops, LNCS, vol. 7155, pp. 481–490. Springer (2012)
2. Daubechies, I., Sweldens, W.: Factoring wavelet transforms into lifting steps. Journal of Fourier Analysis and Applications **4**(3), 247–269 (1998)
3. Franco, J., Bernabe, G., Fernandez, J., Acacio, M.: A parallel implementation of the 2D wavelet transform using CUDA. In: 17th Euromicro International Conference on Parallel, Distributed and Network-based Processing, pp. 111–118, February 2009
4. Galiano, V., López, O., Malumbres, M., Migallón, H.: Improving the discrete wavelet transform computation from multicore to GPU-based algorithms. In: Proceedings of the 11th International Conference on Computational and Mathematical Methods in Science and Engineering (CMMSE), pp. 544–555 (2011)
5. Galiano, V., López, O., Malumbres, M., Migallón, H.: Parallel strategies for 2D discrete wavelet transform in shared memory systems and GPUs. The Journal of Supercomputing **64**(1), 4–16 (2013)
6. Iwahashi, M., Kiya, H.: Non separable two dimensional discrete wavelet transform for image signals. In: Discrete Wavelet Transforms – A Compendium of New Approaches and Recent Applications. InTech (2013)

7. Kucis, M., Barina, D., Kula, M., Zemcik, P.: 2-D discrete wavelet transform using GPU. In: 5th Workshop on Application for Multi-Core Architectures, pp. 1–6. IEEE Computer Society (2014)
8. van der Laan, W., Roerdink, J.B.T.M., Jalba, A.: Accelerating wavelet-based video coding on graphics hardware using CUDA. In: Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, pp. 608–613, September 2009
9. van der Laan, W.J., Jalba, A.C., Roerdink, J.B.T.M.: Accelerating wavelet lifting on graphics hardware using CUDA. IEEE Transactions on Parallel and Distributed Systems **22**(1), 132–146 (2011)
10. Matela, J.: GPU-based DWT acceleration for JPEG2000. In: Annual Doctoral Workshop on Mathematical and Engineering Methods in Computer Science, pp. 136–143 (2009)
11. Tenllado, C., Lario, R., Prieto, M., Tirado, F.: The 2D discrete wavelet transform on programmable graphics hardware. In: Visualization, Imaging and Image Processing Conference 2004, pp. 808–813, September 2004
12. Tenllado, C., Setoain, J., Prieto, M., Pinuel, L., Tirado, F.: Parallel implementation of the 2D discrete wavelet transform on graphics processing units: Filter bank versus lifting. IEEE Transactions on Parallel and Distributed Systems **19**(3), 299–310 (2008)

Deriving CGM Based-Parallel Algorithms for the Optimal Binary Search-Tree Problem

Vianney Kengne Tchendji, Jean Frédéric Myoupo and Gilles Dequen

Abstract This paper presents a methodology to derive CGM (*Coarse Grain Multicomputer*) parallel algorithms for the cost of the Optimal Binary Search Tree Problem (OBST Problem). Depending on the parameter we want to optimize, we derive an algorithm accordingly. Therefore a load balancing, an efficient and a minimum communication rounds algorithms are obtained. Our CGM algorithms use p processors, each with $O(n^2/p)$ local memory. The best one in communication requires $O(p^{1/2})$ communication rounds and $O(n^2/p)$ computations on each processor. Another one with the best time efficiency needs only $O(n^2/p)$ time steps. All local computations on processors are based on the Knuth sequential algorithm for the OBST problem.

Keywords Parallel processing · Coarse Grain Multicomputer · Dynamic programming · Bulk synchronous parallel

1 Introduction

In this paper we tackle the problem of parallelizing OBST algorithm on the Bridging Coarse Grain BSP/CGM (Bulk Synchronous Parallel model/Coarse Grain Multicomputer) [5, 6, 19]. CGM seems the best suited for the design of algorithms that are not too dependent on an individual architecture. A BSP/CGM machine is a set of p processors. Each having its own local memory of size *m (O(m)>>O1)* and connected to a router able to deliver messages in point-to-point fashion.

V.K. Tchendji
Department of Mathematics and Computer Science,
University of Dschang, Dschang, Cameroon
e-mail: vianneykengne@yahoo.fr

J.F. Myoupo(✉) · G. Dequen
MIS Laboratory, University of Picardie Jules Verne, Amiens, France
e-mail: {jean-frederic.myoupo,gilles.dequen}@u-picardie.fr

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_57

A BSP/CGM algorithm consists in alternating local computations and global communication rounds. Each communication round consists in routing a single *h-relation* with $h=O(m)$. A CGM computation/communication round corresponds to a BSP super step with communication cost [4]. g is the cost of the communication of a word in the BSP model. To produce an efficient BSP/CGM algorithm designers effort tend to maximize speedup and minimize the number of communication rounds (ideally independent from the problem size, and, constant in the optimum).

There are some CGM-based parallel algorithms for some problems modeled by dynamic programming among others: The longest common subsequence problem [3, 4, 8], the string editing problem [1]. The special case of MCOP (Matrix Chain Ordering Problem), OBST and OSPP (Optimal String Parenthesizing Problem) problems which are modeled by the same dynamic programming algorithm, each with its own specification, was tackled in [7, 12,17].

The CGM-based parallel algorithm for MCOP in [13] requires $O(p)$ communication rounds and, at most, $O(n^3/p)$ time steps on p processors. The one for OSPP in [17] runs in $O(n^3/p)$ time steps with $\lceil (2p)^{1/2} \rceil$ communication rounds. The former CGM algorithm for OBST [12] needs $O(n^2/p)$ time steps with p communication rounds. A serious drawback of these algorithms is that the loads of processors are unbalanced and for two of them, a processor can detain up to p blocks in the worst case.

In this paper we propose a methodology to derive CGM parallel algorithms for the cost of the OBST Problem. Depending on the parameter we want to optimize, we derive an algorithm accordingly. Therefore a load balancing, an efficient and a minimum communication rounds algorithm is obtained. The best one [16] requires only $O(p^{1/2})$ communication rounds and $O(n^2/p)$ computations time. In our CGM algorithm, local computations on processors are carried out using the Knuth sequential algorithm for the OBST problem.

The remainder of this paper is organized as follow: In the next section we give the definition of the OBST problem and a short review of its sequential algorithms. Section 3 will be devoted to the presentation of our OBST BSP/CGM algorithms, and simulation results obtained. The last section concludes our work.

2 The Optimal Binary Search-Tree Problem

2.1 Classical Sequential Algorithm

Denote by *obst(i,j)* an optimal binary tree whose keys correspond to the interval $int(i,j)=[k_i, \ldots, k_j]$ and denote by *cost(i,j)* the cost of such a tree. Let $w(i,j)= p_i+\ldots + p_j + q_{i-1} + q_i+\ldots+q_j$. Therefore $w(1,n)=1$ and $w(i+1,i)=q_i$.

The costs obey to the following dynamic programming recurrences for $1\le i\le j\le n$ (equation 1) :

- $\cos t(i, j) = w(i, j) + \min\{\cos t(i, k-1) + \cos t(k+1, j) : i \le k \le j\}$.
- $\cos t(i+1, i) = q_i$

The corresponding sequential generic algorithm requires $O(n^3)$ time and $O(n^2)$ space. The values *cost(i,j)* are calculated by their tabulation in an array, often called dynamic programming table (or tasks graph) denoted here by *cost(1,n).* To each sub-problem *(i,j)* corresponds an entry in the dynamic programming table. This entry contains the value *cost(i,j)* at the end of the algorithm. For example, having seven keys *(n=7),* the corresponding dynamic programming table *cost(1,7)* (where *c(i,j)* stands for *cost(i,j)*) is depicted in figure 1.

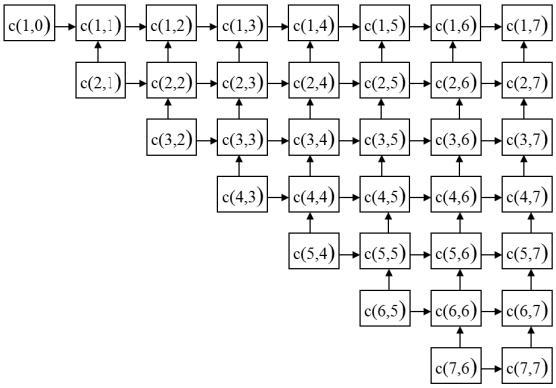


Fig. 1 Dynamic programming table for n = 7

The general structure of this algorithm is done by algorithm 1:

Algorithm 1. Classical algorithm for OBST

Data: The *2n+1* probabilities
Result: The cost of the best binary search tree
1 In diagonal0, Initialize each element (I+1, 1) to q_i and all other elements to 0.
2 **for** each diagonal d from 1 to n **do**
3 **for** each element *cost(i,j)* in diagonal **do**
4 *cost(i,j)*←*w(i,j)+min{cost(i, k+1)+cost(k+1, j): i≤k≤j};*

Once all elements of table *cost(i,n)* are calculated (called first step or step of calculation or search of the optimal solution), a recursive algorithm in $O(n^2)$ is executed for the construction of the tree (called second step or step of construction of the optimal solution). It uses a table, that we denote by *cut(1,n)*, which is obtained from the table *cost(1,n). cut(i,j)* gives the value of *k* (in equation 1) that minimizes *cost(i,j).* It is the first step that gives an optimal decomposition of *obst(i,j)* in two sub-trees. In this paper we are only interested by the parallelization of the first step of this problem, i.e., the search of the optimal cost.

2.2 Stata of Art

The sequential nature of the computation (diagonal-by-diagonal processing) of the dynamic programming table used in the resolution of the OBST problem prevents from its efficient parallelization in the case of the PRAM model [2]. In the case of real target parallel machine (systolic model for example), where one takes into consideration the communication operations, the parallelization of this algorithm becomes more complicated, because of the strong dependence between sub-problems (the calculation of an element of the diagonal k requires two elements of diagonals *0, 1, ..., k-1*, see [9, 10]. The solution becomes, again, more difficult, if we want also to search the OBST solutions on a Bridging model as the BSP model [18, 19]. Especially if the goal is to benefit from the characteristic of the majority of current standard parallel machines (where local memory size >> *1*) in order to minimize the communication overhead by regrouping them in global communication rounds between pure calculation super steps, which is the objective of the CGM approach.

The fastest sequential algorithm is due to Knuth [14]. It requires $O(n^2)$ time and $O(n^2)$ space. This algorithm has the same principle than the classical one. This principle is the so-called property of monotonicity verified by some elements of the matrix *cut(1,n)* that permits this acceleration [14].

2.3 Knuth-Acceleration-Based Sequential Algorithm in $O(n^2)$

Knuth acceleration results from this observation in the property of the matrix cut [14]:

$$i \le i' \le j \le j' \Rightarrow cut(i, j) \le cut(i', j')$$

It transforms the algorithm 1 (in section 2.1 viewed supra) to the following (algorithm 2):

Algorithm 2. Knuth algorithm for OBST

```
Data: The 2n+1 probabilities
Result: The cost of the best binary search tree
1 In diagonal0, Initialize each element (I+1, 1) to q_i and
all other elements to 0.
2 for each diagonal d from 1 to n do
3   for each element cost(i,j) in diagonal do
4     cost(i,j)←w(i,j)+min{cost(i, k+1)+cost(k+1, j):
      cut(i,j-1)≤k≤cut(i+1,j)};
5     cut(i,j) ← the value of k that minimizes cost(i, j)
```

It is clear that the acceleration comes from the internal for loop which requires only $O(n)$ comparison operations instead of $O(n^2)$ in the classical algorithm. For the calculation of an element of the diagonal $(j-i+1)$, say $cost(i,j)$, the number of comparison operations is reduced from $(j-i)$ to $(cut(i+1,j)-cut(i,j-1))$. This acceleration is not regular. In the same diagonal, the gain (the reduction) of the number of comparison operations defer from an element to another. Thus, contrary to the version without acceleration, if elements of a diagonal are distributed fairly on the different processors, nothing guarantees that these processors will have the same load. Therefore, a new constraint must be managed in the conception of our CGM algorithm. It is "the inequality of necessary calculation loads for elements of the same diagonal of the dynamic programming table". This property of the Knuth algorithm provokes an unmanageable unbalanced load that can make vanished the acceleration quoted supra. The only solution to correctly obtain from this acceleration is to calculate each diagonal by a unique processor. Whereas such choice helps to exploit the Knuth acceleration calculation and eliminates the global unbalanced loads, the parallelization process can penalize (to reduce) the exploitation of the calculation independence. The main source of parallelism of the Knuth algorithm, as for the classic $O(n^3)$ version, comes from the existing independence between the calculation of elements of each diagonal of the dynamic programming table.

3 A Methodology to Derive a Set of CGM Algorithms

In this section, we propose an original methodology to derive CGM parallel algorithms for the cost of the OBST problem. Depending on the parameter we want to optimize, we derive an algorithm accordingly. Therefore a load balancing, an efficiency and a minimization of communication rounds are obtained. Since the resolution of the problem reduces to calculating the cost corresponding to each node in task graph, the subdivision of work to be done also reduces to partitioning of the graph. In [16, 17], it is shown that for this problem, it is difficult to optimize simultaneously all these performance criteria in the same algorithm. It is shown that load balancing (and thus efficiency) and minimization of communications are two conflicting objectives for partitioning the tasks graph.

We want to give to the final user, the choice to optimize one criterion or another. He can also choose an intermediate solution. Our idea is based on a graph partitioning technique and a mapping of data onto the processors already defined. The steps of the algorithm based on user parameters are then deduced. The performance (complexity) of the algorithm is expressed from these parameters. These parameters are:

- g (grain) : number of rows (or columns) after the desired graph partitioning (we deduce from this partitioning the block size of the graph);
- p : the number of processors.

We will apply our idea to the tasks graph partitioning technique and data distribution introduced in [11] and [16]. In these approaches, the tasks graph is

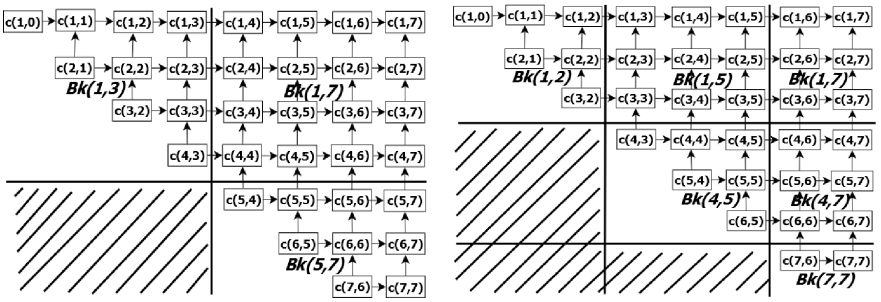
partitioned in the same way (in rows and columns of blocks) and data dependencies are the same. These approaches differ in the data distribution technics used on the processors.

The next sub-sections presents data structures that are involved in the partition of the dynamic programming table in sub-tables, in order to map them on a coarse grained multicomputer of p processors.

3.1 Tasks Graph Partitioning Technique

We partition the upper triangular matrix (or tasks graph) into $g(g+1)/2$ sub-matrices (sub-graphs) or blocks (*Bk(i,j)* for short). *Bk(i,j)* being a $\lceil (n+1)/g \rceil \times \lceil (n+1)/g \rceil$ matrix, except those at the extremities (as shown in figures 2a and 2b). Hereafter, each sub-matrix *Bk(i,j)* will be called a block. And to ease the readability of the text, we will frequently use the following: $f(g,n) = \lceil (n+1)/g \rceil$.

Each block *Bk(i,j)* $(1 \leq j \leq n)$ is a dynamic programming table of *c(i,j),* it contains the temporary values of the cost corresponding to node *(i,j).*



(a) For $g=2$ we have 3 blocks (b) For $g=3$ we have 6 blocks

Fig. 2 Task graph portioned according to the parameter g

Every block *Bk(i,j), for* $i < j$, is a sub-problem of (1). A block (a sub-problem) will be assigned to a processor responsible for sequentially solving the associated OBST problem.

3.2 CGM Algorithms

The communication scheme derived from the mapping of figure widely discussed in [16] is easily implemented. Let's denoted by $P_{i,j}$ the processor which evaluates the block *Bk(i,j)*. To evaluate its block, processor $P_{i,j}$ waits for a block from each of the processors $P_{i,l}$ $(1<l<j)$ and $P_{d,j}$, $(i<d<n)$. After the evaluation of the block *Bk(i,j)*, processor $P_{i,j}$ sends it to processors $P_{i,r}$, $(j<r \leq n$ and $P_{u,j}$, $(1 \leq u<i)$. From there, the execution time of the algorithm can be predicted.

A step of our algorithm is the evaluation of at most p consecutive blocks all independent. When there is at must p blocks in each diagonal, a step of the algorithm consists of evaluating one diagonal of blocks. But when there are more, a diagonal block is evaluated in several steps (rounds) .

Lemma 1
The number of steps of our algorithm is defined as follows:

$$r = \begin{cases} g, if g \le p \\ \left\lfloor \frac{g(g+1) - p(p-1)}{2p} \right\rfloor + (p-1) + \theta, if g > p \end{cases}$$

$$where \theta = \left\lceil \frac{(\frac{g(g+1) - p(p-1)}{2}) \bmod p}{p} \right\rceil$$

Proof: It is straightforward

Our distribution shows that at step k, at most p blocks are exchanged between the processors. A processor executes algorithm 2 for calculation phases. Thus algorithm 3 is the overall CGM algorithm that is run by every processor to solve the OBST problem.

Algorithm 3. CGM algorithm

Data: The initial cost(1,n) and cut 1, n) tables and the list of *2n+1* probabilities
Result: The tables cost(1,n) and cut(1, n) calculated
1 In diagonal0, Initialize each element (I+1, 1) to q_i and all other elements to 0.
for d=1 to R 1 **do**
1 Compute local solutions of the blocks of round d using Algorithm 2.
2 Communicate the entries (of cost and cut tables) necessary for calculation of each block of rounds {d+1, d+2, …, R}

We can now express the performances of our algorithms.

Theorem 1
The CGM algorithm runs in the worst case in $O((n^2/g^2) \times g)$ time steps with R rounds of communications.
Proof: It is straightforward

3.3 Simulation Results

In this section, we present the evolution of the performance of the algorithms that we derive based on parameters p and g. To do this, the values of p are taken from the set {2,4,10, 20,30, . . . ,100} and those of g in the set {2,4,6,8, . . . , 1000}. Figure 3 shows that for the same grain of calculation, the number of rounds of communication which is a factor which reflects the communication time, decreases as the number of processors increases.

Figure 4 shows that this family of algorithms have better efficiency when few number of processors are used. Similarly, with a large number of processors, it also obtains a good efficiency when the grain is large (i.e., when the blocks are small). Note however that this efficiency is calculated based only on the amount of calculations to be performed. However, if one also takes into account the amount of data communicated (i.e. the communication time), this efficiency will keep the same progression but will be lower than the one observed in figure 4.

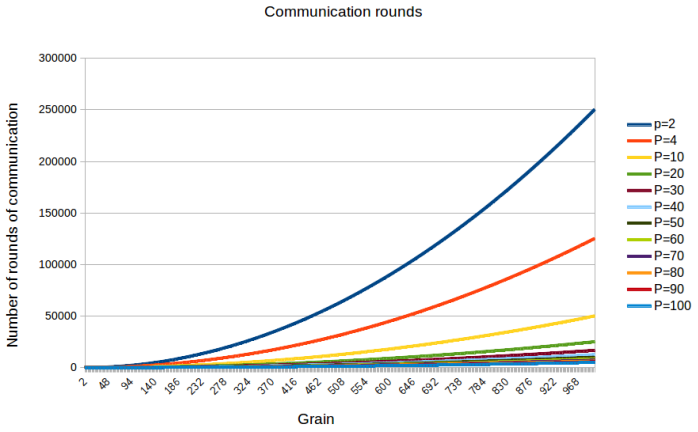


Fig. 3 Number of communication rounds from parameters p and g.

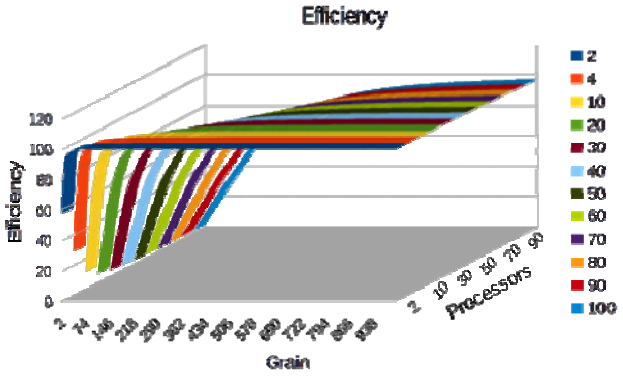


Fig. 4 Efficiency from parameters p and g.

4 Conclusion

In this paper we have proposed an original methodology to derive CGM parallel algorithms for the cost of the OBST problem. Depending on the parameter we wanted to optimize, we derived an algorithm accordingly. Therefore a load balancing, an efficient and a minimum communication rounds algorithm is obtained. We applied our idea to the tasks graph partitioning technique and data distributions introduced in [11] and [16], and the CGM algorithms obtained use p processors, each with $O(n^2/p)$ local memory. They run in the worst case in $O(n^2/g^2)\times R$ time steps with R rounds of communications.

Our idea can be applied to other graphs partitioning techniques and data distribution. For example, the one introduced in [15] where each diagonal block is evaluated by a single processor. It can also be applied to techniques for solving other problems, for example the one introduced in [12] to solve the Matrix Chain Ordering Problem using the sequential Yao's algorithm [20].

References

1. Alves, C.E.R., Cáceres, E.N., Dehne, F.: Parallel dynamic programming for solving the string editing problem on a CGM/BSP. In: SPAA 2002, pp. 275–281. ACM, New York (2002)
2. Alves, C.E.R., Cáceres, E.N., Dehne, F., Song, S.W.: A parallel wavefront algorithm for efficient biological sequence comparison. In: ICCSA 2003, pp. 249–258 (2003)
3. Alves, C.E.R., Caceres, E.N., Song, S.W.: A coarse-grained parallel algorithm for the all-substrings longest common subsequence problem. Algorithmica **45**, 301–335 (2006)
4. Bradford, P.G.: Parallel dynamic programming. Ph.D. thesis, Indiana University (1994)
5. Cheatham, T., Fahmy, A.F., Stefanescu, D.C., Valiant, L.G.: Bulk synchronous parallel computing-a paradigm for transportable software. In: HICSS, vol. 2, pp. 268–275 (1995)
6. Dehne, F., Fabri, A., Rau-Chaplin, A.: Scalable parallel computational geometry for coarse grained multicomputers. International Journal of Computational Geometry and Applications **6**(3), 379–400 (1996)
7. Fotso, L.P., Tchendji, V.K., Myoupo, J.F.: Load balancing schemes for parallel processing of dynamic programming on BSP/CGM model. In: PDPTA, pp. 710–716 (2010)
8. Garcia, T., Myoupo, J.F., Seme, D.: A coarse-grained multicomputer algorithm for the longest common subsequence problem. In: PDP, pp. 349–356 (2003)
9. Guibas, L.J., Kung, H.T., Thompson, C.D.: Direct VLSI implementation of combinatorial algorithms. In: Proceedings of Conference on Very Large Scale Integration, California Institute of Technology, pp. 509–525 (1979)
10. Karypis, G., Kumar, V.: Efficient parallel mappings of a dynamic programming algorithm: a summary of results. In: IPPS, pp. 563–568 (1993)

11. Kechid, M., Myoupo, J.F., Khalufa, A.S., Alghamdi, M.S.: Mapping dynamic programming problems on coarse grained multicomputer. In: ICIA2012, pp. 487–499. The Society of Digital Information and Wireless Communication (2012)
12. Kechid, M., Myoupo, J.F.: A coarse grain multicomputer algorithm solving the optimal binary search tree problem. In: Proceedings of the Fifth International Conference on Information Technology: New Generations, pp. 1186–1189. IEEE Computer Society, Washington, DC (2008)
13. Kechid, M., Myoupo, J.F.: An efficient BSP/CGM algorithm for the matrix chain ordering problem. In: PDPTA, Las Vegas, pp. 327–332, July 2008
14. Knuth, D.E.: Optimum binary search trees. Acta Informatica **1**(3), 14–25 (1972)
15. Myoupo, J.F., Tchendji, V.K.: An efficient cgm-based parallel algorithm solving the matrix chain ordering problem. International Journal of Grid and High Performance Computing (IJGHPC) **6**(2), 74–100 (2014)
16. Myoupo, J.F., Tchendji, V.K.: Parallel dynamic programming for solving the optimal search binary tree problem on CGM. International Journal of High Performance Computing and Networking **7**(4), 269–280 (2014)
17. Tchendji, V.K., Myoupo, J.F.: An efficient coarse-grain multicomputer algorithm for the minimum cost parenthesizing problem. The Journal of Supercomputing **61**, 463–480 (2012)
18. Valiant, L.G.: Bulk-synchronous parallel computers. In: Parallel Processing and Artificial Intelligence, pp. 15–22 (1989)
19. Valiant, L.G.: A bridging model for parallel computation. Communications of the ACM **33**, 103–111 (1990)
20. Yao, F.F.: Speed-up in dynamic programming. SIAM Journal on Matrix Analysis and Applications (formerly SIAM Journal on Algebraic and Discrete Methods) **3**(4), 532–540 (1982)

Experimental Evaluations of MapReduce in Biomedical Text Mining

Yanqing Ji, Yun Tian, Fangyang Shen and John Tran

Abstract In this paper, we demonstrate our development of two biomedical text mining applications: biomedical literature search (BLS) and biomedical association mining (BAM). While the former requires less computations, the latter is more computationally intensive. Experimental studies were conducted using Amazon Elastic MapReduce (EMR) with an input of 33,960 biomedical articles from TREC (Text REtrieval Conference) 2006 Genomics Track. Our experiment results indicated that both applications' scalabilities were not linear in term of the number of computing nodes. Meanwhile, BAM achieved better scalability than BLS since BLS performed less computations and were primarily dominated by overheads such as JVM startup, scheduling, disk I/O, etc. These observations imply that existing MapReduce framework may not be suitable for on-line systems such as literature search that needs quick response.

Keywords MapReduce · High-performance computing · Big data · Text mining · Biomedical literature

Y. Ji(✉)
Gonzaga University, Spokane, WA 99202, USA
e-mail: ji@gonzaga.edu

Y. Tian
Eastern Washington University, Cheney, WA 99004, USA
e-mail: ytian@ewu.edu

F. Shen
NY City College of Tech, Brooklyn, NY 11201, USA
e-mail: fshen@citytech.cuny.edu

J. Tran
Frontier Behavioral Health, Spokane, WA 99202, USA
e-mail: jtran@smhca.org

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_58

1 Introduction

MapReduce represents a parallel and distributed programming paradigm for large-scale data analysis on clusters of commodity servers. Using a divide-and-conquer approach, MapReduce solves big-data problems in two phases: map and reduce. In the map phase, the input data is split into different partitions that are processed in parallel by different map tasks. In the reduce phase, results of each partition are combined to get the final results or the input for another map-reduce job. With this simple programming model, programmers can concentrate on high-level data processing strategies, while the MapReduce framework takes care of all details of parallel execution.

MapReduce has been employed to process large volume of data in different domains because of its simplicity and fault-tolerance [1][2]. However, its performance and uses in biomedical text mining have not been extensively studied. In this study, we developed two biomedical text mining applications: biomedical literature search (BLS) and biomedical association mining (BAM). Both applications had two MapReduce jobs and utilized the same input text data. These two applications were used to study the advantages and limitations of MapReduce in the context of biomedical text mining.

The input of both applications is a large volume of full-text biomedical articles published by biomedical researchers and healthcare providers. These articles represent an important source of information that not only enables researchers to discover in-depth knowledge about various biological systems, but also helps medical practitioners do evidence-based medicine in real clinical settings [3]. They have attracted a lot of researchers in the field of biomedical information retrieval and text mining [5]. In this study, these articles were preprocessed by matching them into standard biomedical concepts defined in the Unified Medical Language System (UMLS) Metathesaurus. Each concept has a unique concept identification (ID).

Our first application, named as BLS, was developed to help biomedical researchers and healthcare providers more efficiently find relevant articles from a biomedical database. At present, the most influential biomedical database is PubMed [6] which is maintained by the National Library of Medicine. PubMed contains more than 24 million citations and approximately 10,000 citations are added to the database every week. PubMed supports keyword and constraint queries. However, a keyword query generally returns a long list of hits. For example, the keyword “lung cancer” retrieves more than 250,000 articles. Adding a couple of constraints could narrow down the results but the returned list is still likely too long for users to review each hit. Furthermore, the quality of the query results is poor when users only vaguely know their information needs and cannot provide precise keywords. Another limitation of the current PubMed search algorithm is that it does not support complex query search. For example, if a question “How do interactions between insulin-like GFs and the insulin receptor affect skin biology?” is searched in PubMed, nothing is returned. Using UMLS Metathesaurus, our method can more effectively retrieve highly relevant articles and support queries with any levels of complexity. Each query can include any number of keywords, questions, or sentences.

Our second application, referred as BAM, was designed to efficiently extract associations between biomedical concepts from large text data. Association mining aims at discovering association rules in the form of $X \rightarrow Y$, where X and Y are two disjoined itemsets, i.e., $X \cap Y = \emptyset$. Since the introduction of the first association mining algorithm in 1994 [7], this technique has been used for knowledge discovery in various areas including biomedicine. For example, Shetty et al. studied whether applying association mining to PubMed could discover important drug-adverse event associations [8]. Xu et al. demonstrated that drug-adverse effect pairs extracted from biomedical literature using various approaches including association mining could significantly improve post-marketing drug safety surveillance [9]. The strength of an association rule is normally assessed by various interestingness measures. In this study, we developed a MapReduce algorithm that could be used to compute a category of interestingness measures defined on the basis of a 2×2 contingency table.

A comparative study of the above two applications was conducted using 33,960 biomedical articles from TREC 2006 Genomics Track. More specifically, we explored the scalabilities of both applications in terms of the number of computing nodes under Amazon?s EMR platform. Our experiment results indicated that both algorithms were not linearly scalable in terms of the number of nodes, even though BAM exhibited better scalability than BLS. This implies that MapReduce may not be suitable for on-line biomedical systems that requires quick responses since it is difficult to reduce the execution time to a user-acceptable level (e.g., a couple of seconds) by simply increasing the number computing nodes.

2 Mapreduce Algorithm Design

In this section, we present our MapReduce algorithms for both applications: BLS and BAM. In the MapReduce programming paradigm, two methods must be defined: Map() and Reduce(). The Map() method reads a list of (key1, value1) pairs from input data and computes an arbitrary number of intermediate key-value pairs, (key2, value2). In this study, input data is a set of documents or articles. Key1 denotes a document id and value1 is the content of the document. The intermediate pairs are grouped together according to the key and (key2, list(value2)) is generated by the MapReduce framework. After all map tasks are completed, the Reduce() method is then applied to all values (i.e., list(value2)) associated with the same intermediate key in order to generate output key-value pairs. In both phases, the framework schedules map tasks and reduce tasks to different workers so that they are executed in parallel.

2.1 MapReduce Algorithm for BLS

In this study, we define accumulative term frequency-inverse document frequency (A-TF-IDF) to rank the articles for users? queries. TF-IDF is an established weighting scheme in information retrieval and text mining [10]. It upweights a term by its

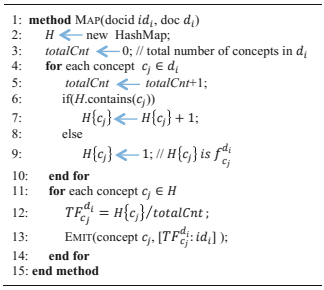


Fig. 1 Map() Method of the First MapReduce Job for BLS.

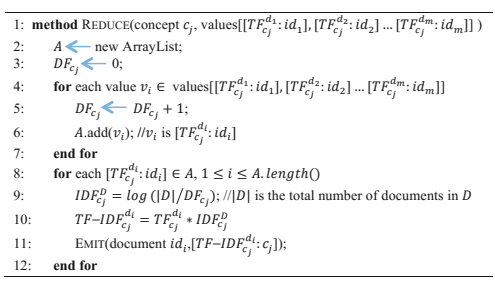


Fig. 2 Reduce() Method of the First MapReduce Job for BLS.

frequency in the document and downweights it by the logarithm of how common it is in a collection of documents. It essentially makes the TF-IDF value higher for a term that has high frequency in a document but is less likely contained by the other documents in a collection. In the context of this study, a term is actually a concept ID and a document refers to a biomedical article. To be consistent with the notation of TF-IDF, we use the term "document" to represent an article in the following discussions.

Let $D = \{d_1, d_2, \ldots, d_i, \ldots, d_m\}$ be a set of documents in a database. Let $C = \{c_1, c_2, \ldots, c_j, \ldots, c_n\}$ be a set of unique biomedical concept IDs contained by a document d_i. *Term frequency (TF)* measures how often a term appears in a document. Terms that appear in a document more often are more likely to be important within the document. The term frequency for a concept c_j in a document d_i is defined as the frequency of c_j in d_i divided by the total number of concepts in d_i. That is,

$$TF_{c_j}^{d_i} = \frac{f_{c_j}^{d_i}}{\sum_{j=1}^{n} f_{c_j}^{d_i}} \tag{1}$$

where $f_{c_j}^{d_i}$ is the frequency of a concept c_j in a document d_i. The *inverse document frequency (IDF)* examines the general importance of a term in a set of documents D. It is defined as,

$$IDF_{c_j}^{D} = log \frac{|D|}{DF_{c_j}} \tag{2}$$

where $|D|$ represents the total number of documents in D, and DF_{c_j} is the total number documents that contain the concept c_j at least once. The TF-IDF weight of a concept c_j in d_i is defined as its TF multiplied by IDF. That is,

$$TF\text{-}IDF_{c_j}^{d_i} = TF_{c_j}^{d_i} * IDF_{c_j}^{D} = \frac{f_{c_j}^{d_i}}{\sum_{j=1}^{n} f_{c_j}^{d_i}} * log \frac{|D|}{DF_{c_j}} \tag{3}$$

```
1: method MAP( docid id_i, doc d_i)
2:     A ← new ArrayList;
3:     for each concept c_j ∈ d_i
4:         if(A.contains(c_j) = false)
5:             A.add(c_j);
6:     end for
7:     SORT(A);
8:     for each concept c_j ∈ A, 1 ≤ j ≤ A.length()
9:         H ← new HashMap;
10:        H{c_j} ← 1;
11:        If(j ≠ A.length())
12:            for each concept c_k ∈ A, j < k ≤ A.length()
13:                H{c_j: c_k} ← 1;
14:            end for
15:        end if
16:        EMIT(concept c_j, hashmap H);
17:    end for
18: end method
```

Fig. 3 Map() Method of the First MapReduce Job for BAM.

```
1: method REDUCE(concept c_j, hashmaps[H_1, H_2 ... H_m] )
2:     H ← new HashMap;
3:     for each hashmap H_i ∈ hashmaps[H_1, H_2 ... H_m]
4:         for each key k ∈ H_i
5:             if(H.contains(k))
6:                 H{k} ← H{k} + H_i{k};
7:             else
8:                 H.add(k, H_i{k});
9:         end for
10:    end for
11:    EMIT(concept c_j, hashmap H);
12: end method
```

Fig. 4 Reduce() Method of the First MapReduce Job for BAM.

Let $Q = \{c_1, c_2, \ldots, c_k, \ldots, c_l\}$ be a query that contains l concepts. We use a metric accumulative TF-IDF to rank all the documents in D. That is, for each document, we first compute the TF-IDF weight of each concept in Q and then sum up these weights. We define accumulative TF-IDF, named A-TF-IDF, for a document d_i relative to a query Q as below:

$$A\text{-}TF\text{-}IDF_Q^{d_i} = \sum_{k=1}^{l} TF\text{-}IDF_{c_k}^{d_i} = \sum_{k=1}^{l} \left(\frac{f_{c_k}^{d_i}}{\sum_{j=1}^{n} f_{c_j}^{d_i}} * log \frac{|D|}{DF_{c_k}} \right) \tag{4}$$

After the A-TF-IDF is computed for each document, the documents are ranked according to their A-TF-IDF values. The document with a higher A-TF-IDF value will be ranked higher. The ranked list is then returned to users.

We developed two MapReduce jobs to in order to compute the A-TF-IDF value for each document given a query. Figure 1 gives the Map() method of the first MapReduce job for BLS. The input of this method is a document d_i. The MapReduce framework guarantees that each document is processed by one mapper, a worker that is assigned a map task. The first loop (line 4-10) reads the document d_i and computes the total number of concepts in d_i as well as the count for each concept. Each concept and its count are stored in a hashmap structure where the concept is the hash key and its count is the mapped hash value. We use $H\{c_j\}$ represents the mapped value for concept c_j. Therefore, $H\{c_j\}$ is actually the frequency of c_j in d_i (i.e., $f_{c_j}^{d_i}$). The next loop gets the frequency of each concept c_j and calculates the term frequency $TF_{c_j}^{d_i}$. The method then emits each concept c_j and its $TF_{c_j}^{d_i}$ attached with the document identification id_i.

After the above map method, all $TF_{c_j}^{d_i}$ values (and their related document IDs) associated with the same concept will be brought to the same reduce task by the MapReduce framework. Figure 2 provides the Reduce() method of the first MapReduce job for BLS. The first loop (line 4-7) counts the total number of documents that contains a particular concept (i.e., DF_{c_j}). It also adds each value (i.e., $[TF_{c_j}^{d_i} : id_i]$) to an

Table 1 A 2 × 2 Contingency Table for X and Y

	X	$\overline{X}$	
Y	f_{XY}	$f_{\bar{X}Y}$	f_Y
$\overline{Y}$	$f_{X\bar{Y}}$	$f_{\bar{X}\bar{Y}}$	$f_{\bar{Y}}$
	f_X	$f_{\bar{X}}$	N

Table 2 Statistics of HTML Files, Text Files, and MetaMap Files

	Max size	Min size	Ave size	Total size
Html files	780 KB	0.86 KB	780 KB	2.78 GB
Text files	325 KB	0.29 KB	38.6 KB	1.31 GB
MetaMap files	11.2 MB	3.16 KB	465 KB	15.8 GB

array. The next loop computes $IDF_{c_j}^D$ and $TF\text{-}IDF_{c_j}^{d_i}$. The method then outputs each document ID and its $TF\text{-}IDF_{c_j}^{d_i}$ attached with the concept c_j.

The Map() method of the second MapReduce job simply reads the output of the first job and emits it to the Reduce() method of the second job. That is, the second reduce method is applied to all the $TF\text{-}IDF_{c_j}^{d_i}$ values (and their concepts) associated with the same document ID. This method then computes the $A\text{-}TF\text{-}IDF_Q^{d_i}$ related to a particular query for each document. As the second job is straightforward, its Map and Reduce methods are not provided.

2.2 MapReduce Algorithm for BAM

To evaluate the strength of an association rule, researchers have developed various interestingness measures such as confidence [11], interest(I) [12], and so forth. Many of these measures are computed based on the frequency counts tabulated in a 2 × 2 contingency table. In this study, we focus on this category of interestingness measures whose computations depend on the 2 × 2 contingency table. Note that, even though both X and Y may contain one or more concepts, we only consider the associations between two single concepts in this study primarily because it is often difficult to biologically or clinically interpret the mining results when X and Y contain two or more concepts.

Table 1 shows an example of a 2 × 2 contingency table for X and Y. The notation $\bar{X}(\bar{Y})$ is used to indicate that $X(Y)$ is absent from a document. Each entry f in the 2×2 table denotes a frequency count. For instance, f_{XY} represents the number of times that X and Y occur in the same document, while $f_{X\bar{Y}}$ is the number of documents that contain X but not Y. In this table, $f_X = f_{XY} + f_{X\bar{Y}}$ and $f_Y = f_{XY} + f_{\bar{X}Y}$, which represent the number of documents that contain X and Y, respectively. N is the total number of documents in a dataset D. The I measure, for instance, is defined as $Nf_{XY}/f_X f_Y$. Therefore, our goal is to develop an efficient parallel algorithm using MapReduce in order to compute f_{XY}, $f_{X\bar{Y}}$, $f_{\bar{X}Y}$ and $f_{\bar{X}\bar{Y}}$. Figure 3 shows the pseudo-code of our Map() method of the first MapReduce job for BAM. The first loop (line 3-6) in the Map() reads each concept in a document and creates a list of unique concepts. Line 7 lexicographically sorts this list. This step makes sure that, when two concepts are paired together in the following nested loop, the concept with lower

lexicographical order will always precede the concepts with higher order. This avoids the possibility of generating two different pairs with the same two concepts. The next nested loop (Line 8-17) generates each concept and its associated pairs (i.e., c_j:c_k), then emits the results. We use a *one occurrence per document* counting scheme. That is, the count for each concept and each of its associated pairs in a document is one.

We use a "stripes" approach. Each concept and its associated pairs are stored in a hashmap H, and the whole H is considered as a value emitted together with the concept, which is the key. An alternative approach is to emit each concept and each co-occurrence concept pair. Obviously, our approach generates much fewer intermediate key-value pairs. For example, if a document contains m unique concepts, our approach generates $O(m)$ number of pairs, while the other approach produces $O(m^2)$ pairs. Since the intermediate outputs produced by the Map() method are first sorted locally in order for grouping key-value pairs sharing the same key, in our approach the MapReduce execution framework performs less sorting and thus be more efficient.

The Reduce() method of the first job for BAM is shown in Figure 4. This method performs an element-wise sum of all hashmaps associated with the same key. The MapReduce framework makes sure that all hashmaps with the same key will be brought to the same reducer for processing. The frequency counts associated with the same concept or pair in different hashmaps are accumulated and the results are stored in a new hashmap in each reducer. The final hashmap is emitted using the same key.

After the above MapReduce job is completed, we still do not have enough information to compute the interestingness measures defined on the 2×2 contingency table. For each pair in the final hashmap generated by the Reduce() method, its frequency count is known. In addition, the count for the first concept of the pair can also be obtained from the same hashmap. That is, given a concept pair (X, Y), f_{XY} and f_X are known after the MapReduce job. The next step is to efficiently compute f_Y given the output of the first MapReduce job. Once f_Y is ready, the following equations can be used to find $f_{X\bar{Y}}$, $f_{\bar{X}Y}$ and $f_{\bar{X}\bar{Y}}$:

$$f_{X\bar{Y}} = f_X - f_{XY} \tag{5}$$

$$f_{\bar{X}Y} = f_Y - f_{XY} \tag{6}$$

$$f_{\bar{X}\bar{Y}} = N - f_{XY} - f_{X\bar{Y}} - f_{\bar{X}Y} \tag{7}$$

where N is the total number of documents in the dataset.

To compute f_Y, we designed the second MapReduce job. The input of the Map() method of the second job is the final hashmap emitted by the Reduce() method from the first job. The Map() method reads the contents of the hashmap and first emits the single concept and its count from the hashmap. For instance, $\{X, [X : f_X]\}$ is emitted for concept X. That is, if $X = A$ and $f_X = 2$, then$\{A, [A : 2]\}$ is emitted. Concept X is also the first concept of each pair in the hashmap. The method then appends the count of the single concept to each pair entry in the hashmap. After that,

each new pair entry is emitted using the second concept of the pair as the key. That is, for an entry $(XY\ f_{XY})$, $\{Y, [(XY : f_{XY} : f_X)]\}$ is emitted.

In the reduce phase of the second job, each reducer receives f_Y and the information of all pairs whose second concept is Y since they have the same key. As f_X is already attached to a concept pair (X, Y), the reducer has all the three values f_{XY}, f_X, and f_Y. With these values, $f_{X\bar{Y}}$, $f_{\bar{X}Y}$ and $f_{\bar{X}\bar{Y}}$ can be computed using Eq. (5)-(7). The reducer now has enough information to compute the value of any interestingness measure based on the 2×2 contingency table. As the pseudo codes for the Map() and Reduce() methods of the second job are straightforward, they are not provided in the text.

3 Experiments

3.1 Experiment Data and Environment

To test our two MapReduce algorithms, we obtained 33,960 articles from TREC 2006 Genomics Track [13]. These data files were provided as html files and their total size was 2.78 GB. Before testing our biomedical text mining applications, biomedically meaningful concepts need to be extracted from these files. We achieved this using UMLS Metathesaurus, which is a large vocabulary and standard database that contains health and biomedical concepts. Specifically, we first converted the html files into text files by removing all html tags. Next, these text files were sent to UMLS servers one by one through Java-based APIs provided by UMLS. The UMLS servers are maintained by National Library of Medicine (NLM) and hold the Metathesaurus and a set of lexical tools. Once a text file is received by the server, the file is broken down into phrases, each of which is mapped to a standard concept in Metathesaurus by a lexical tool called MetaMap [14]. The server generates a MetaMap file containing each phrase and its matched concepts and returns it to users? local computer. All the returned MetaMap files are the input of our applications. Note that each phrase may be mapped to multiple concepts, each of which is associated with a score. The higher the score, the closer the phrase matches the concept. When the MetaMap files are processed by our applications, only the concept with the highest score is retrieved to represent an associated concept. Table 2 provides the statistics of the three types of files used in the test. The I measure, as described in section 2.2, was selected as a study case for BAM in the experiments.

Amazon offers a web-based service called Amazon Elastic MapReduce (EMR) that uses Hadoop to manage a cluster of virtual servers. With EMR and Amazon's cloud computing architecture, it is simple to launch a Hadoop cluster without worrying about details such as cluster setup, Hadoop configuration, etc. The experiments in this study were performed on Amazon EMR.

Table 3 Execution Time Of Both Applications Using Different Number of Nodes (Seconds)

Application		# nodes 10	20	30	40	50
BLS	Job 1	139	126	105	87	80
	Job 2	60	59	62	62	63
	Total	199	185	167	149	143
BAM	Job 1	678	395	291	216	178
	Job 2	188	120	103	92	87
	Total	866	515	394	308	265

Table 4 Data-Local Map Tasks and Total Map Tasks Created in The Map Phase of The First MapReduce Job of Both Applications

Application	# nodes	10	20	30	40	50
BLS	Average data-local map tasks	20.5	41.6	29.8	23.3	22.7
	Average total map tasks	72.1	152.3	233.7	312.4	393.2
	Data-local map task ratio	0.284	0.273	0.128	0.075	0.058
BAM	Average data-local map tasks	16.8	23.9	24.3	24.1	21.5
	Average total map tasks	72.3	152.1	233.2	312.5	393.4
	Data-local map task ratio	0.232	0.157	0.104	0.077	0.055

3.2 Experiment Results

We conducted our experiments using Amazon EMR that employs Hadoop to distribute users' data and computations across a resizable cluster of computing nodes. Each node is an Amazon Elastic Compute Cloud (EC2) instance, a virtual machine that has CPU, memory, storage and networking capacity. When creating an EMR cluster, users are allowed to choose EC2 instance types, the number of instances, Hadoop versions, etc. A moderate instance type (named as m3.xlarge by Amazon) was chosen in this study. This instance type has 4 virtual CPUs and 15 GB memory. The most recent Hadoop version (i.e., 2.6.0) supported by Amazon EMR was selected. The input and output data were stored in the Hadoop Distributed File System (HDFS) on the created cluster. HDFS represents the standard data storage layer for a Hadoop system. Note that the MapReduce framework employs a runtime scheduling scheme to executes its map and reduce tasks. Due to the dynamic nature of the framework, we run both applications 10 times and the average execution time is reported.

We examined the scalabilities of the two applications in terms of the number of nodes in the cluster used for experiments. Table 3 gives the execution time of both applications using different number of nodes. Since each application consists of two MapReduce jobs, the execution time for each job is also provided. One can observe that the first job consumes most of the execution time for both applications since it has to read the original input data and find all the concepts and their counts. To get a better view of the scalabilities, speedup rates for different number of nodes are calculated and presented in figure 5. The speedup rate for n nodes is defined as the total execution time for 10 nodes divided by the total execution time for n nodes.

We observe that the scalabilities of both applications are not linear. This implies that the computation power of all nodes is not fully utilized when the cluster includes more nodes. For example, when the number of nodes is increased from 10 to 20 (i.e, the number of nodes doubles), the speedup rate for BLS is only 1.08 while the rate for BAM is 1.72. In addition, BAM achieved a much better scalability than BLS, because BAM requires more computations that can be performed in parallel. In such a case, the computation power of the nodes are more efficiently utilized. As for BLS, since it performs less computations, its time cost is dominated by overheads that cannot be reduced by parallelism, such as JVM startup, scheduling, disk I/O, etc.

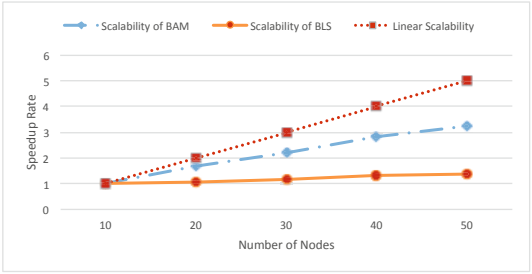


Fig. 5 Scalabilities of Our Two Applications in Terms of the Number of Nodes.

The non-linearity of both applications is attributed to the existing design of the Hadoop framework. Hadoop takes a master-slave model. The master node accepts jobs, assigns them to different slaves, and monitors the execution of each job, while slave nodes actually store data and run jobs. Each slave contains a DataNode and TaskTracker. While a DataNode is a HDFS component that actually stores smaller data blocks, a TaskTracker represents a server that accepts and executes map or reduce tasks. When the cluster contains more slaves, the input data will be distributed on more nodes and more map tasks will be generated to read and process the data. During execution, Hadoop attempts to schedule map tasks to slaves on which the data required by these tasks are stored. These tasks are called *data-local* tasks since the data are physically on the same slave node as the computation. On the other hand, Hadoop may have to schedule many tasks to different slaves in order to balance the overall loads and to maximize parallelism. Table 4 presents the data-local map tasks and the total map tasks in the map phase of the first MapReduce job of both applications. We observe that the ratio of data-local map tasks decreases when the number of nodes increases, meaning more non-local map tasks are created and incurs more network data accesses. This is the main cause of the non-linearity.

4 Conclusions and Future Work

We have developed two MapReduce applications in biomedical text mining: BLS and BAM. While the former has limited computations, the latter is much more computationally intensive. Our experiment results indicated that both applications' scalabilities were not linear in term of the number of computing nodes. When the number of nodes was increased, more data became non-local to map and reduce tasks and incurred more network data accesses. Thus the time cost to read the data does not decrease proportionally, as the number of nodes increases. BAM achieved better scalability than BLS since BLS performed less computations and were primarily dominated by overheads, such JVM startup, scheduling, disk I/O etc. These observations imply that existing MapReduce framework may not be suitable for on-line systems such as literature search that needs quick response. With the existing MapReduce

framework, mainly designed to maximize parallel throughput for batch processing, it is impossible to reduce the execution time to a user-acceptable level (e.g., within a couple of seconds) by simply increasing the number of nodes.

In the future, we will investigate a real-time Biomedical Text Mining system by optimizing the existing Hadoop framework for high performance.

References

1. Lee, K.-H., Lee, Y.-J., Choi, H., Chung, Y.D., Moon, B.: Parallel data processing with mapreduce: a survey. AcM sIGMoD Record **40**(4), 11–20 (2012)
2. Doulkeridis, C., Nørvåg, K.: A survey of large-scale analytical query processing in mapreduce. The VLDB Journal-The International Journal on Very Large Data Bases **23**(3), 355–380 (2014)
3. Horvath, A.R.: From evidence to best practice in laboratory medicine. The Clinical Biochemist Reviews **34**(2), 47 (2013)
4. Lee, M., Cimino, J.J., Zhu, H.R., Sable, C., Shanker, V., Ely, J.W., Yu, H.: Beyond information retrieval—medical question answering. In: AMIA annual symposium proceedings, vol. 2006, p. 469. American Medical Informatics Association (2006)
5. Hersh, W., Voorhees, E.: Trec genomics special issue overview. Information Retrieval **12**(1), 1–15 (2009)
6. National Center for Biotechnology Information. U.S. national library of medicine, January 2016. http://www.ncbi.nlm.nih.gov/pubmed
7. Agrawal, R., Srikant, R., et al.: Fast algorithms for mining association rules. In: Proc. 20th int. conf. very large data bases, VLDB, vol. 1215, pp. 487–499 (1994)
8. Shetty, K.D., Dalal, S.R.: Using information mining of the medical literature to improve drug safety. Journal of the American Medical Informatics Association **18**(5), 668–674 (2011)
9. Rong, X., Wang, Q.Q.: Large-scale combining signals from both biomedical literature and the fda adverse event reporting system (faers) to improve post-marketing drug safety signal detection. BMC bioinformatics **15**(1), 17 (2014)
10. Chowdhury, G.: Introduction to modern information retrieval. Facet publishing (2010)
11. Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to data mining (2006)
12. Geng, L., Hamilton, H.J.: Interestingness measures for data mining: a survey. ACM Computing Surveys (CSUR) **38**(3), 9 (2006)
13. TREC Genomics Track, April 2015. http://skynet.ohsu.edu/trec-gen/
14. Aronson, A.R., Lang, F.M.: An overview of metamap: historical perspective and recent advances. Journal of the American Medical Informatics Association **17**(3), 229–236 (2010)

Algorithmic Approaches for a Dependable Smart Grid

Wolfgang Bein, Bharat B. Madan, Doina Bein and Dara Nyknahad

Abstract We explore options for integrating sustainable and renewable energy into the existing power grid, or even create a new power grid model. We present various theoretical concepts necessary to meet the challenges of a smart grid. We first present a supply and demand model of the smart grid to compute the average number of conventional power generator required to meet demand during the high consumption hours. The model will be developed using Fluid Stochastic Petri Net (FSPN) approach. We propose to model the situations that need decisions to throttle down the energy supplied by the traditional power plants using game-theoretic online competitive models. We also present in this paper the power-down model which has shown to be competitive in the worst case scenarios and we lay down the ground work for addressing the multi-state dynamic power management problem.

Keywords Power grid · Renewable energy · Sustainable energy · Petri nets · Power down problem · Online algorithm · Competitive analysis

1 Introduction

The electricity grid is said to be the most challenging engineering endeavor in human history. Electric power is the driving engine of our society. With the shift

W. Bein(✉) · D. Nyknahad
Department of Computer Science, University of Nevada Las Vegas,
Las Vegas, NV 89154, USA
e-mail: wolfgang.bein@unlv.edu, nyknahad@gmail.com

B.B. Madan
Modeling, Simulation & Visualization Engineering, Old Dominion University,
Norfolk, VA 23529, USA
e-mail: bmadan@odu.edu

D. Bein
Department of Computer Science, California State University,
Fullerton, Fullerton, CA 92834, USA
e-mail: dbein@fullerton.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_59

towards renewable energy sources, such as biomass, wind, and solar on the power supply side and smart appliances and electric vehicles on the load side, there are enormous challenges in designing a reliable, effective and secure power infrastructure. The existing grid is primarily an open loop system designed to support only one-way dependable flow of electrical power supplied by conventional fossil, hydro and nuclear power plants. Tremendous transformations in information technology benefiting sectors such as industrial automation, health, finance, and communication, support the need for a radically different grid, which manages variable and distributed supply and demand from large battery storage systems, web-enabled appliances or electrical vehicles. There is a pressing need for new models, algorithmic techniques, and appropriate security designs such that the new grid can take advantage of technological advances.

A smart grid will incorporate sustainable energy sources to reduce the dependence on conventional power generation sources. The smart grid initiative seeks to augment the existing power grid by integrating sustainable and renewable power sources on the supply side, EVs, smart buildings and smart appliances on the demand side and extensive use of available information and communication technologies. Whereas, the existing grid has focused only on the unidirectional flow of electricity from the suppliers to the consumers, the goal of the smart grid is to introduce environmental friendliness and high degree automation through bidirectional flow of electricity and information as shown in the conceptual diagram of smart grid (Fig. 1).

National energy policy mandates 40% increase in the current installed renewable bulk wind energy and solar photo-voltaic (SPV). Both these sources of energy are geographically distributed. In addition, homes are increasingly

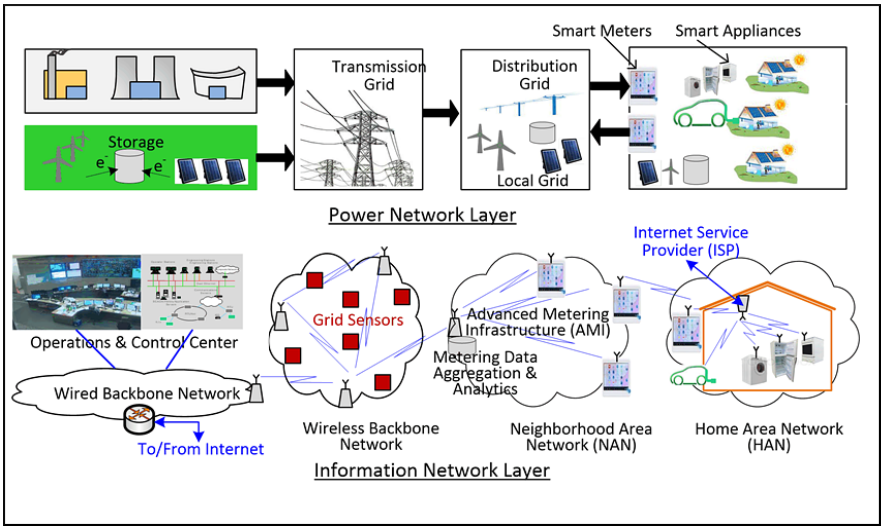


Fig. 1 Smart Grid Power and Information Flow Layers

installing solar panels and SPV shingles, thus becoming widely and renewable energy sources feeding the grid. However, electrical energy produced by both wind and solar power generation are highly weather dependent and weather cannot be controlled. Similarly, supply from home based renewable sources is hard to predict not only due to unpredictable weather but also due to uncertain demand for power in homes. Transition of the existing grid to the smart grid, while solving many important problems, also present challenges. Success of the smart grid is contingent upon meeting the following challenges:

- Minimize the number of non-renewable power generators that are idled during off-peak hours through "Supply-Load" and "Load-Supply".
- Identify new models to forecast load and supply for the smart grid since existing models used in the existing grid do not work for the smart grid [1, 2, 3].
- Identify and address smart grid cybersecurity challenges created by the integration of the electric grid with the information networks [4,5].

Sustainability, as defined by the National Science Board (NSB), involves meeting present needs without compromising the ability of future generations to meet their own needs, and cuts across disciplinary boundaries, because it has many aspects that cannot be addressed by individual disciplines [6]. Green computing is a relatively new area of computer science, and many questions regarding it are still open; see papers by Pruhs et al. [7, 8, 9] and articles [1], [10], [11] by other authors.

We propose to develop a supply and demand model of the smart grid to compute the average number of conventional power generator required to meet demand during the high consumption hours. The model will be developed using Fluid Stochastic Petri Net (FSPN) approach. The parameters, namely the transition firing time probability distributions, the number of tokens in discrete places and the capacity of continuous (fluid) places, will be estimated using real smart grid data. When renewables produce a surplus of energy, such surplus generally does not affect the operation of traditional power plants. Instead, renewables are throttled down or the surplus is simply ignored. In a future, where more than two-thirds of all domestic power would be generated by renewables this is not tenable. Power-down problems play an important role in managing a renewable energy infrastructure. It be could argued that a game-theoretic approach which assumes an omniscient adversary may not be so realistic for modeling the grid. However, in order to construct a truly resilient grid, worst case assumptions should thus be taken into account. We present in this paper the power-down model which has shown to be competitive in the worst case scenarios and we lay down the ground work for addressing the multi-state dynamic power management problem.

The paper is organized as follows. Preliminary notions and the Petri Nets model are given in Section 2. Models for managing a renewable energy infrastructure using algorithms for the power-down problem are presented in Section 3. We conclude in Section 4 and propose directions for future work.

2 Smart Grid Supply and Load Management Modeling Using Petri Nets

Designing and analyzing a grid is non-trivial. Performances of the management in real conditions strongly depend on the accuracy of the forecasts and of the mode of operations. The existing grid has been remarkably successful and reliable, which can be attributed to the accurate modeling of load pattern on the grid to accurately predict the demand. The demand prediction allows power generation to be easily adapted to the predicted demand. Fig. 2 (left), reproduced from [12], shows the demand curve along with three different power generation sources that are used to meet the demand.

Fig. 2 (left) highlights two important issues:

- Average load on the grid is much smaller than the peak load. However, to keep the grid stable, the power generation capacity was designed to meet the peak demand. This led to almost 80% of the available generation capacity being idled during the off-peak hours resulting in poor utilization, which is a tremendous financial overhead on power generation and further to the consumers.
- The power sources employed to meet the demand have large carbon footprint. Since large carbon foot print and nuclear waste have serious environmental ramification, there is strong push towards alternative and sustainable energy technologies such as wind, solar, geothermal, wave, etc., and bulk energy storage to address peak load situations.

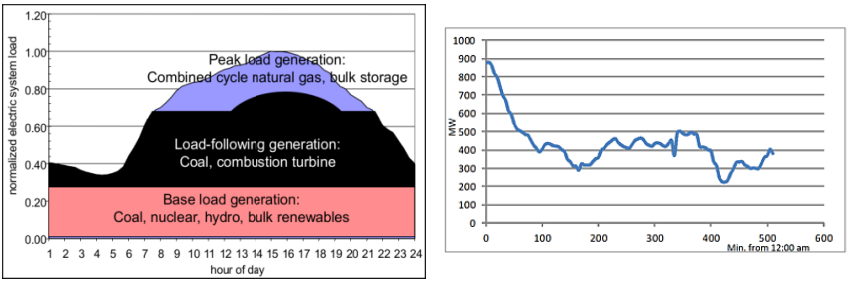


Fig. 2 Electrical Grid Load Pattern and Wind Power Supply Variations

Looking at the load curve shown in Fig. 2 (right), it is quite possible that there may be strong winds during 1-5 am and dying winds during the peak demand hours of 3-7pm. Fig. 2 (right) plots the variations in the supplied wind power to the Washington DC area on 2/10/2015 from 12:00am - 8:30am (based on manual digitization of the graph in 13).

Stochastic Petri Nets (SPN) [14] and Fluid Stochastic Petri Nets (FSPN) [15,16] are a subset of Petri Nets PN [17]. Each PN is a bipartite graph with two type of vertices: Places (denoted by a circle with label P_{XX}) and Transitions

(denoted by a rectangle with T_{XX}), such that a Place can only have an edge to or from a Transition and vice-a-versa. In other words, no two places (or transitions) are connected to each other by an edge. The primary objective of developing the FSPN is to minimize the conventional power generation capacity that has to be turned on when the demand exceeds the available base supply (Fig. 2, left). Since such phenomena occur only for a small fraction of a 24-hour daily cycle, in a conventional grid, it results in conventional power being active to meet high demand and then being idled.

An FSPN consists of discrete (shown as single line circles) and continuous (shown as double lined circles) places. Discrete places contain tokens, while continuous places contain fluid (electrical power in our case). Transitions are used to model state changes in a system. A transition is said to be *ready* if the place (the head-end of a directed edge) contains at least one token. If the number of tokens k >1, then we label the edge with k, which is referred to as the edge multiplicity. In general a transition can be either immediate, i.e. it fires as soon as it is ready, or timed, i.e. it fires after a delay of random time with known distribution. Both SPN and FSPN are suitable for modeling distributed systems that need to deal with many random and concurrent events. Since renewable energy sources are geographically distributed and their energy production depends on the random weather behavior, we conjecture that SPN and FSPN are also well suited for modeling and prediction of supply and load in a smart grid. While SPN can deal with only discrete items, e.g., number of working or failed units. The other entity to be dealt with is the power supplied and power consumed. While the later can be quantized (e.g., KW or MW), it is in fact best modeled as a continuous variable, similar to fluid flow. Since FSPN can better model such continuous variables, this project will propose to develop an FSPN model of the smart grid. Fig. 3 shows a simple FSPN that captures varying load (demand) and supply from conventional and sustainable energy sources of a smart grid.

We only consider timed transitions. For example, the transition TSHM takes tokens from the place PSGH, which models sustainable energy sources (e.g., wind turbines) working at their highest capacity, and KSGH, which models the number of such sources at any arbitrary time. The power supplied by these is modeled by double-lined directed edge from TSHM to the continuous place PSUP, which models the amount of energy that the grid can provide to the load. The transition TSUL models the delivery of electrical power from the continuous place PSUP to the continuous place PLD, and can be modeled as an immediate or a timed transition. We choose to model it as a timed transition to represent the time that be taken to condition the supply (e.g., power factor) to the load. Since sustainable (e.g., wind energy) sources depend on the prevailing atmospheric conditions, for simplicity in this example FSPN models such changes as 3-level phenomenon, i.e. "High", "Medium" and "Low" wind speeds, and transitions "High" to "Medium" to "Low" to "High" wind speed transitions. The FSPN model to be developed can have other types of transitions, e.g., "Medium" to "High", with specified probability.

The objective of smart grid is to minimize the contribution of the non-sustainable conventional power generation plants through the use of sustainable energy sources. The red shaded block of the FSPN model (Fig. 3) models the contribution of the conventional power generation sources towards meeting the load following and peak demand. It provides conceptual underpinning of how such a model will achieve this goal. The variable KBGW (# tokens in the place PBGW) models the number of conventional power generation plants that a conventional grid requires to meet demand above the base level demand.

Since one of the primary objectives of the smart grid is to substitute such big commercial scale plants with sustainable energy sources, KBGW is the key metric that needs to be minimized. The transition TBWR models the retirement of such plants, whose firing rate transition is controlled by a guard function gRG(LH,LP), where LB, LH and LP denote the power supply levels required to meet the demand during the `Base', `Load following' and `Peak' phases of the demand curve (Fig. 2) at any time. For example, when the supply rises to level $\geq$ LH, the guard function [14] behavior is such that it increases the rate of switching OFF (or retiring) conventional power plants that were put in place to meet the load following demand requirements. On firing, this transition removes a token from the place PBGW and moves this token to PBGR, thus reducing (increasing) by 1, the number of working (retired) conventional plants. Similar behavior is achieved by making gRG a function of LP for retiring conventional power plants that meet peak load requirements. The transition TRRW models the need to replace or rejuvenate retired plants to model situation in which sustainable sources are unable to meet the demand and conventional power plants need to be increased.

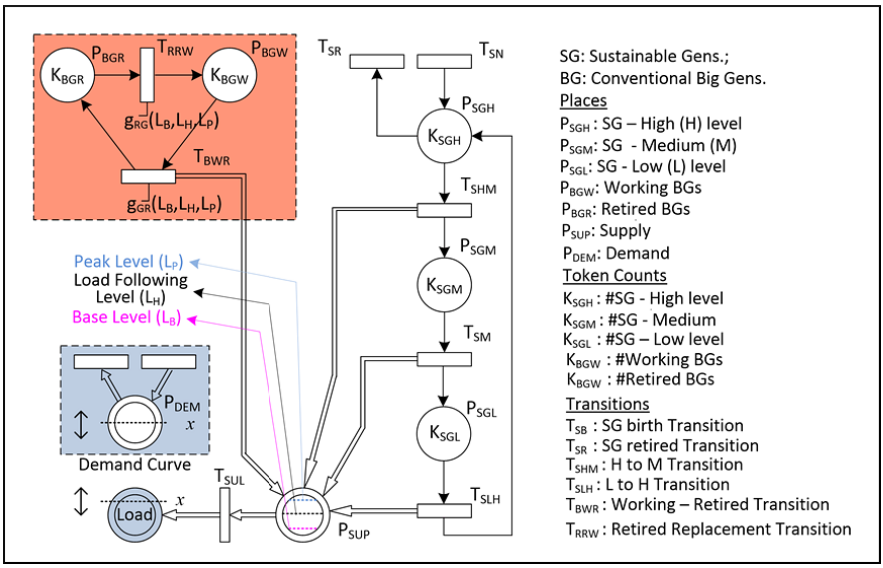


Fig. 3 Fluid Stochastical Petri Net Model for the Smart Grid

It should be emphasized that since both - the demand and the contribution of the sustainable power - are random functions of time, all FSPN transitions in Fig. 3 fire at random instant time according some probability distribution. Therefore, instantaneous KBGW values are meaningless and statistical expectation E[KBGW] will need to be used as metric of the effectiveness of sustainable sources in the smart grid. We will apply the FSPN solution technique described in [15,16] and use real field data to estimate the distribution of firing times of different transitions and obtaining flow rates to/from continuous places, state of FSPN places and probability of reaching different states of the stochastic process underlying the FSPN.

3 Generator Power Down

Electrical energy supplied by sustainable energy sources is unpredictable due to its dependence on the weather, for example. When renewables produce a surplus of energy, such surplus generally does not affect the operation of traditional power plants. Instead, renewables are throttled down or the surplus is simply ignored. In a future, where more than two-thirds of all domestic power would be generated by renewables this is not tenable. Instead it may be the traditional power plant that will need to be throttled down. Taking a large power plant offline is not a simple matter and causes considerable extra cost. Thus power-down problems play an important role in managing a renewable energy infrastructure. Assuming that the base power load generators are always active (as shown in Fig. 2), the problem to be solved is:

Given n conventional generators generating c_1,... c_n units of power and m renewable energy sources generating r_1, ... r_m units of power, for high load situations, decide the time instants at which conventional power generators can be turned OFF to support the current demand.

The transition cost from the ON state to the OFF state is assumed to be 0, while the cost of transition from OFF to ON for the i^{th} conventional generator is $a_i > 0$. For $n > 1$ the problem becomes the problem becomes similar (though not identical) to the multi-state dynamic power management problem. Dynamic power management or DPM refers to the problem of judicious application of various low power techniques based on runtime configurations in an embedded system to minimize the total energy consumption. The power manager (PM) can change the power consumption of a device through selection of shutdown/sleep/wakeup states for the device. (See for example [18, 19,20])

We have identified four distinct problems relevant to the smart grid which can leverage the field of online algorithms:

- Single conventional power plant and 2-state power generation (ON/OFF) model.
- Single conventional power plant and multi-state power generation model.
- Multiple conventional power plants with 2-state power generation model.
- Multiple conventional power plants with multi-state power generation model.

Suppose now that a device has two states, the ON state and the OFF state, and there is a need to develop a control algorithm for changing between these states so as to minimize energy when given a sequence of service requests. When a user makes a request, the device must remain in the ON state for some interval of use. At the end of that interval, the device can be left in the ON state, or can be turned off at any time afterward, unless it is being used again. In the ON state, the device uses one unit of cost per unit of time. We call this the running cost. There is no cost in the OFF state. There is also a constant switching cost a>0 for turning the device on and no cost for turning it off. A well-known optimal online algorithm for this problem is an adaptation of the "ski rental" algorithm SkiRental [21,22]. As Pruhs pointed out in his tutorial on sustainable and green computing at FOCS 2011 (see [8,9]), there are many results on algorithmic issues in sustainable computing, perhaps most notably, on speed scaling (e.g., [23],[8]). However, solutions to the power down problem are based on variants of the simple ski rental algorithm. The algorithm SkiRental [21, 22] turns the device on when a user requests service. After use, SkiRental waits for the next service request for time t, where t is the switching cost divided by the running cost (see Fig. 4). If another user requests service within the period, SkiRental does not turn off; otherwise, SkiRental turns off after time t. Algorithm SkiRental is 2-competitive, which is optimal.

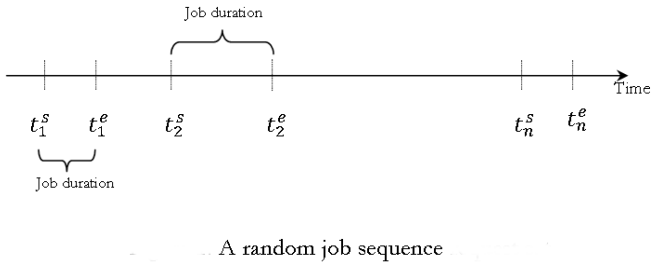


Fig. 4 Device states, running cost, and switching of states costs

SkiRental is quite unsatisfactory for use in sustainable computing. For example, if the arrival interval of users is spaced long enough, then SkiRental keeps the device in the ON state for extra *a* units for each user, which seems quite wasteful. Since SkiRental is optimal, it appears as if there is not much hope for improvement. However, if we allow the worst-case competitive ratio to increase by only a small positive constant ε from 2, we can design numerous algorithms other than SkiRental. We call such algorithms ε -OPT [24], [25]. We can gradually decrease the duration of keeping the ON state "standby time" when the frequency of the device usage becomes low.

Gradually decreasing the duration of the standby time for DRA algorithm can be explored further in various directions. If the waiting time - the time from the preceding user leaving to the latest user arriving -- was more than a, something

like an “off-peak" situation, then the system may be slack and the algorithm may elect to make the standby time shorter. On the other hand, if the preceding waiting time was less than or equal to a, the algorithm may reset the standby time. The slackness parameter considered here is only one example of interest, others should be considered. We propose to work on systems where there are more than two states which are of interest both in theory and in practice.

4 Conclusions and Future Work

In this paper we explore options for integrating sustainable and renewable energy into the existing power grid and we lay down the ground work for addressing the multi-state dynamic power management problem, or even create a new power grid model. We propose to model the situations that need decisions to throttle down the energy supplied by the traditional power plants using game-theoretic online competitive models. We present various theoretical concepts necessary to meet the challenges of a smart grid. In each area we have identified the main technical challenges and elaborated how game theory can be applied to address these challenges.

We first present a supply and demand model of the smart grid to compute the average number of conventional power generator required to meet demand during the high consumption hours. The model will be developed using Fluid Stochastic Petri Net (FSPN) approach. We also present in this paper the power-down model which has shown to be competitive in the worst case scenarios. Moreover, we proposed several future directions for extending these two approaches, Petri Nets and power down algorithms, so as to reduce the gap between game theoretical models and practical implementations of future smart grids.

It will be also of interest for the multi-state dynamic power management problem to investigate other dynamic game theoretic models, in cooperative and non-cooperative settings, and their applications in smart grid systems, due to the pervasive presence of time-varying parameters in smart grids such as generation and demand.

Another venue of research relevant to the smart grid refers to the cyber security problems. Smart grid information traffic and its control systems are real time systems. However, the existing cyber security paradigm is based on preventing, detecting and finally recovering from an attack. This paradigm does not fit well with the requirements of real time systems. Developing a resilient cyber infrastructure architecture with the intrinsic ability of attack tolerance, similar to the failure tolerance model for critical systems is of high interest.

Acknowledgment Discussions with Rüdiger Reischuk of Universität Lübeck during the sabbatical visit of Wolfgang Bein are acknowledged. The work of author Wolfgang Bein was supported by National Science Foundation grant IIA 1427584. The work of Doina Bein was supported by California State University Professional Development Support.

References

1. Balantrapu, S.: Load Forecasting in Smart Grid (2013). http://www.intelligentutility.com/article/13/10/load-forecasting-smart-grid (available online)
2. Borges, C.E., Penya, Y.K., Fernandez, I.: Evaluating Combined Load Forecasting in Large Power Systems and Smart Grids. IEEE IEEE Transactions on Industrial Informatics **9**(2), 1570–1577 (2013)
3. McDonald, J., McGranaghan, M.: Strategic R&D Opportunities for the Smart Grid (2013). http://www.nist.gov/smartgrid/upload/Final-Version-22-Mar-2013-Strategic-R-D-Opportunities-for-the-Smart-Grid.pdf (available online)
4. Karlin, A., Manasse, M., McGeoch, L., Owicki, S.: Randomized competitive algorithms for non-uniform problems. In: Proceedings of the 1st Symposium on Discrete Algorithms (SODA), pp. 301–309 (1990)
5. Irani, S., Shukla, S., Gupta, R.: Online Strategies for Dynamic Power Management with Multiple Power-Saving States. ACM Trans. Embedded Computing Systems **2**(3), 325–346 (2003)
6. National Science Board Building a Sustainable Energy Future: U.S. Actions for an Effective Energy Economy Transformation (2009)
7. Pruhs, K.: Green computing algorithmics. In: Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS), pp. 3–4 (2011a)
8. Irani, S., Pruhs, K.R.: Algorithmic problems in power management. SIGACT News **36**(2), 63–76 (2005)
9. Pruhs, K.: Introduction to special issue on theoretical aspects of green computing. Sustainable Computing: Informatics and Systems **1**(3), 165–166 (2011)
10. Goth, G.: Chipping away at greenhouse gases. Communications of the ACM **54**(2), 13–15 (2011)
11. Mitchell, R.L.: Power struggle: how IT managers cope with the data center power demands. Computerworld (2006)
12. Bushby, S.T.: NIST Smart Grid Interoperability Program (2015). http://www.nist.gov/el/upload/4-2-Bushby-Smart-grid.pdf (available online)
13. Wind Power. http://www.pjm.com/about-pjm/renewable-dashboard/wind-power.aspx (available online)
14. Trivedi, K.: Probability and Statistics with Reliability, Queueing, and Computer Science Applications, 2nd edn. John Wiley (2001)
15. Horton, G., Kulkarni, V.D., Nicol, D.M., Trivedi, K.S.: Fluid stochastic Petri Nets: Theory, applications, and solution techniques. European Journal of Operational Research **105**(1), 184–201 (1998)
16. Gribaudo, M., Sereno, M., Bobbio, A., Horvath, A.: Fluid Stochastic Petri Nets augmented with Flush-out Arcs: Modeling and Analsys. Discrete Event Dynamic Systems **11**(1&2), 97–117 (2001)
17. Murata, T.: Petri nets: properties, analysis and applications. Proceedings of the IEEE **77**(4), 541–579 (1989)
18. Albers, S.: Energy efficient algorithms. Communications of the ACM **53**(5), 86–96 (2010)
19. Augustine, J., Irani, S., Swamy, C.: Optimal power-down strategies. SIAM Journal on Computing **37**(5), 1499–1516 (2008)
20. Baghzouz, Y.: Evaluation of an energy storage system on a distribution feeder with distributed PV systems. In: CIRED, Paper No. 0123, Lyon, France (2010)

21. Karlin, A., Manasse, M., McGeoch, L., Owicki, S.: Competitive randomized algorithms for nonuniform problems. Algorithmica **11**, 542–571 (1994)
22. Karlin, A., Manasse, M., Rudolph, L., Sleator, D.: Competitive snoopy caching. Algorithmica **3**, 79–119 (1988)
23. Bansal, N., Kimbrel, T., Pruhs, K.: Speed scaling to manage energy and temperature. J. ACM **54**(1), 1 (2007)
24. Bein, W., Hatta, N., Hernandez-Cons, N., Ito, H., Kasahara, S., Kawahara, J.: An online algorithm optimally self-tuning to congestion for power management problems. In: Proceedings of the 9th Workshop on Approximation and Online Algorithms (WAOA 2011), pp. 35–48 (2012)
25. Andro-Vasko, J., Bein, W., Ito, H., Nyknahad, D.: Evaluation of online power-down algorithms. In: Proceedings of the 12th International Conference on Information Technology (ITNG) (2015)

Performance Evaluation of Data Migration Methods Between the Host and the Device in CUDA-Based Programming

Rafael Silva Santos, Danilo Medeiros Eler and Rogério Eduardo Garcia

Abstract CUDA-based programming model is heterogeneous – composed of two components: host (CPU) and device (GPU). Both components have separated memory spaces and processing units. A great challenge to increase GPU-based application performance is the data migration between these memory spaces. Currently, the CUDA platform supports the following data migration methods: UMA, zero-copy, pageable and pinned memory. In this paper, we compare the zero-copy performance method with the other methods by considering the overall application runtime. Additionally, we investigated the aspects of data migration process to enunciate causes of the performance variations. The obtained results demonstrated in some cases the zero-copy memory can provide an average performance on 19 % higher than the pinned memory transfer. In the studied situation, this method was the second most efficient. Finally, we present limitations of zero-copy memory as a resource for improving performance of CUDA applications.

1 Introduction

By considering the involved architectures, CUDA-based programming model is heterogeneous – composed of two components: host (CPU) and device (GPU) [5]. Both components have separated memory spaces and processing units. A great challenge for performance increasing in CUDA-based programming is the data migration between host and device memory spaces [7]. It is not a simple task to manage this data transfer, even without focusing in performance, since in some methods it is necessary to perform explicit requests for memory copy and the control of concurrent data access [8]. At present, the latest CUDA version provides four main methods for data migration between host and device memory spaces: zero-copy memory, UMA (Unified Memory Access) model, pageable and pinned memory [11].

R.S. Santos · D.M. Eler(✉) · R.E. Garcia
Faculdade de Ciências e Tecnologia, UNESP - Univ Estadual Paulista,
Presidente Prudente, SP, Brazil
e-mail: rafael.silva.sts@gmail.com, {daniloeler,rogerio}@fct.unesp.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_60

In this paper, we investigate the zero-copy migration method, by comparing the performance with the other three memory migration methods provided by CUDA API and evaluating all runtime tests. We conducted a case study for the migration process. First, we developed an application to perform a scalar multiplication and the sum from the elements of a vector. Then, we adapted this application to execute and to evaluate two situations. In the first, less memory access is performed by the GPU during threads execution; in the second, a greater number of access transactions is performed. Finally, we present the limitations of using zero-copy memory method as an approach to increase a CUDA application performance.

Besides the introduction, this document consists of another five sections. Related works are introduced in Sect. 2. In Sect. 3, we describe the main data migration methods between host and device on the CUDA architecture. Then, in Sect. 4, we present the results. In this section, we evaluate and discuss the performance of the methods described in Sect. 3, comparing them with zero-copy memory method. In Sect. 5, we present some limitations in the use of zero-copy memory. Finally, Sect. 6 concludes this paper.

2 Related Works

In the last years, various studies on the transfer of data between the different architectures of the GPU-based programming model were produced. Kaldewey et al. [4] conducted a study on the efficiency of use of the bandwidth of the PCI-E bus in communication between the host and the device in the UVA model. By zero-copy memory they demonstrated the performance is close to the theoretical maximum bandwith memory. However, in that study, a comparison of the impact of the use of this memory in the global application performance compared to other transfer methods was not performed. Bai et al. [1] used the zero-copy memory to optimize algorithms performance for lattice-based cryptographic systems. In the tests, they analyzed both single GPU and multiple GPUs communicating with the CPU. Landaverde et al. [7] conducted a research of UMA model performance from the pageable memory transfer. In the methodology, they established a benchmark model to the methods similar to the model used in this work. Based on the results, they demonstrated that the use of UMA model causes a performance loss regarding to the pageable and pinned memory.

3 Data Migration Between Host and Device

3.1 CUDA Main Memory Access Model

Compute Unified Device Architecture (CUDA) is a platform developed by NVIDIA to allow use of GPUs in general-purpose computing. CUDA-based programming

Fig. 1 Heterogeneous CUDA-programming model

model is heterogeneous and each component, host and device, owns processing unit and memory space [11].

The rate that a processing unit can access the main memory is limited by its self **memory bandwidth**. In CUDA-based programming model, in addition to memory bandwidth of GPU and CPU, there is a **bus** between the host and device, as shown in in Fig. 1. Typically, the bus that mediates this communication is the PCI (Peripheral Component Interconnect). The PCI memory bandwidth is generally less than CPU and GPU [10]. Thus, it is important to analyze the memory transfer impacts on the design of a CUDA application.

3.2 CUDA Data Transfer Methods

Along the versions, NVIDIA introduced different memory management methods on the CUDA platform. By the time this study was conducted, the latest version of CUDA SDK – CUDA 7 SDK – offers the following methods: pageable memory, pinned memory, zero-copy memory, and Unified Memory Access (UMA) model [11].

Standard Transfer Method: Standard transfer method can be executed with pageable memory or pinned memory. In a **pageable memory**, during the execution of an application, the physical address of the data may change, given the memory pages may undergo swap for the secondary memory. In this way, before the data is copied from the host to the device, architecture migrates the desired data portion for pinned memory buffer on the host [11]. Then, through the PCI resource Direct Memory Access (DMA) performs the data transfer from the buffer to the device [3]. In contrast to pageable memory, **pinned memory** does not undergo swap. Thus, in standard transfer method of pinned memory, it does not occur to data migration to the buffer. This feature allows the PCI use DMA to directly transfer the data from the current physical location in host memory to the device memory space [10]. In pageable memory, the additional transfer contributes to a decrease in performance (memory bandwidth) regarding to pinned memory [2].

Zero-copy Method: Zero-copy memory is a "kind" of pinned memory provide by Unified Virtual Addressing (UVA). The use of zero-copy method discards the need

to explicit requests (i.e., a function call) for data migration. UVA model provides a single shared virtual space of memory between the host and the device, providing the address of the data on the host is mapped to the device [9]. Data migration occurs implicitly whenever a portion that is not presented in context is referenced, either the host or the device [10]. Thus, the zero-copy memory is used when a set of data does not fit completely in the device memory; once the CUDA API maintains only a portion of data that is accessed at a specific time in the device [6]. The remaining data set is maintained in host memory [10]. By considering the use of multiple GPUs, UVA model also provides a shared memory space and eliminates the intermediate step to copy the host memory during data transfer from one GPU to another. In the state of the art, the aspects of data movement in zero-copy memory provide by CUDA API are not disclosed.

UMA Model: Unified Memory Access (UMA) model is similar to zero-copy memory regarding no need for an explicit request for data migration. On the other hand, the data is not allocated in a pinned memory. As in zero-copy memory, the API is responsible for managing the entire data migration process.

4 Case Study: Evaluation of Memory Management Methods

In this section, we evaluate the performance of data migration between the host and device memory spaces. In addition to the performance comparison, we investigated the aspects that justify the variation between migration methods.

In the tests, we setup two different situations for the scenario: in the first, described in Subsect. 4.1, we created a simulated application in which the GPU makes a small amount of memory access transactions for each kernel thread; in the second, we adapted this application to the GPU performs a greater amount of memory access transactions, as shown in Subsect. 4.2.

4.1 Situation 1: Less Amount of Memory Access Transactions During a Thread Execution

Application Model: For the tests, we have developed an application that performs scalar multiplication and sum from vector elements. The implemented algorithm can be divided into the following steps:

1. Initialization of vector elements
2. Data transfer to the device (HtoD)
3. Vector scalar multiplication (kernel)

4. Data transfer to the host (DtoH)
5. Sum of vector elements

The step that comprehends the vector scalar multiplication was parallelized and runs on the GPU. The code snippet that is executed in this step can be seen in Fig. 2.

```
1  _ _global_ _ void mult(int *vect, int num, int N){
2          //Thread index
3          int id = blockIdx.x * blockDim.x + threadIdx.x;
4          //Multiplication = vector element * constant value
5          if (id < N)
6                  *(vect + id) = *(vect + id) * num;
7  }
```

Fig. 2 Code snippet that runs on the device for vector multiplication by a constant.

The host is responsible for the initialization steps and sum of vector elements. These steps are performed sequentially. Throughout the algorithm, all steps use the vector as the Input and Output Dataset. After the initialization of the vector, it is necessary to migrate this set of data from the host to device (HtoD). Then, after multiplication by a scalar value, the vector should be migrated back to the host (DtoH). The context of CUDA application comprehends a single stream. A pseudo-code containing all algorithm steps is presented in Algorithm 1. The code snippet that contains the parallelized instructions was previously shown in Fig. 2.

To compare and investigate the performance of each memory management method, we adapted the application with the necessary function calls to allocate and copy memory. It is worth mentioning that we focus exclusively on evaluating the performance differences and research aspects of the data migration between the host and the device. Thus, the implemented algorithm is not complex and does not require a large computational effort.

Algorithm 1. Scalar multiplication and sum of vector elements

Require: $n > 0$
Ensure: $vec \leftarrow vec * c$
  $vec \leftarrow Allocation$
  $c \leftarrow k_1$ {k_1 is a constant}
  for $i = 1, i \leq n, i++$ **do**
    $vec[i] = k_2$ {k_2 is a constant}
  end for
  $\Rightarrow$ *migration data* – HtoD
  mult<<<N, 1>>> (vec, c, n) {CUDA kernel}
  $\Rightarrow$ *migration data* – DtoH
  $sum = 0$
  for $i = 1, i \leq n, i++$ **do**
    $sum = sum + vec[i]$
  end for

Methodology: We developed an application which performs a scalar multiplication and sum from the vector elements (see the Subsect. 4.1) for our microbenchmarks. NVIDIA Visual Profiler and nvprof tools were used to measure processing time and to obtain details of the data transfer aspects throughout the runtime. We run the application with four different configurations of data migration methods provided by CUDA platform.

In the tests, we use five different vector sizes. The arrangement of threads and threads blocks on kernel function was performed based on these sizes. Number of elements in each vector size: 1048576 (1024 threads x 1024 blocks), 2097152 (1024 threads x 2048 blocks), 4194304 (1024 threads x 4096 blocks), 8388608 (1024 threads x 8192 blocks) and 16777216 (1024 threads x 16384 blocks) elements.

The timeline during runtime for each test was divided into five ranges as the steps that have been described in the implemented algorithm. To define each time range, we use NVTX (NVIDIA Tools Extension). We named these ranges as follows: initialization, HtoD transfer, multiplication, DtoH transfer and sum. It is easy to associate each range with the described algorithm steps, which are presented in Subsect. 4.1. Data transfer ranges were not considered in the UMA model and the zero-copy memory, once there is no explicit memory copy in these methods.

All results represent an average of five executions. All the tests were performed on computer with Windows 8.1 64 Bits Operating System, Intel I5-2320 CPU, 4 GB RAM, PCI Express x16 2.0 bus, NVIDIA Geforce GT 740 GPU, with 1 GB RAM DDR3 and 384 CUDA cores. The application was implemented using the CUDA 7 SDK.

Experimental Results: In this section, we present results of the tests. In order to provide a better understanding, all values were normalized.

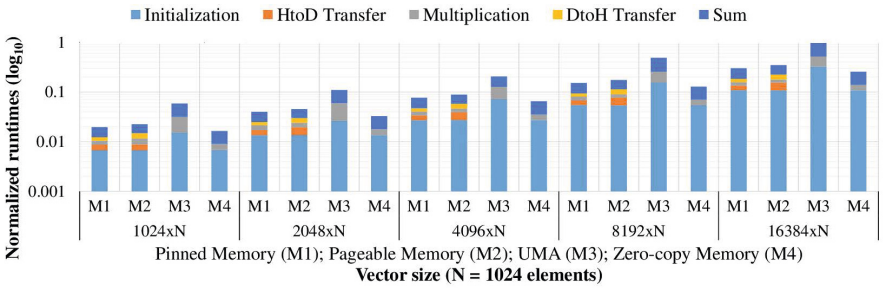


Fig. 3 Global normalized runtime for the application of scalar multiplication and addition of vector elements (Situation 1).

Figure 3 shows the normalized runtimes for application that performs the scalar multiplication and the sum from the vector elements. In particular, the results match the configured application in four memory management methods: pageable memory (represented by M1), pinned memory (represented by M2), UMA

(represented by M3) and zero-copy memory (represented by M4). Apart from the global runtime, it is also shown the runtime of each range that the application was divided: initialization, HtoD transfer, multiplication, DtoH transfer and sum.

From Fig 3, we can see that there was a distinct variation in application performance among the data migration methods. Additionally, it is possible to observe the results were consistent, once the performance variation pattern in retained once the array size is increased. From the obtained results, we ordered the migration data methods regarding best described performance: zero-copy memory, UMA model, pageable memory and pinned memory. In all presented tests, this order of efficiency is the same. On average, the application configured with the standard transfer method with pinned memory spends 19.40% more time than zero-copy memory; with pageable memory, 37.24% more time was spent; and in UMA model, it was 256.57%.

In order to figure out the reasons for performance variations among the methods, we evaluate aspects of data migration during tests runtime. In the tests, the configuration with pageable memory spends on average 62.47% more time than the time spent by the pinned memory in HtoD transfer range and 77.40% in DtoH transfer range. Through the NVIDIA Visual Profiler, it is possible to measure the throughput of pinned memory – on average 6.68 GB/s (HtoD) and 6.697 GB/s (DtoH), whereas in the pageable memory was 3.69 GB/s (HtoD) and 3.87 GB/s (DtoH). In the other analyzed ranges, the time consumed by both methods is the same. On average, there is a difference of 0.59% in relation to the run time in the initialization, 1.52% in multiplication, and 2.43% in the sum of vector elements. In the standard transfer method, the data migration does not affect the performance of other ranges that the timeline of the tests was divided. Thus, we used this method to compare and to investigate the data migration aspects and performance in UMA model and zero-copy memory.

In UMA model, there is no explicit memory copy. However, from Fig. 3, of course notice a high discrepancy between the processing time of this method regarding the standard transfer method, when we analyze the ranges of initialization, multiplication and sum from vector elements. On average, the UMA model spent 152% more time than the average of the pageable and pinned memories during initialization, 670% more than multiplication and 251% more than sum range.

In all tests, the zero-copy memory is more efficient. As in UMA model, there is no explicit memory copy. However, the data migration takes place at different times. Figure 4 shows the time consumed during the execution of multiplication range for zero-copy memory and standard transfer methods for all tests. As we can see, once the vector has more than 2097152 (2048x1024) elements, the kernel runtime with zero-copy method becomes greater than the time taken by the standard transfer method with pageable memory and pinned. NVIDIA Visual Profiler tool does not support a graphical analysis of the data migration with zero-copy memory. Although, it is possible to collect information about reading and writing transactions in system memory, i.e., in the host, while running the kernel. In all executed tests, regardless of the vector size, there are on average 2097152 (2048x1024) access transactions to host memory (reading and writing) for the kernel running with zero-copy memory. Each Access transaction features 32-bit width and the transfer rate is on average 5.88 GB/s.

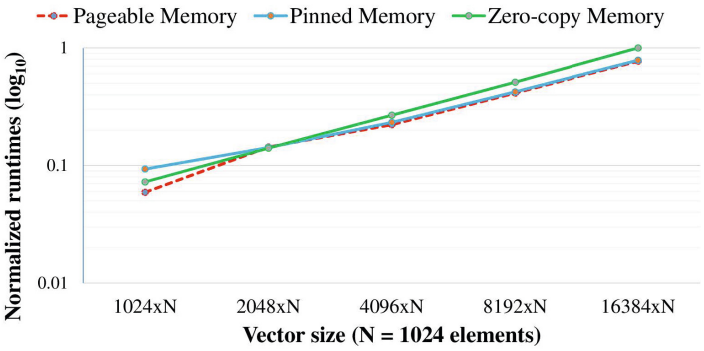


Fig. 4 Multiplication range runtime. Configured application with the memory management methods: zero-copy memory, pageable memory and pinned memory.

Discussion: The results demonstrate the data migration between host and device memory spaces effectively impacts the application performance. Additionally, different memory management methods provided by CUDA, exhibit high performance variation.

Zero-copy memory method was more efficient than other methods. However, we cannot assert that for any application of this method will be more efficient. The results demonstrate that the use of zero-copy memory affects the performance of the kernels, once that occur access transactions to host memory during the execution. Thus, to further investigate the behavior of this method, we modified the kernel function to run the tests again. The modified kernel and the results obtained are shown in Subsect. 4.2.

The pinned memory obtained the second best performance. Regarding pageable memory, the performance difference is caused by the memory throughput. Pinned memory has a transfer rate about 77% higher than the pageable memory. The performance of UMA was lower than all other methods.

4.2 Situation 2: Greatest Amount of Memory Access Transactions During the Execution of a Thread

In order to better investigate the behavior of zero-copy memory and another memory management methods, we adapt the original kernel function shown in Fig. 2. The modified code snippet can be seen in Fig. 5. The kernel modification does not cause changes in the application results. The modification added 49 redundant instructions of each thread execution. By performing this kernel, we intend to simulate an increased amount of memory access transactions while running the kernel. Note that the remaining steps of the algorithm (Algorithm 1) have not been modified. Thus, the produced results by the application are the same.

```
1  _ _global_ _ void mult(int *vect, int num, int N){
2            //Thread index
3            int id = blockIdx.x * blockDim.x + threadIdx.x;
4            //Multiplication = vector element * constant value
5            if (id < N)
6                    for (int i = 0; i < 50; i++)
7                            *(vect + id) = *(vect + id) * num;
8  }
```

Fig. 5 Code snippet of modified kernel.

Methodology: To perform the tests, we use the same methodology from the original application (see Subsect. 4.1), i.e., with unmodified kernel.

Experimental Results: Fig. 6 shows the normalized runtime for the application in Situation 2. In comparison with Fig. 3, which represents the application execution time in Situation 1, we can observe that the relative time consumed by the multiplication range increased in all tests. Moreover, in all other observed ranges (initialization, HtoD transfer, DtoH transfer and sum), the spent runtime was the same. In all tests, zero-copy memory has the lowest performance. Whereas, the standard transfer method with pinned memory is the most efficient.

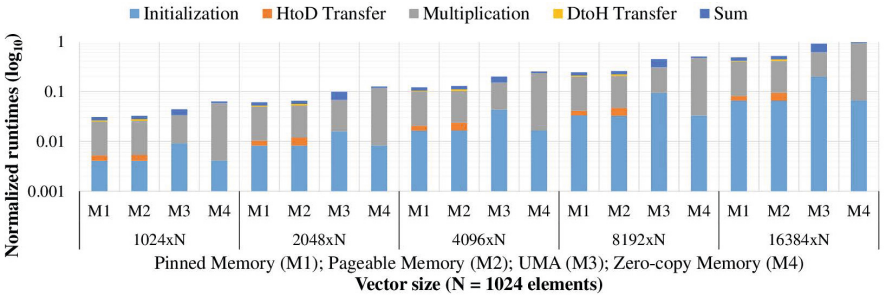


Fig. 6 Global normalized runtime for the application with kernel adapted (Situation 2).

The ratio of the multiplication range runtime with both modified kernel and unmodified kernel is shown in Figure 7. From this figure, we can see that the standard transfer method increases of about 20 times the runtime. Whereas the runtime increase in zero-copy memory was at least 40 times. Through NVIDIA Visual Profiler, we can observe in zero-copy memory, the kernel runtime increase is accompanied by an increase in the number of host memory access transactions. In all tests, the number of access transactions increased 50 times, which corresponds to the increased number of instructions in the modified kernel.

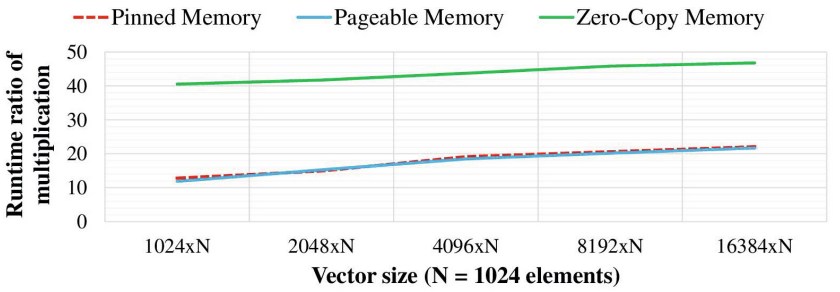


Fig. 7 Runtime ratio of multiplication range. Ratio between runtime multiplication in Situation 2 and Situation 1

Discussion: The tests demonstrated that an increase in access to data during kernel execution affects the migration of memory between the host and the device when using the zero-copy memory method. Additionally, we can observe that there is not a self optimization CUDA API when using zero-copy memory, once the kernel modification aims to increase the amount of memory access transactions and leads to the execution of redundant instructions.

5 Limitations of Zero-copy Memory Usage

The first limitation on the use of zero-copy memory resides in the fact that the data is allocated in a pinned memory. The allocation of large amounts of pinned memory can affect the operating system performance [10]. Unfortunately, it is not possible to measure a precise relationship between the amount of memory allocated and total memory installed in the system. Beyond the amount available memory, the operating system and other applications that are running in the environment influence the performance loss of the entire system. On this way, a good practice is not to use zero-copy memory when it is not known in advance, the maximum amount of data that will be allocated. As discussed in Subsect. 4.2, a problem in the use of *zero-copy* memory is the occurrence of performance loss when the amount of memory access transactions increases. Particularly, part of these transactions may include redundant copies performed on a kernel. In some cases, it can use a local variable within the kernel function that receives a copy of the data used for to avoid the redundant accesses. **Basically, the zero-copy memory must be used on data undergoing a lesser amount of access during transactions execution**. In case of more amount of access transactions, other transfer methods are recommendable.

6 Conclusion

Zero-copy memory presents implicit and transparent memory copies to the programmer, hiding the complexity of managing the data migration. Originally, this method was conceived as a feature that allows the use of data sets that may not be entirely stored in the memory device.

This study showed that in cases in which the kernel function performs a small amount of memory access transactions (in particular, a single access transaction), zero-copy memory can be used to provide performance increase. In certain cases, the use of zero-copy memory can provide a performance gain of more than 19% in the runtime application when compared to pinned memory. Based on the obtained results, we demonstrated that the total number of memory access transactions during execution of the kernel reduces the overall performance of the application and establishing a barrier in using zero-copy memory.

In the tested situations, we do not use multiple streams and memory copy process was not overlapped by the running kernel. Therefore, further works may include analysis of the performance of the zero-copy memory in a concurrent streams scenario. In a other future study, we will investigate the performance of zero-copy memory in other models of GPUs and also in the OpenCL API.

Acknowledgment This work was supported by Brazilian financial agency São Paulo Research Foundation (FAPESP) – grants 2015/00622-7 and 2013/03452-0.

References

1. Bai, T., Davis, S., Li, J., Jiang, H.: Analysis and acceleration of ntru lattice-based cryptographic system. In: 2014 15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 1–6, June 2014
2. Fatica, M.: Accelerating linpack with cuda on heterogenous clusters. In: Proceedings of 2Nd Workshop on General Purpose Processing on Graphics Processing Units,GPGPU-2, pp. 46–51. ACM, New York (2009)
3. Hennessy, J.L., Patterson, D.A.: Computer Architecture, Fifth Edition: A Quantitative Approach, 5th edn. Morgan Kaufmann Publishers Inc., San Francisco (2011)
4. Kaldewey, T., Lohman, G., Mueller, R., Volk, P.: Gpu join processing revisited. In: Proceedings of the Eighth International Workshop on Data Management on New Hardware, DaMoN 2012, pp. 55–62. ACM, New York (2012)
5. Kim, Y., Shrivastava, A.: Memory performance estimation of cuda programs. ACM Trans. Embed. Comput. Syst. **13**(2), 21:1–21:22 (2013)
6. Kirk, D.B., Hwu, W.M.W.: Programming Massively Parallel Processors: A Hands-on Approach, 1st edn. Morgan Kaufmann Publishers Inc., San Francisco (2010)
7. Landaverde, R., Zhang, T., Coskun, A., Herbordt, M.: An investigation of unified memory access performance in cuda. In: High Performance Extreme Computing Conference (HPEC), 2014 IEEE, pp. 1–6, September 2014

8. Li, W., Jin, G., Cui, X., See, S.: An evaluation of unified memory technology on nvidia gpus. In: 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), pp. 1092–1098, May 2015
9. Tang, K., Yu, Y., Wang, Y., Zhou, Y., Guo, H.: Ema: Turning multiple address spaces transparent to cuda programming. In: ChinaGrid Annual Conference (ChinaGrid), 2012 Seventh, pp. 170–175, September 2012
10. NVIDIA Corporation: CUDA C Best Practices Guide, March 2015
11. NVIDIA Corporation: CUDA C Programming Guide, March 2015

Design of a Deadlock-Free XY-YX Router for Network-on-Chip

Sang Muk Lee, Eun Nu Ri Ko, Young Seob Jeong and Seung Eun Lee

Abstract With the increasing number of cores in multiprocessor System-on-Chip (SoC), the design of an efficient communication fabric is essential to satisfy the bandwidth requirements of multiprocessor systems. Nowadays, scalable Network-on-Chips (NoC) are becoming the standard communication framework to replace bus-based networks. In this paper, we propose a deadlock-free XY-YX router for on-chip networks. In order to prevent deadlocks, we exploit additional physical channels in horizontal direction and optimize the priority of output channel allocation. Experimental results prove that the proposed deadlock-free router enhances the throughput of NoC.

Keywords Network-on-Chip (NoC) · Interconnection network · Router algorithm · Router design

1 Introduction

Future integrated systems will contain billions of transistors [1], composing tens to hundreds of IP cores. These IP cores, which support emerging complex multimedia and network applications, should be able to engender various multimedia and networking services. An efficient cooperation among these IP cores can be achieved through innovations of on-chip communication strategies [2, 3, 4]. Usually, buses that are fine-crafted for supporting a specific processor are successfully implemented in

S.M. Lee · E.N.R. Ko · S.E. Lee(✉)
Department of Electronic Engineering,
Seoul National University of Science and Technology, Seoul, Korea
e-mail: seung.lee@seoultech.ac.kr

Y.S. Jeong
Technology Division R&D Center, Technology Leaders & Innovators Inc.,
Seongnam, Korea

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_61

virtually all complex System on Chip (SoC) designs. However, the bus cannot adapt to changes in the system architecture especially because of the ever-growing complexity of single-die multiprocessor systems. For effective communication, conventional bus-based architectures have been replaced with network architectures that use a networking technology to establish effective communication among chip multiprocessors and large-scale IPs. The Network-on-Chip (NoC) layered approach has advantages of 1) better electrical properties, 2) higher bandwidth, 3) energy efficiency, 4) scalability, and the key point is the separation of IP design and functionality from chip communication requirements and interfacing.

While designing an NoC architecture, a wide variety of topologies that could comprise a network and routing algorithm to select a path, from a source to a destination, for effective communication should be considered [5]. Conventional XY routing, which is often used in torus and mesh topology, is deadlock- and livelock-free and is easy to design too. However, it has a drawback of uneven traffic distribution during communication, resulting in congestion. Therefore, modified XY routing techniques need to balance the load in order to improve the performance of the router. We proposed a deadlock-free XY-YX router architecture [6]. The main idea is the use of XY routing for a request packet and YX routing for a response packet. In this paper, we present the design details of the deadlock-free XY-YX router and verify the performance and feasibility of our router, using 65-nm CMOS technology. Experimental results demonstrate that our router improves the throughput of network compared with conventional XY routing.

The organization of this paper is as follows: Section 2 provides the related work regarding the NoC. Section 3 introduces the proposed XY-YX router. Section 4 presents the implementation details, demonstrating the feasibility of our router. Finally, Section 5 concludes this paper.

2 Related Works

As a new SoC design paradigm, NoC has been proposed to support the trend for SoC integration. In designing NoC systems, the following issues must to be considered: topologies, routing algorithms, performance, latency, and complexity. As there are topological variations, many researchers have proposed various interconnected architectures such as SPIN [7], CLICHE [8], 2D torus [9], OCTAGON [10], and BFT [11]. Considering NoC systems, the mesh is more practical because of its modularity; it can be easily expanded by adding new nodes and links without any modification of the existing node structure. As the mesh nodes can be used as basic components in on-chip communication, the mesh nodes are potentially important components to accomplish a scalable communication model in NoC environments [12].

Other challenges in NoC design are the routing algorithm and switching techniques. A routing algorithm has to maximize the throughput and minimize the complexity of the router implementation [13, 14, 15]. To determine a path among

the set of possible paths from source to destination, the routing algorithms are classified as deterministic/oblivious and adaptive ones [16]. The deterministic/oblivious routing algorithms choose a route without considering any information about the network's present condition. Adaptive routing algorithms use the state of the network like the status of a node or link, the status of buffers for network resources, or history of channel load information.

The deterministic/oblivious and adaptive routing algorithms, which are simple and easy to analyze, have been proposed. Dimension-ordered routing (DOR) is an algorithm that is suitable for meshes [17]. Unfortunately, as a DOR algorithm has no routing flexibility, it is sensitive to topology changes worst-case throughput was not enhanced. Class-based-Deterministic Routing (CDR) takes advantage of both XY and YX routing [18, 19]. In [20, 21], the path of CDR is deterministic by message class. Memory request packets use XY routing, which is compared against other routing algorithms. Valiant and Brebner [20, 21] proposed two-phase randomized routing in which a random node is used as an intermediate destination. By increasing routing flexibility, the worst-case throughput is improved. However, the average-case throughput was not enhanced and the number of network hops was increased. In [22], Ted Nesson *et. al.* presented the behavior of Randomized Oblivious Multi-phase Minimal (ROMM) routing. ROMM has the benefit of minimum number of network hops, routing flexibility, and good average throughput. Their simulation showed that the ROMM algorithm can perform better than the fully randomized routing and DOR on wormhole routed mesh and torus networks. However, the worst-case throughput was not improved. In order to overcome performance degradation of wormhole router, the use of multiple clocks to implement an adaptive wormhole router was investigated [23]. They reduced latency, increased throughput, and optimized power consumption. In [24, 25], Xiaohui Li *et. al.* proposed modified deterministic XY routing technique.

XY-YX routing (O1TURN, Orthogonal one-turn routing) [26] is comparable to DOR (X-first Y-next OR Y-first X-next) in terms of router implementation complexity, and it allows packets to traverse through one of the two possible dimension-ordered routes, which are different from DOR. While XY-YX routing is much more restrictive than ROMM considering the number of potential routes, the limited amount of routing flexibility allowed by XY-YX routing is sufficient to match the average-case throughput achieved by ROMM. Moreover, in terms of latency, XY-YX routing guarantees the minimum number of network hops. Research combining XY routing and YX routing has been performed [24, 27, 28]. In [24], D. Seo *et. al* tried to balance the XY and YX traffic in order to improve the performance of the XY-YX router. However, splitting the traffic evenly between XY and YX routes in toggle XY (TXY) routing does not always guarantee optimal load balancing. Moreover, TXY leads to out-of-order arrivals, requiring large re-order buffers. In order to insure the in-order arrivals, R. Gindin *et. al* proposed the weighted ordered toggle algorithm that assigns XY and YX routes to source-destination pairs, which reduces the maximum network capacity for the given traffic pattern [28]. In [28], the routes are calculated by using the information of the actual traffic pattern.

Algorithms in [24] and [28] try to balance the traffic of XY and YX packets to maintain the optimal load balancing in XY-YX routing in order to improve the performance of the router. The objective of these studies is to determine the ratio of the two optimal routing techniques considering the traffic circumstances in order to obtain better performance when implementing the combined XY-YX routing algorithm. An adaptive mode technique that depends on the circumstances of the congested network has been proposed [29, 30].

3 XY-YX Router Architecture

3.1 XY-YX Routing

The XY-YX routing can be used in applications owing to the benefits that it offers such as performance improvement and power reduction by predicting response packet paths in advance and eliminating wake up delays [28]. However, in the XY-YX combined routing networks, the performance varies depending on the degree of the traffic congestion. In other words, efforts to adjust the transmission ratios of the XY and YX packets are required in order to combine the two routing techniques. Moreover, deadlocks that occur when two routing techniques are combined have to be eliminated from the network. Algorithms in [24] and [28] try to control the transmission ratios of XY and YX packets to maintain the optimal load balancing. This requires multiple virtual channels per physical channel to solve the deadlock.

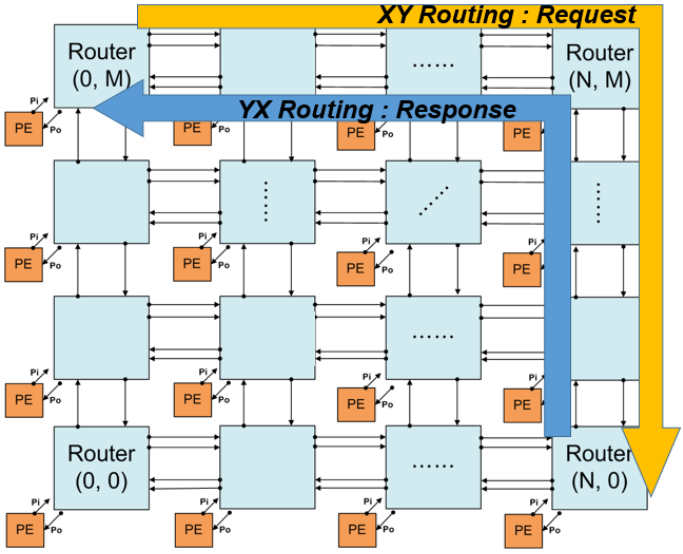


Fig. 1 The proposed XY-YX routing on mesh topology

By adding two disjoint horizontal channels instead of using virtual channels, our proposal prevents the deadlock. Though our approach to provide deadlock freedom requires additional resources to build two physical channels, it can reduce the complexity in routing algorithm, which controls the multiplexing of virtual channels to escape deadlock situation when virtual channels are exploited for this purpose. The overhead of adding physical channels can counterbalance the cost of allocating virtual channel buffers and associated control logics. Therefore, our approach of providing physical channels in horizontal direction can be beneficial in its own way. The use of the horizontal channels is controlled by the direction of delivered data. Figure 1 shows example of the proposed XY-YX routing design on N × M mesh topology.

3.2 Deadlock-Free XY-YX Router Design

Figure 2 illustrates a detailed block diagram of the deadlock-free router. It is similar to a typical 2D mesh NoC router. The proposed router contains seven input and seven output channel ports corresponding to the neighboring directions and the processing element (PE). Each input channel has an input FIFO queue. There are two east and two west channel pairs reaching each router, providing two disjoint sub-networks for deadlock freedom. The proposed router consist of an arbiter, a crossbar switch, three architectural blocks named- XY Router, YX Router, and Internal Router, and supports wormhole flow control for packet transmission in a mesh topology. All incoming packets into the router have a 1-bit

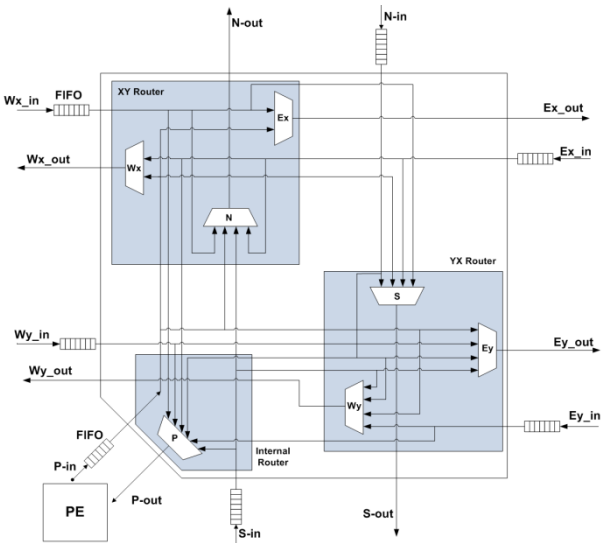


Fig. 2 Deadlock-free XY-YX router architecture

Table 1 Priority assignment on ports

Router	Output channel port	Input channel port
XY router	Ex-out	P-in, Wx-in
	N-out	P-in, S-in, Wx-in, Ex-in
	Wx-out	P-in, Ex-in
YX Router	Ey-out	P-in, N-in, S-in, Wy-in
	Wy-out	P-in, N-in, S-in, Ey-in
	S-out	P-in, N-in, Wx-in, Ex-in
Internal Router	P-out	N-in, S-in, Ex-in, Ey-in, Wx-in, Wy-in

signal to determine whether XY or YX routing is applied. The arbiter allocates outgoing channel for incoming packets and controls the crossbar switch through signals. The XY Router serves the channel set {Wx, Ex, S-in, and N-out} and the YX Router serves the channel set {Wy, Ey, S-out, and N-in}. Each horizontal channel is exclusively used depending on XY- or YX-routing algorithms of the delivered packets. To distinguish their functioning, each horizontal channel is denoted by Wx/Ex for XY-routing and Wy/Ey for YX-routing. The internal router supports the additional interface of a PE.

We assume that a packet coming through an input port does not loop back; thus, each input port is connected to the corresponding output port except for itself in the associated channel set. For the outgoing channel allocation, the router applies a fixed priority scheme to reduce complexity. Initially, we applied the output selection policy, which is a descending order of priority, in a clock-wise direction for each outgoing channel to incoming channels (referred to as origin hereafter). However, we have observed that vertical channels suffer the high congestion because of the additional channels in horizontal direction. Therefore, the output selection policy is modified to have higher priority in vertical channels than in horizontal channels (referred to as modified hereafter), thereby forwarding packets coming from the North and South, first. In order to prevent a certain packet from remaining in the input queue for a long time, we applied an aging counter to increase the priority of the packet.

4 Experimental Results

4.1 Physical Implementation

The XY-YX router is easy to be implemented based on an existing XY router and has higher performance with less overhead. In order to show the performance of a router and its feasibility for VLSI implementation, the XY-YX router was implemented using Verilog™ HDL, and a logic description of our router was obtained using the Synopsys™ with 65-nm CMOS technology. The Synopsys™

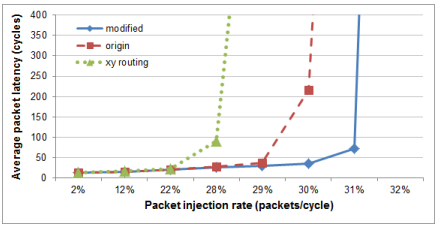
tool chain provided critical path information for a single router up to 537 MHz. The overall router including the input FIFOs with size 8 occupies an area of approximately 0.45 mm^2 using the 65-nm CMOS technology. Table 2 summarizes the physical characteristics of routers and processor. If the router is integrated within a chip multi-processor as an interconnection network, the area imposed by the network will be suitable for indicating the feasibility.

Table 2 Physical characteristics of the routers and processor

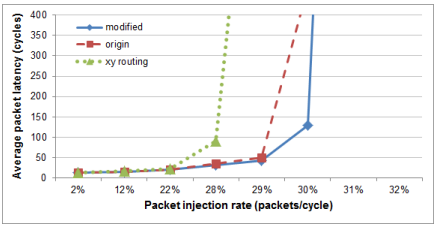
	XY router	**The proposed XY-YX router**
Maximum Frequency	534 MHz	537 MHz
Area	0.29 mm^2	0.45 mm^2
Dynamic Power @ 500 MHz	5.954 mW	9.2621 mW

4.2 Simulation Results

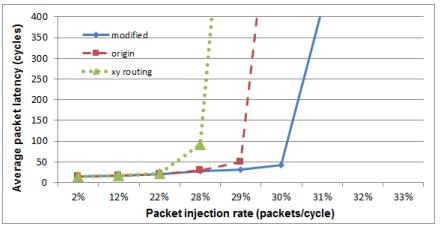
In this simulation, we evaluate a network of deadlock-free routers connected in the 8 × 8 2-D mesh topology. For the measurement of throughput and average packet latency, we adopted a standard interconnection network measurement setup where the packet generation is placed in front of an infinite depth source queue, and an input timing of each packet is measured whenever it is generated. The packet used in simulation is a randomly synthetic data. The packet length is fixed to eight flits (one head flit, six body flits, and one tail flit) even though the packet format supports various packets sizes. Figure 3 compares the average packet latency of three routing algorithm for different traffic loads. Each graph includes three curves: the original priority of XY-YX routing (origin), the modified priority of XY-YX routing (modified), and the existing XY routing. Each graph represents the offered traffic (flit/node/cycle) on X-axis and the average latency (cycles) on Y-axis. Multi-flit packets composed of eight 32-bit flits are injected into the network, and the performance metric considered is the average packet latency under different traffic loads. In our experiment, we calculated the average packet latency of the three routers by exploiting the various ratios of XY and YX packets in network. As shown in Figure 3, the origin router shows better performance than existing XY routing, and the modified router outperforms in terms of throughput for all traffic patterns, demonstrating the feasibility and superiority of our proposal. In addition, when packets are injected with same rate, the disproportionately XY-YX ratio simulation results (see Fig. 3(b), (e), (f), and (g)) indicate higher packet latency than the well-balanced XY-YX ratio simulation results (see Fig. 3(a), (c), and (d)) of the modified router.



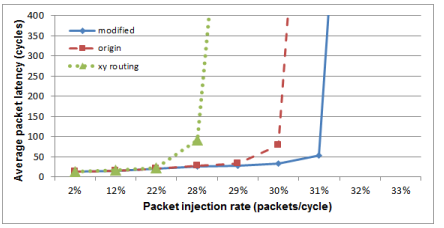
a. xy-50% and yx-50%



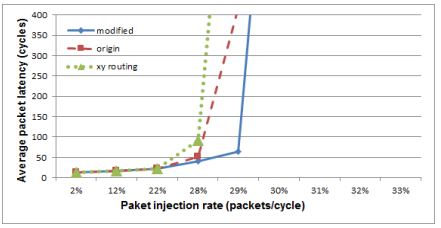
b. xy-10% and yx-90%



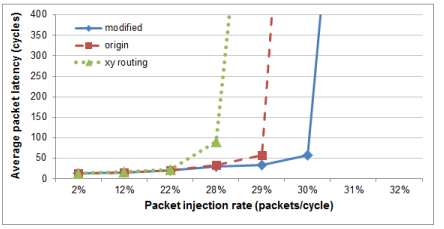
c. xy-60% and yx-40%



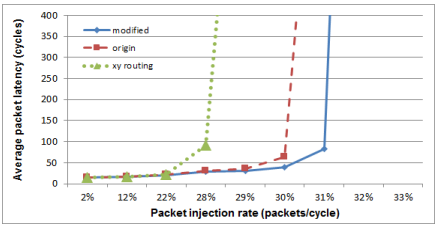
d. xy-40% and yx-60%



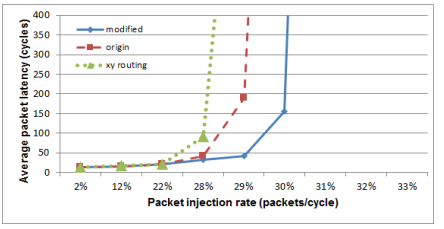
e. xy-90% and yx-10%



f. xy-70% and yx-30%



g. xy-20% and yx-80%



h. xy-80% and yx-20%

Fig. 3 The 8×8 mesh network comparison based on XY-YX ratios: xy-50% vs yx-50% (Fig. 2a), xy-10% vs yx-90% (Fig. 2b), xy-60% vs yx-40% (Fig. 2c), xy-40% vs yx-60% (Fig. 2d), xy-90% vs yx-10% (Fig. 2e), xy-70% vs yx-30% (Fig. 2f), xy-20% vs yx-80% (Fig. 2g), and xy-80% vs yx-20% (Fig. 2h)

5 Conclusion

In this paper, we proposed and implemented an XY-YX router architecture for flexible on-chip interconnection. The key idea is using XY and YX routing for different packets and avoiding deadlocks by adding two disjoint horizontal

channels instead of using virtual channels, thereby reducing the hardware cost to avoid the deadlock. The synthesis result in the 65-nm CMOS technology proved the feasibility of the proposed XY-YX router. The experimental results demonstrated the performance enhancement in terms of throughput, increasing the accepted load of NoC systems. In the future, we plan to explore the deadlock-free XY-YX router in FPGA and evaluate a packet latency of the deadlock-free router.

Acknowledgement This work was supported by National Research Foundation of Korea (NRF) Grant funded by Korean Government (MSIP) [NRF-2014R1A1A1004150 Development of Fault Injection Platform for Resilient System-on-Chip Design] and the IT R&D program of MSIP/IITP [R0101-15-0186 Development of open type Hadoop storage appliance to support more than 48 TB per single data node].

References

1. Millberg, M., Nilsson, E., Thid, R., Kumar, S., Jantsch, A.: The nostrum backbone-a communication protocol stack for networks on chip. In: 17th International Conference on VLSI Design, pp. 693–696 (2004)
2. Gebali, F., Elmiligi, H., El-Kharashi, M.W.: Networks-on-Chips: Theory and Practice. CRC Press (2011)
3. Lee, S.E., Bahn, J.H., Yang, Y.S., Bagherzadeh, N.: A generic network interface architecture for a NoC based multiprocessor SoC (NePA). LNCS, vol. 4934, pp. 247–260. Springer, Heidelberg (2008)
4. Lee, S.E., Bagherzadeh, N.: A high-level power model for network-on-chip (NoC) router. Computers & Electrical Engineering **35**(6), 837–845 (2009)
5. Lee, S.E., Bagherzadeh, N.: Increasing the throughput of an adaptive router in network-on-chip (NoC). In: International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), pp. 82–87 (2006)
6. Jeong, Y.S., Lee, S.E.: Deadlock-free XY-YX router for on-chip interconnection network. IEICE Electronics Express **10**(20), 1–5 (2013)
7. Guerrier, P., Greiner, A.: A generic architecture for on-chip packet-switched interconnections. In: Proceeding of Design and Test in Europe (DATE), pp. 250–256 (2000)
8. Kumar, S., Jantsch, A., Soininen, J. Forsell, M., Millberg, M., Oberg, J., Tiensyrja, K., Hemani, A.: A network on chip architecture and design methodology. In: Proceeding of ISVLSI, pp. 117–124 (2002)
9. Dally, W.J., Towles, B.: Route packets, not wires: on-chip interconnection net-works. In: Proceeding of Design Automation Conference (DAC), pp. 683–689 (2001)
10. Karim, F., Nguyen, A., Dey, S.: An interconnect architecture for networking systems on chips. IEEE Micro **22**(5), 36–45 (2002)
11. Pande, P.P., Grecu, C., Ivanov, A., Saleh, R.: Design of a switch for network on chip applications. Proceeding of ISCAS **5**, 217–220 (2003)
12. Duato, J., Yalamanchili, S., Ni, L.M.: Interconnection Networks: An Engineering Approach. IEEE Computer Society Press (2003)
13. Lee, S. E., Bagherzadeh, N.: NePA: networked processor array for high performance computing. In: US-Korea Conference on Science, Technology, and Entrepreneurship, August 2008

14. Hu, W.H., Lee, S.E., Bagherzadeh, N.: DMesh: a diagonally-linked mesh network-on-chip architecture. In: Int'l Workshop on Network-on-Chip Architectures (NoCArc), pp. 14–20 (2008)
15. Wang, C., Hu, W.H., Lee, S.E., Bagherzadeh, N.: Area and power-efficient innovative congestion-aware network-on-chip architecture. Journal of Systems Architecture **57**(1), 24–38 (2011)
16. Dally, W.J.: Towles B: Principles and Practices of Interconnection Networks. Morgan Kaufmann Publishers, San Francisco (2004)
17. Sullivan, H., Bashkow, T.R.: A large scale, homogeneous, fully distributed parallel machine, I. In: Proceedings of the 4th Annual Symposium On Computer Architecture, ISCA 1977, pp. 105–117 (1977)
18. Abts, D., Jerger, N.D.E., Kim, J., Gibson, D., Lipasti, M.H.: Achieving predictable performance through better memory controller placement in many-core CMPs. In: Proceedings of the 36th Annual International Symposium on Computer Architecture, ISCA 2009, pp. 451–461 (2009)
19. Fang, A., Hallnor, E.G., Li, B., Leddige, M., Dai, D., LEE, S.E., Srihari, M., Iyer, R.: Boomerang: reducing power consumption of response packets in NoCs with minimal performance impact. IEEE Computer Architecture Letters **9**(2), July–December 2010
20. Valiant, L.G.: A scheme for fast parallel communication. SIAM Journal on Computing **11**(2), 305–361 (1982)
21. Valiant, L.G., Brebner, G.J.: Universal schemes for parallel communication. In: Proceedings of the Thirteenth Annual ACM Symposium on Theory of Computing, STOC 1981 pp. 263–277 (1981)
22. Nesson, T., Johnson, S.L.: ROMM routing on mesh and tours networks. In: Proceedings of the Seventh Annual ACM Symposium on Parallel Algorithms and Architectures, SPAA 1995, pp. 275–287 (1995)
23. Lee, S.E., Bagherzadeh, N.: A variable frequency link for a power-aware network-on-chip (NoC). Integration-the VLSI Journal **42**(4), 479–485 (2009)
24. Seo, D., Ali, A., Lim, W.T., Rafique, N., Thottethodi, M.: Near-optimal worst-case throughput routing for two-dimensional mesh networks. In: Proceedings of the 32nd annual international symposium on Computer Architecture, ISCA 2005, pp. 432–443 (2005)
25. Li, X., Cao, Y., Wang, L., Cai, T.: Fault-tolerant routing algorithm for network-on-chip based on dynamic xy routing. Wuhan University Journal of Natural Sciences **14**(4), 343–348 (2009)
26. Bahn, J.H., Lee, S.E., Yang, Y.S., Yang, J.S., Bagherzadeh, N.: On design and application mapping of a network-on-chip (NoC) Architecture. Parallel Processing Letters (PPL) **18**(2), 239–255 (2008)
27. Fang, Z., Hallnor, E.G., Li, B., Leddige, M., Dai, D., Lee, S.E., Makineni, S., Iyer, R.: Boomerang: reducing power consumption of response packets in nocs with minimal performance impact. IEEE Computer Architecture letters **9**(2), July 2010
28. Gindin, R., Cidon, I., Keidar, I.: NoC-based FPGA: architecture and routing. In: First International Symposium on Networks-on-Chips (NOCS), pp. 253–264 (2007)
29. Ghosal, P., Das, T.S.: Network-on-chip routing using structural diametrical 2D mesh architecture. In: Emerging Applications of Information Technology (EAIT), Kolkata, November 2012, pp. 471–474 (2012)
30. Lee, S.E., Bahn, J.H., Bagherzadeh, N.: Design of a feasible on-chip interconnection network for a chip multiprocessor (CMP). In: International Symposium on Computer Architecture and High Performance Computing, pp. 211–218 (2007)

An FPGA Based Compression Accelerator for Forex Trading System

Ji Hoon Jang, Seong Mo Lee, Oh Seong Gwon and Seung Eun Lee

Abstract In this paper, we propose an FPGA based hardware accelerator for forex trading system. In the forex trading market, the trading volume of currencies is growing larger every year. In order to provide a real-time processing of large volume and high availability service, we focused on the two types of workload, where a bottleneck occurs. The bottleneck between an application server and an internal hard disk is caused by the overhead from storing the transaction logs, due to the bandwidth limitation of a hard disk. Our key idea is to suppress the overhead of transaction logging through the high throughput hardware compression. Compared to software compression, our hardware accelerator scored 6x better performance in compression throughput.

Keywords FPGA · Hardware accelerator · Compression · Forex trading

1 Introduction

In the forex trading market (a.k.a. foreign exchange trading market), high-frequency trading has been increasing. Therefore, conventional forex trading system needs the development of real-time processing technique to process the transaction in real-time. We conducted workload analysis of forex trading system with a financial solution company managing the forex trading system. Fig. 1 shows the conventional forex trading system. The forex trading system has two bottlenecks. The bottleneck between the DB server and the storage is caused by frequent updates of the specific data. The result of the workload analysis shows that 163876 SQL queries are executed during 6 minutes. From among these, 31410 SQL queries (19.17%) are repeatedly updating the specific data into the

J.H. Jang · S.M. Lee · O.S. Gwon · S.E. Lee(✉)
Department of Electronic Engineering, Seoul National University of Science and Technology, 232 Gongneung gil, Nowon-gu, Seoul 139-743, Korea
e-mail: seung.lee@seoultech.ac.kr

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_62

storage. This specific data indicates stock quotes, state of the stock exchange, or other real-time information of forex trading. In our previous work, we proposed the novel methods to solve this bottleneck [1]. Another bottleneck between the application server and the internal hard disk is caused by the overhead of transaction logging. In the financial trading system, the transaction logging is mandatory. However, in the case of high-frequency trading, the transaction logging causes the bottleneck due to the transmission bandwidth limitation of hard disk. In order to suppress the overhead, we compress the transaction logs for lower bandwidth usage. Thanks to the transaction log that is a text data, we can achieve a high compression ratio and a real-time processing. However, software utilities are often slower than hardware solution [2, 3]. The hardware acceleration is a promising way to increase the throughput of compression application [4]. An FPGA facilitates hardware parallelism and this feature is suitable for processing of big data [5]. Therefore, in order to achieve high throughput compression, we adopted the PCIe based FPGA for the implementation of compression algorithm. In this paper, we propose an FPGA based compression accelerator for forex trading system. Our compression accelerator provides computation off-loading to application server and improves the data processing performance. Furthermore, hardware accelerators can improve the response latency and energy efficiency [6].

The remainder of this paper is organized as follows: In section 2, we introduce the background knowledge about LZ4 software compression and workload analysis results. Next, section 3 describes the system architecture of our compression accelerator. In section 4, we detail the micro-architecture of LZ4 compression core. Section 5 provides experimental results. Finally, we conclude in section 6 by proposing the future work.

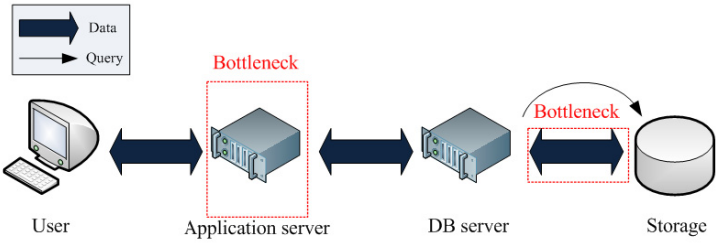


Fig. 1 The conventional forex trading system

2 Background

2.1 LZ4 Software Compression

LZ4 compression algorithm is adopted because LZ4 is the fastest scheme among the Lempel Ziv-family [7]. In [7], according to open source benchmark result,

LZ4 is a little faster than LZO, snappy, and 5-times faster than zlib (LZ77). We explored the LZ4 software compression throughput on an ARM Cortex-A9 processor in FPGA and analyzed the workload in order to use the hardware parallelism. Fig. 2 depicts the system architecture of our software LZ4 test-bed. The software LZ4 test-bed is composed of PCIe endpoint IP, AXI4-stream to EBI interface module, buffer, AXI4-lite interface module, AXI4 memory mapped interconnect, and Cortex-A9 processor. The PCIe endpoint IP, which is responsible to physical layer, data link layer, and transaction layer of PCIe protocol, has AXI4-Stream interface as user interface and physical interface with x4 lane between other PCIe system and software LZ4 test-bed.

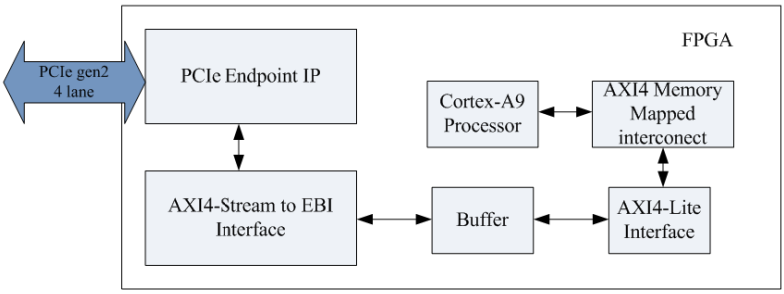


Fig. 2 System Architecture of Software LZ4 test-bed

The AXI4-Stream to EBI interface module extracts the data of transaction layer packet (TLP) received from PCIe endpoint IP and the data is fed to buffer. Cortex-A9 processor is interconnected to the buffer through AXI4-lite protocol. LZ4 software is running on the Cortex-A9 processor and compresses the transaction logs. In order to measure the throughput of software LZ4 test-bed, we used SCU (Snoop Control Unit) timer built in Cortex-A9 on an FPGA. We set the parameters as follows. Timer load value is 666, prescaler value is 0, and operating clock frequency is 667MHz. The transaction logs had been offered from financial solution company. Table I shows the compression result of software LZ4 test-bed. In the compression result, the compression throughput is roughly 200Mbps [8].

Table 1 Compression Result of Software LZ4

Transaction Log size (Bytes)	Compressed Log size (Bytes)	#Ticks	Throughput (Mbps)
3424	1246	189	138.217
3413	654	113	230.434
3398	1377	203	127.707
3387	470	86	300.473
3382	588	104	248.102

2.2 *Workload Analysis of LZ4 Software*

We conducted workload analysis of the LZ4 software in order to use the hardware parallelism. During the LZ4 compression, the redundant data can be substituted by 3-byte LZ4 data composed of token, match length, and offset. In order to find the interesting point that is causing the performance degradation, we analyzed and modified the LZ4 source code, where LZ4 source code is grouped into the blocks according to its functionality. Fig. 3 shows the proportion of executed blocks during LZ4 software compression. The `catch_up` block finds the first matching between the redundant data and plain text.

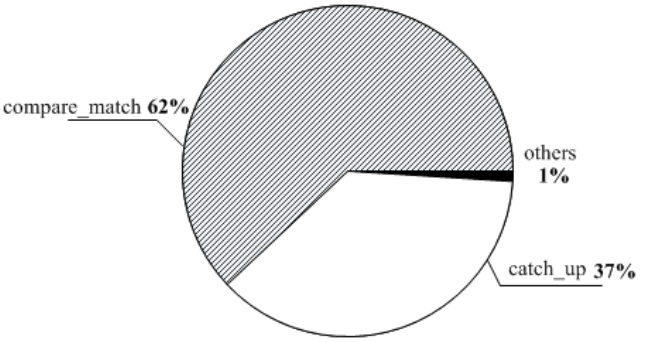


Fig. 3 Proportion of executed blocks during LZ4 software compression

The `compare_match` block compares the redundant data to plain text to calculate match length. Both blocks occupy the 99% of whole process, where the `compare_match` block uses a byte-aligned memory access. We concluded that these blocks need an acceleration by using a hardware parallelism.

3 System Architecture

Fig 4. shows the system architecture of compression accelerator. Compared to software LZ4 test-bed, Cortex-A9 processor is replaced with the LZ4 compression IP. Two compression cores and a router are integrated in LZ4 compression IP. The compression core is designed as the fully pipelined architecture, where the iteration is completed during 20 cycles and compresses the 16 bytes (parallelization size). Thanks to our stall-free pipelined architecture, we expected the throughput performance of compression accelerator as follows: total throughput = 16 Bytes / (20 cycles * period) * 2 cores. Fig. 5 shows the pipeline stages of LZ4 compression core. An operation of each stage is as follows:

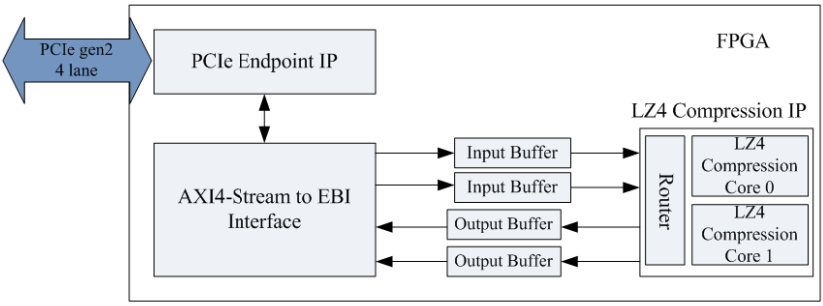


Fig. 4 System Architecture of Compression Accelerator

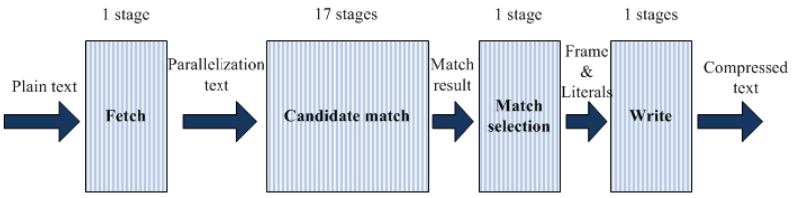


Fig. 5 Pipeline Stages of LZ4 Compression Core

- Fetch: The parallelization text (processed text at one iteration in parallel) is slid into the current window from the lookahead window, where the current window indicates the text processed in current iteration and the lookahead window indicates the text processed in next iteration. The parallelization text is prepared from plain text while other stages are conducted.
- Candidate match: The parallelization text is compared with each dictionary for candidate matching, where 16 match lengths data are calculated.
- Match selection: In order to obtain the best compression ratio, the best match length is found among the 16 match results. LZ4 frame composed of the token, the literals, the match length, and the offset is built by using the best match result.
- Write: The compressed data is fed to the output buffer through the write logic, which has an extra FSM. According to plain text, the execution time of the write logic is varied up to 20 clock cycles.

4 Micro-Architecture

Fig. 6 shows the micro-architecture of a compression core named `lz4_compress`. The `lz4_compress` core is composed of 16 `dictionaries`, an `allocator`, an `fsm`, a `manager`, a `position`, a `compare_match`, and a `compressed_data_write` module.

To parallelize the LZ4 algorithm, we designed the proper dictionary architecture for a hardware implementation. In the LZ4 software in our workload analysis, dictionary finds a first matching between the inside dictionary text and window text. In addition, dictionary provides the relative position between the window text and literals in order to calculate match length. The simple hash function can be used for reducing the depth of memory, where the simple hash function returns the short bucket bits. We used ASCII as a hash function, where the first character of parallelization text in current window is used. Therefore, the bucket size is 7-bit and the depth of memory is 128 because the bucket is used as an address of memory. Multiple dictionaries as the same number of parallelization text are used for parallelization [9]. The `allocator` module has a sliding window and performs the data allocation to other submodules. The `fsm` module

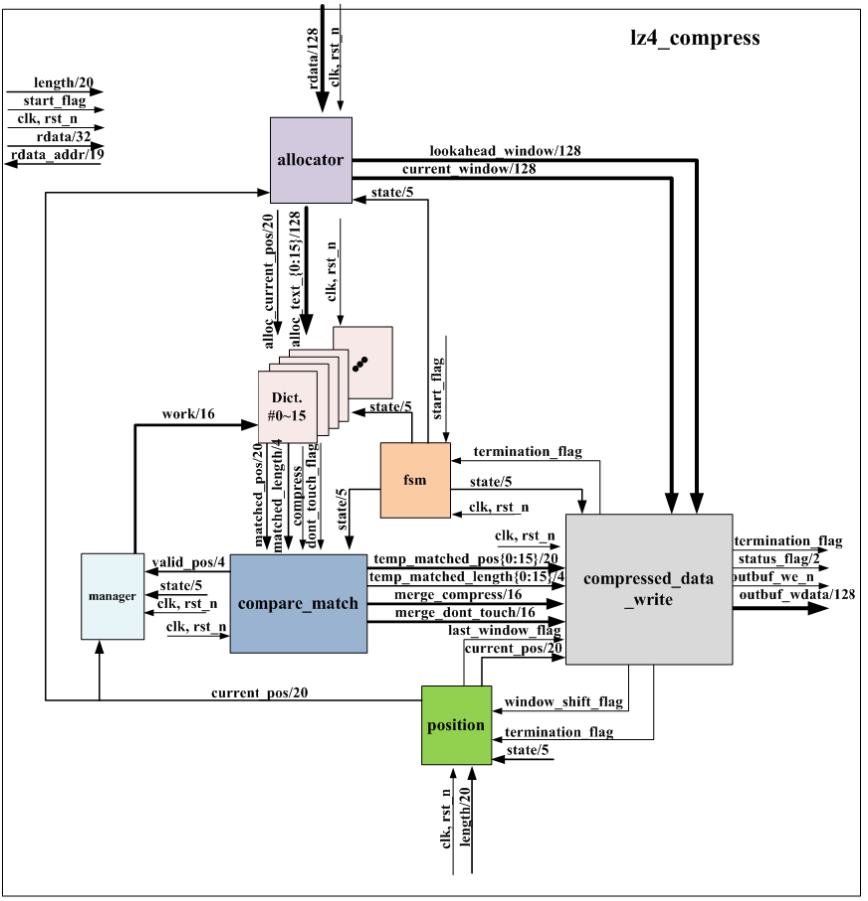


Fig. 6 Micro-Architecture of lz4_compress core

controls the pipeline operation and transmits the state to other submodules. The `manager` module controls the dictionary operation based on `valid_position` signal, where `valid_position` signal indicates the valid position on the parallelization text of current window because the text in lookahead window can be compressed in previous iteration. Therefore, the compressed text in previous iteration is invalid at the next iteration. The `position` module forwards a current position from the start position of plain text to other submodules. The `compare_match` module finds the best matched length among the 16 match lengths from dictionaries. The `compressed_data_write` module has the steering logic, where it builds the LZ4 data and writes compressed data to the output buffer.

5 Experimental Results

We realized LZ4 compression accelerator on the Xilinx ZC706 evaluation board with a Zynq-7000 (XV7Z045-2FFG900C). Xilinx PlanAhead tool was used to implement the compression accelerator targeting 100MHz operation. Table II shows the resource utilization of Zynq-7000 for the compression accelerator.

Table 2 Zynq-7000 Resource Utilization

Resource	Number used	Utilization
Register	41197	9%
LUT	65047	29%
Slice	22852	41%
IO	12	3%
RAMB36E1	72	6%
RAMB18E1	104	9%

To measure the compression time, we used a logic analyzer to detect physical signals indicating start and completion for LZ4 compression in real-time. The compression accelerator includes LZ4 compression IP, where two LZ4 compression cores are integrated. A 64KB transaction log is divided into two 32KB logs and fed to buffer. An application program on host server operates as follows: First, one of 32KB transaction log is written to an input buffer of LZ4 compression IP in accordance with the requested data length. After writing operation, the remainder transaction log is written to the other input buffer of LZ4 compression IP. Next, application program sets control register of each LZ4 compression core to request the compression start. When compression is completed, each LZ4 compression core writes compressed data to an output buffer respectively and sets the completion bit on control register. Finally, the application program reads the compressed data from the output buffer. Fig. 7 shows the experimental environment for estimating the compression throughput of our compression accelerator.

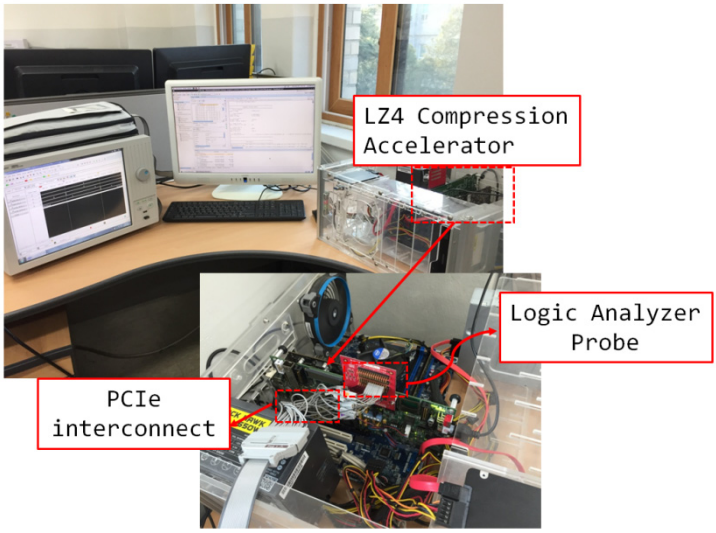


Fig. 7 The experimental environment

Compression throughput is measured through total transaction log size divided by compression time. In the experiment, total transaction log size was 64KB and the compression time was 411.528us, where we measured the time between the start request flag and completion flag by using a logic analyzer. Fig. 8 shows the completion flags of each LZ4 compression core on logic analyzer. Consequently, our LZ4 compression accelerator recorded the compression throughput of about 1.2Gbps. In terms of each LZ4 compression core, the compression throughput recorded of about 600Mbps. Compared to LZ4 software compression, our LZ4 compression accelerator has 6x better throughput performance thanks to hardware parallelism. Furthermore, the compression throughput can be improved through adoption of the many-core scheme, where hardware scheduler is needed.

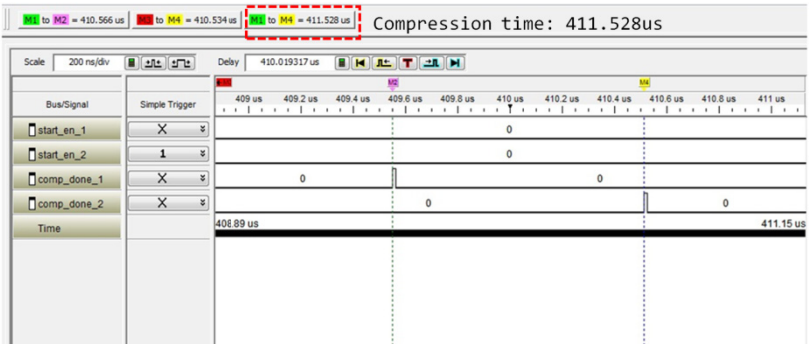


Fig. 8 Compression time measurement with logic analyzer

We integrated our compression accelerator in the application server of forex trading system in data centers to check the feasibility of our proposal. In high-frequency trading condition, our compression accelerator provided the compression off-loading to the application server and accelerated the transaction logging successfully. Fig. 9 shows the demonstration environment in the application server of forex trading system.

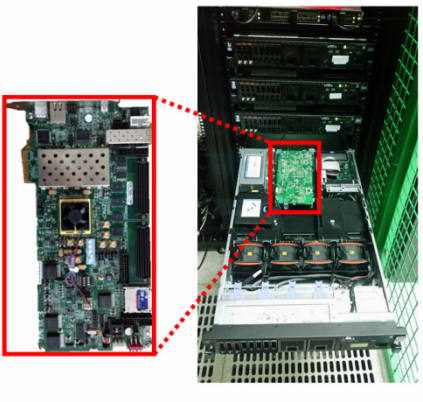


Fig. 9 The demonstration environment in the application server of forex trading system

6 Conclusion and Future Work

We proposed an FPGA based compression accelerator for enhancing the performance of forex trading system which needs a real-time processing technique to process high-frequency transaction. In the experimental results, our LZ4 compression accelerator recorded the 6x better throughput performance compared to software based compression. In addition, the compression throughput can be improved to integrate more LZ4 compression cores. We plan to integrated more compression cores along with a hardware scheduler in the hardware accelerator in order to achieve the high compression throughput. We expect that hardware acceleration of compression will bring forth a new spectrum of novel usage models for computing systems.

Acknowledgments This work was supported by the IT R&D program of MSIP/IITP [R0101-15-0186, Development of open type Hadoop storage appliance to support more than 48 TB per single data node] and [10043896, Development of virtual memory system on multi-server and application software to provide realtime processing of exponential transaction and high availability service].

References

1. Kim, S.J., Lee, S.M., Jang, J.H., Kim, S.D., Lee, S.E.: In-time transaction accelerator architecture for RDBMS. In: Advanced Technologies, Embedded and Multimedia for Human-centric Computing, pp. 329–334 Springer, Netherlands (2014)
2. Lee, S.E., Zhang S., Srinivasan, S., Fang, Z., Iyer, R., Newell D.: Accelerating mobile augmented reality on a handheld platform. In: IEEE Int'l Conf. on Computer Design (ICCD), pp. 419–426 (2009)
3. Lee, S.E., Min, K.W., Suh, T.W.: Accelerating Histograms of Oriented Gradients descriptor extraction for pedestrian recognition. Computers and Electrical Engineering **39**(4), 1043–1048 (2013)
4. Sukhwani, B., Abali, B., Brezzo, B., Asaad, S.: High-throughput, lossless data compresion on FPGAs. In: 19th Annual IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM), pp. 113–116 (2011)
5. Guha, R., Al-Dabass, D.: Performance prediction of parallel computation of streaming applications on FPGA platform. In: 12th International Conference on Computer Modelling and Simulation (UKSim), pp. 579–585 (2010)
6. Lyer, R., Sirinivasan, S., Tickoo, O., Fang, Z., LLLikkal, R., Zhang, S., Chadha, V., StillWell, P., Lee, S.E.: Cogniserve: Heterogeneous Server architecture for large-scale Recognition. IEEE Micro **3**, 20–31 (2011)
7. LZ4 algorithm. https://code.google.com/p/lz4
8. Jang, J.H., Lee, S.M., Kim, S.D., Gwon, O.S., Ko, E., Lee, S.M., Shin, J.W., Lee, S.E.: Accelerating forex trading system through transaction log compression. In: 2014 International SoC Design Conference (ISOCC), pp. 74–75 (2014)
9. Abdelfattah, M. S., Hagiescu, A., Singh, D.: Gzip on a chip: High performance lossless data compression on fpgas using opencl. In: Proceedings of the International Workshop on OpenCL 2013 & 2014, No. 4. ACM (2014)

Part VI
Agile Software Testing and Development

An Academic Case Study Using Scrum

Luciana Rinaldi Fogaça, Luiz Alberto Vieira Dias and Adilson Marques da Cunha

Abstract This paper aims to present an academic experience in which a real world software development problem can be tackled by four classes of students. Within this context, it was proposed by the professors the development of a project named SI-GAC, an acronym in Portuguese of what could be freely translated to English as "Integrated Accident and Crisis Management System". This project, with social community benefits, had as its main objective to develop an embedded software system for crisis management. The Scrum framework and Agile Methods were used, with the purpose of training the students in these technologies with a real world problem. Four sprints were exercised and it was possible to show that, with some adaptations, even for geographically dispersed teams the results achieved were satisfactory.

Keywords Component · Software engineering · Agile methods · Software development · Scrum · Scrum of Scrums

1 Introduction

At the present time, there is a growing expectation for software to be delivered as quickly as possible, without leaving aside quality. And this is one of the reasons why the Agile Methods have become so popular among software experts. "Promoted by senior software practitioners, agile methods were intended to avoid traditional engineering practices and rather focus on delivering working software as quickly as possible."[4]

Universities, in the role of supplying skilled labor to fulfill market`s demands, have to include practical activities in their courses, so that the students may have a better understanding of the concepts, tools and methods being used outside the academic world, such as the Agile Methods.

Taking into account this idea, during the second semester of 2015, encompassing both the undergraduate and graduate programs in Electrical and Computer

L.R. Fogaça(✉) · L.A.V. Dias · A.M. da Cunha
Computer Science Division, ITA - Instituto Tecnológico de Aeronáutica,
Sao Jose dos Campos, Brazil
e-mail: fogaca.lu@hotmail.com, {vdias,cunha}@ita.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_63

Engineering, option Informatics, it was adopted a case study using Scrum (the most widespread agile method [5]).

Within this context, it was proposed by the professors the development of a project named SI-GAC, an acronym in Portuguese of what could be freely translated to English as "Integrated Accident and Crisis Management System". A project from which the social community could benefit.

However, besides having this social orientation, the professors presented a challenge: to use SCADE, a tool developed by ESTEREL Technologies/Ansys, which is "used to design critical software, such as flight control and engine control systems, landing gear systems, automatic pilots, power and fuel management, cockpit displays, rail interlocking systems and signaling, automatic train operation, computer based train control, emergency braking systems, overspeed protection, train vacancy detection, nuclear power plant controls, and many other aerospace, railway, energy, automotive, or industrial applications."[1]

The project involved students from four different courses: Embedded Systems Project (CES-65), Real-Time Embedded Systems (CE-235), Software Quality, Reliability and Safety (CE-230), and Advanced Topics on Software Testing (CE-237).

This article focuses on the Scrum techniques applied to the academic project, as well as the experience gained from its implementation inside the classroom.

2 Methodology

2.1 Definition of the Teams

Considering the structure of a small agile team, as shown in Figure 1, the participating students of the four courses were divided into eight teams named TSXX, where XX represents the number of the team (from 01 to 08) and TS stands for Team Scrum.

Due to the relatively large size of the problem, it was used the Scrum of Scrums approach, with each individual team (about 10 people) using the Scrum for the management of its developments.

The 80+ students, from four different courses, and not knowing each other personally, had to organize themselves into the Team Scrums, each team composed by students of all the four courses.

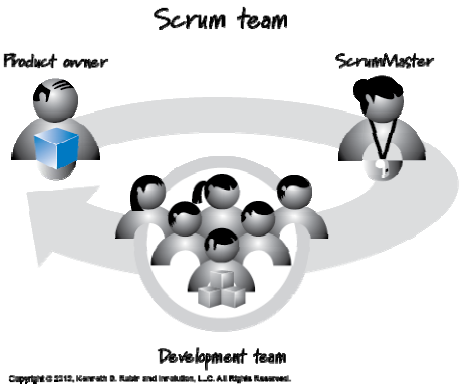


Fig. 1 Scrum Roles [2]

It was defined by the Stakeholders (professors) that the roles attributed to the students would be chosen according to the course they were attending: Embedded System Project (CES-65), Software Quality, Reliability and Safety (CE-230) and most of the Real-Time Embedded Systems (CE-235) students would constitute the Development Team. Some of the Real-Time Embedded Systems (CE-235) students would be the Scrum Masters and the Advanced Topics on Software Testing (CE-237) class would be the Product Owners and the Testers.

2.2 Project Architecture

In the context of the management of an accident or crisis, the architecture of the project was created, where the Scrum Teams (TS) have been grouped two by two into four segments. Figure 2 gives a better understanding of the responsibilities of each TS.

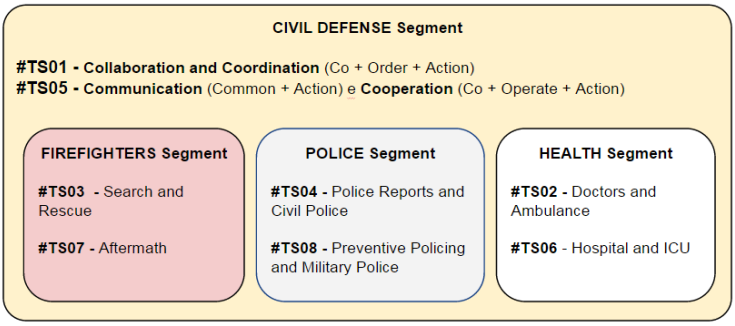


Fig. 2 Project Architecture

2.3 Project Schedule

As it usually happens in the corporate area, project schedules are very tight. In the SI-GAC it was not different, with only 17 weeks to complete the job. Considering the available timeframe, the project timeline defined may be seen in Figure 3.

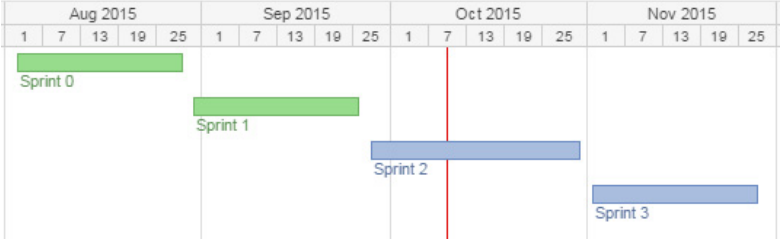


Fig. 3 Timeline of the project

Figure 3 shows that the whole time was divided into four regular, repeatable work cycles, known as sprints or iterations. “The heart of Scrum is a Sprint, a time-box of one month or less during which a “Done”, useable, and potentially releasable product Increment is created” [6]. Each sprint had about 4 weeks (one month).

2.4 Product Backlog

“The Product Backlog is an ordered list of everything that might be needed in the product and is the single source of requirements for any changes to be made to the product” [6]. It is composed of User Stories (US), a term coined in Extreme Programming by Kent Beck in 1999 [7]. A user story describes functionality that will be valuable to either a user or purchaser of a system software [8].

In the context of the SI-GAC project, the US were created based on the suggestions of all team integrants. On the other side, the definition of which US should have the highest priority and which one should have the lowest was a task of the product owners that, as mentioned before, were the students from the Advanced Topics on Software Testing (CE-237) class.

This product backlog was kept in an excel spreadsheet available to all the teams through a website created to keep all information concerning the project.

2.5 Sprint Backlog

Once all US have been written down and prioritized, some of them were selected to be part of the First Sprint Backlog.

It is important to notice that, when the project schedule has been defined, there was a sprint called “Sprint 0”. In this sprint, exceptionally, no sprint backlog has been defined. Sprint 0 was used for learning purposes, to level the knowledge of all students in two of the tools that were to be used during the semester: SCADE and the version control system for software development, GIT. Also during this Sprint, all students participated in a simulation called “LEGO from Scrum”. This simulation was very important for the students to understand the basic concepts of Scrum and be able to apply them to the project.

Then, officially, it can be said that the First Sprint was the one named “Sprint 1”. For following up of each sprint backlog, the students decided to adopt the free application named Trello. In Figure 4, it is possible to see how the sprint backlog was set inside Trello. The column PBI (Product Backlog Items) contains all US that the team expects to finish by the end of the sprint. The other columns (To do, In Progress and Done) contain smaller tasks that derived from the user stories and that are moved from “To Do” to “Done” as the project evolves and time passes by. The ideal is that each US is broken into smaller tasks that can be done within a maximum of 2 days so that the burn-down chart (a concept that will be clarified in the following sessions of this article) is constantly updated.

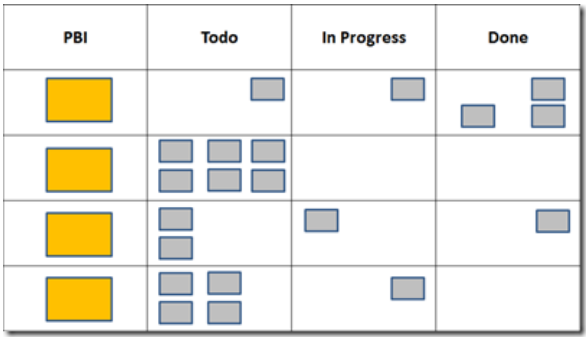


Fig. 4 Sprint Backlog [12]

2.6 Scrum Ceremonies

Meetings (or "ceremonies", as they are called inside Scrum) play an important role during the software development. They are essential to keep the team focused, working towards a common objective and to encourage interaction and communication between team members.

Considering the difficulty of gathering all team members from SI-GAC project together, these ceremonies happened, most of the time, online.

Here follows four of the ceremonies that are usually adopted by agile organizations (and that were adopted also by the SI-GAC participants):

- Sprint Planning:

During the Sprint Planning, the team decides which user stories will be part of the next sprint. In this meeting, it is also estimated the effort to complete the tasks. This can be done using the Planning Poker, a technique where each team member receives cards like the ones showed in Figure 5 and raises the one that best represents the effort of the team to complete each US. The higher the number, also higher the effort to finish the US. E.g. 100 would be a US that is almost impossible to do, one would be very easy. Zero and ½ are tasks done by inspection. At the end of the Sprint Planning, the team must reach a consensus on the story points and have as output the Sprint backlog.

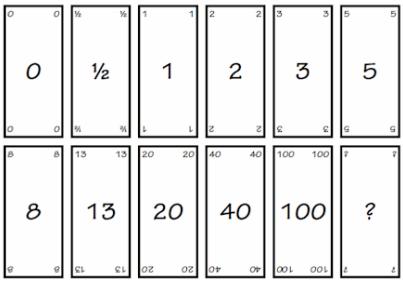


Fig. 5 Scrum cards for Planning Poker [13]

As mentioned before, it was not possible to have every team member physically present at these meetings. What was done in order not to skip the planning poker was to make everyone fill an online form with what was considered by them the necessary effort for each US. They also made suggestions on what would be the tasks for the sprint. Later, during an online meeting, and having everyone's answer to the form, it was easier and faster for the TS to reach the expected consensus.

- Sprint Review:

During the Sprint Review, the teams need to show their work to the stakeholders. They have to present their deliverables, after the sprint is over. In the case of project SI-GAC, the stakeholders were the professors so the work of the team was presented by the Product Owners in the format of a presentation containing a quick, narrated video (of max. 4 minutes). This happened during the class hours of the Advanced Topics on Software Testing (CE-237) class.

- Sprint Retrospective:

Sprint Retrospectives happen always at the end of each iteration and the idea behind them is to make improvements for the next iterations. In this meeting, the team has the chance to point out what worked well during the sprint and what could be better. In the context of SI-GAC, this ceremony also did not take place physically. A form was made so that the students could assess their team mates' contribution to the sprint. Besides that, it was possible to leave opinions regarding the previous work, what could be improved, what should not be changed because it was already working well.

This form provided subsidy to determine the students grades at the end of the semester, together with the professors evaluations of the sprint review presentations.

- Daily Scrum Meeting:

As the name suggests, the Daily Scrum Meeting is supposed to happen daily and last no longer than 15 minutes. However, within the project, it was not possible to do it every day. What was done instead was an online weekly meeting. Also, it lasted longer than 15 minutes, usually around 45. During this meeting, every member answered three questions:

— What did I do last week?
— What will I do in the next week?
— Am I blocked by anything?

Even though the meetings happened weekly, the team communicated through e-mail and whatsapp every time they found necessary.

2.7 Burn-Down Chart

"Burndown chart is an important measurement tool for planning and monitoring of progress in agile methods based on software working as an agile principle.

In most of methods and among many of teams, it is used for representation of amount of remained work" [9].

In the case of the SI-GAC project, the base for tracking the progress of the activities and updating the burn-down chart was the story points that the teams attributed during the sprint planning to each user story of the sprint backlog. The sum of the story points of the user stories selected to compose the sprint backlog represents the total points to be "burned" during the sprint. Each time a member of the team completes a task, he/she can "burn" points in Trello and consequently makes the red line of Figure 6 go down.

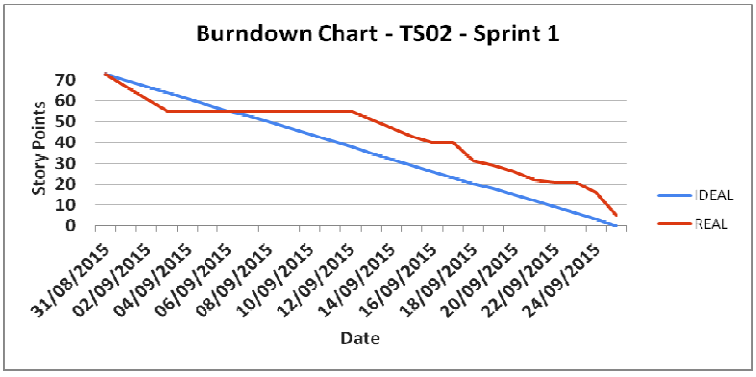


Fig. 6 Burn-down Chart of TS02 during Sprint 1

Figure 6 shows, for instance, that at the beginning of the sprint the students had a very good progress, burning more points than what was expected (evidenced by the ideal line). But for about one week, they stayed constant, not burning any points. At the end of the sprint it is possible to see that the line did not reach zero, showing that there have been some tasks left undone. In this case, two things could happen: the tasks are transferred to be finished in the next sprint (which was the choice of the team in this sprint) or the tasks are canceled and a good justification is given to the stakeholders of why this had to be done. The team may say, for example: "we noticed this task would require much more money and time than what was initially expected so we decided to cancel it. However, we have completed these others tasks instead, solving the same problem in a more efficient way."

3 Results

During the development of the SI-GAC project, it was possible to confirm what was stated by Schwaber in [3]:

"Scrum is not a prescriptive process; it doesn't describe what to do in every circumstance. Scrum is used for complex work in which it is impossible to predict

everything that will occur. Accordingly, Scrum simply offers a framework and set of practices that keep everything visible"

The students succeeded in the sense of keeping the information easily accessible to everyone, even though there was the impossibility of making face-to-face meetings, the main obstacle to communication.

Also, the team integration in the first sprint was even harder by the fact that most of the TS integrants had never personally met.

Constant online communication was the solution for this problem. Also, the support of tools like google sites, google forms, trello and others was essential for the teams to succeed.

At the end of this first sprint the students provided part of the intended system, with working software, developed with quality, reliability, safety, and testability. The students from Advanced Topics on Software Testing, besides being product owners played the role of testers, which was important to guarantee that the developed software fulfilled the requirements.

An important feature of the Agile Methods is that all the software has to be tested since the beginning of the development and has to be continually tested along its development life time until it is ready to be put into production. This process reduces considerably the risk of the development [10].

4 Conclusion

The main objective of this paper was fulfilled in the sense that the idea was to teach Agile Methods with Scrum through a case study and show that is possible to take to classroom a practical engineering problem and solve it, using a framework Scrum.

The authors wanted to stress that the objective was not to prepare an optimized working system, but to provide the students with an experience that will enable them to tackle real world problems in correlated areas, using Scrum.

Despite the fact that the final product is not expected to be complete, but an academic version, with real features, the same kind that the students will encounter in their professional life, the resulting software to be produced has to be quality oriented, reliable, safe, and properly tested.

References

1. Esterel Technologies/Ansys, "SCADE Suite". http://www.esterel-technologies.com/products/scade-suite/ (accessed October 06, 2015) (online available)
2. Rubin, K.: Essential Scrum: A Practical Guide to the Most Popular Agile Process. Pearson Education (2013)
3. Schwaber, K.: Agile Project Management with Scrum. Microsoft Press (2004)
4. Ktata, O., Levésque, G.: Designing and implementing a measurement program for Scrum teams: what do agile developers really need and want? In: Proceedings of the Third C* Conference on Computer Science and Software Engineering. ACM (2010)

5. Mahnic, V., Zabkar, N.: Measuring Progress of Scrum-based Software Projects. Electronics & Electrical Engineering **18**(8), 73–76 (2012). ISSN: 13921215
6. Schwaber, K., Sutherland, J.: The Scrum Guide. Scrum.Org and ScrumInc., (2013)
7. Beck, K.: Embracing change with extreme programming. Computer **32**(10), 70–77 (1999)
8. Cohn, M.: User Stories Applied: For Agile Software Development. Addison-Wesley Professional (2004)
9. Javdani, T., et al.: On the current measurement practices in agile software development. arXiv preprint arXiv:1301.5964 (2013)
10. Gregory, J., Crispin, L.: Agile Testing. Addison Wesley, Boston (2009)
11. Gregory, J., Crispin, L.: More Agile Testing. Addison Wesley, Boston (2015)
12. Image of Sprint Backlog. https://www.scrum.org/portals/0/Images/Sprint-Backlog.png (accessed January 13, 2016) (online available)
13. Image of Planning Poker Cards. http://arquivo.devmedia.com.br/artigos/Roger_Ritter/tecnicas_ageis/image006.gif (accessed January 13, 2016) (online available)

Distributed Systems Performance for Big Data

Marcelo Paiva Ramos, Paulo Marcelo Tasinaffo, Eugenio Sper de Almeida, Luis Marcelo Achite, Adilson Marques da Cunha and Luiz Alberto Vieira Dias

Abstract This paper describes a methodology for working with distributed systems, and achieve performance in Big Data, through the framework Hadoop, Python programming language, and Apache Hive module. The efficiency of the proposed methodology is tested through a case study that addresses a real problem found in the supercomputing environment of the Center for Weather Forecasting and Climate Studies linked to the Brazilian Institute for Space Research (CPTEC / INPE), which provides Society a work able to predict disasters and save people lives. In all three experiments involving the issue, using the Cray XT-6 supercomputer: (i) the first issue involves programming in Python and a sequential and monoprocessed arquitecture; (ii) the second uses Python and Hadoop framework, over parallel and distributed arquitecture; (iii) the latter combines Hadoop and Hive in a parallel and distributed arquitecture. The main results of these experiments are compared, discussed, and topics beyond the scope in this research are exposed as recommendations and suggestions for future work.

Keywords Big Data · Hadoop · Hive · Python · Climate prediction · Distributed systems · Cluster HPC

1 Introduction

Over time the volume of accumulated data reaches considerable proportion, being a classical problem to find valuable information from an expressive data base. Some

M.P. Ramos · P.M. Tasinaffo · L.M. Achite · A.M. da Cunha · L.A.V. Dias
Computer and Electronic Engineering Graduate Program,
Brazilian Aeronautics Institute of Technology, Sao Jose Dos Campos, Sao Paulo, Brazil
e-mail: lmachite@gmail.com, {tasinafo,cunha,vdias}@ita.br

M.P. Ramos(✉) · E.S. de Almeida
Center for Weather Forecasting and Climate Studies, Brazilian Institute for Space Research,
Cachoeira Paulista, Sao Paulo, Brazil
e-mail: marcelopaivaramos@gmail.com, eugenio.almeida@cptec.inpe.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_64

issues are raised, related on how to handle this volume and diversity, how to store all this data, how to analyze it and generate information in a timely manner, and how to find out what is important. The term that refers to the whole related problem is Big Data [1, 2].

The technologies of Big Data relate a new generation of technologies and architectures designed to extract valuable information from very large volumes and a wide variety of data in a timely manner, to increase process efficiency and reduce the waste of resources involved. What provides higher speed capture, discovery and analysis, compared with the traditional model [1].

The term Big Data works with a new paradigm of architecture and programming, which requires a change in the current method of data analysis [3].

By mastering these methods of data analysis, companies will have a major competitive advantage in their area of operation; they can define action plans far more effective and focused on solving problems by analyzing the explosion of data generated. In 2001, Doug Laney defined the three Vs of Big Data. Other researchers, since that year, added new definitions. Therefore, presently there are five factors of dimension, or 5 Vs of Big Data: Volume, Velocity, Variety, Veracity and Value [3].

The Figure 1 illustrates these five factors dimension.

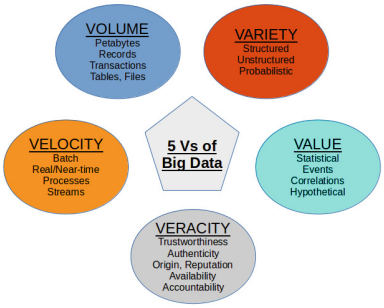


Fig. 1 The five Vs of Big Data [3].

The "volume" refers to the amount of accumulated data. The Big Data allows researchers to precisely deal with such large volumes of data, storing in different localities. The "velocity" is the rate of increase in the volume data as a function of time. A "variety" is how data is represented. A few years ago, most of the data was structure. Today, it is estimated that 80% of the data is not structured. With Big Data, messages, photos, images, videos, audios, etc. Should be worked in parallel to structured data, and this causes a processing overhead. The "veracity" is an important factor, because the processed data should lead to truthful information. Depending on the application, the accuracy of the data is essential and analysis tools should facilitate the cleaning of the data. Finally, the "value" is the most important factor of Big Data. It measures the usefulness of the data for objective decision making. This factor is directly related to return on investment (ROI) [1, 2].

There are at least three fundamental issues that surround the topic Big Data: transportation, management and processing. The first addresses the limitations of hard drives and computer networks, the second refers to the ownership and management of distributed data, aiming at solving problems of access, metadata, use, update, governance and reference, and the latter explains the parallel rather than serial processing due to slowness when it comes to a large volume of data [1].

1.1 Related Work

The main difficulty in dealing with Big Data is due to the volume of data that increases rapidly in comparison to computing resources. A technique that minimizes this problem is data deduplication. It removes replicas, saves hard disk space and data traffic on a computer network, but it overloads the CPU with additional processing. The Input / Output (I / O) time in disk is another variable to be considered [2, 4, 5].

Tekiner, and Keane [6] display concepts about the era of Big Data and how important it is for enterprises to get useful information quickly, to decision making and achieve competitive results.

The Hadoop framework is one of the tools used in Big Data and it has parameters that can be adjusted to improve performance. The Herodotus work [7] describes a detailed set of performance for mathematical models on all MapReduce execution phases in Hadoop, involving variables associated with Hadoop parameters, jobs and processed data, specific parameters of I / O cost, CPU and Network. Joshi [8] discusses some of the best practices for configuring hardware and software of clusters through TeraSort jobs in the Apache Hadoop framework. And Kim et al. [9] implement an automatic benchmarking setup method in order to facilitate the identification process of the configuration parameters set, optimized for better performance of Hadoop HDFS system.

The performance is a feature of Hadoop explored and proven through benchmark, as can be seen in the following literature. Heger [10] performs a TeraSort benchmark by using the Hadoop scheduler default settings, consequently generating the following items: Data Disk Scaling, Data Compression, JVM Reuse Policy, HDFS Block Size, Map Side Spills, Copy Phase Tuning, Reduce Side Spills, JVM Configuration Tuning, and OS (Linux) Tuning.

The benchmark performed by Kc, et al. [11] studies the behavior of jobs processed in a Hadoop cluster of 540 nodes, through experiments that alter the amount of data and specific settings for parallelism in order to reduce tasks and number of logs.

In regards to the work of Gunther, et al. [12], it displays an illusory phenomenon of superlinear scalability of Hadoop, caused by the I / O bottleneck in large clusters, where the parallel performance is above the value given by the speedup metric equation $Sp = T1/Tp$.

The utilization of Hadoop in the analysis of servers? logs is indicated by several authors. For example, Hingave and Ingle [13] use Hadoop in Log analysis to generate

statistical reports, improve business strategies, time optimization and efficient use of available resources.

Another application of Hadoop is in the security area, as in the method proposed by Hameed and Ali [14], denominated HADEC, capable to detect in a timely manner four great flood DDoS attacks: TCP-SYS, HTTP GET, UDP e ICMP; using MapReduce and HDFS, present in the Hadoop framework. The Hadoop performance, according to Prabhu, et al. [15], can be optimized to work with log files from web servers, reporting improvements of 32.97% after adjustment.

Therefore, Hadoop has demonstrated performance in the analysis of Big Data and textual log files, which provides strategic management support and ensure more security, while optimizing time and use of resources.

1.2 Big Data Contextualization

During the last twenty years, CPTEC generated a lot of data from the supercomputing, pre-processing or post-processing environment. Another source contributing to the final amount are the log files, coming from computer networking equipment, internal servers or located in the demilitarized zone (DMZ) [16].

Thus an environment is characterized Big Data, where traditional means of obtaining information are obsolete. Not being possible to analyze the whole volume of data and find important information in a timely manner. Information such that it comes to weather and climate, can be a divider between life and death [1, 2, 6, 16].

Within this scenario, it is possible to determine some points of motivation to conduct Research and Development (R & D) in institutional and academic environment.

Next sections describe: the Big Data Methodology; the distributed systems and Hadoop; the programming language python; the Apache Hive; the experiments; an analysis and discussion of some obtained main results; and also some conclusions, recommendations, and suggestions for future works.

2 Background

2.1 Apache Hadoop

The Apache Hadoop project is characterized by being a framework that enables the distributed data processing in cluster of computers, and is open source. Was modeled to work with Big Data, where a single server controls other computers, providing local storage and processing [17, 18].

The processing can be divided into two components: resource management and process scheduler. In resource management the ResourceManager establishes communication with NodeManager and identifies in real time the status of the resources.

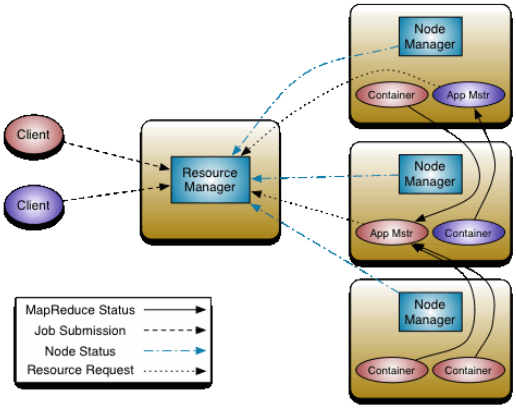


Fig. 2 Distributed Data Processing [17].

To scale processes the ResourceManager creates a ApplicationMaster module for application which may be responsible for managing the operation of MapReduce [17, 18].

The Figure 2 illustrates the architecture used by Hadoop to perform the distributed data processing.

The data storage is called Hadoop Distributed File System (HDFS) and works also following a client / server architecture. The HDFS has three components: NameNode and SecundaryNameNode are located in the server; and DataNode is present in each node of the cluster [17, 18].

The NameNode is responsible for storing metadata, linking the blocks of a file relating to the DataNode. The SecundaryNameNode has the function of monitoring and managing the NameNode, ensuring their availability. Already DataNode is where the blocks of a file are stored in fact [17, 18].

To write a file in HDFS, the Hadoop divides it into one or more blocks and distributes it among DataNodes, these blocks can be replicated to provide fault tolerance [17, 18].

The Figure 3 shows the method used by Hadoop to work with distributed data storage.

MapReduce is a model proposed by Google, which addresses a new programming paradigm for working with Big Data. It allows the manipulation of Big Data in a parallel and distributed manner, and provide fault tolerance, scheduling I / O and monitoring [17, 18, 19, 20].

The Figure 4 shows the Map / Reduce flow data used by Haddop for parallel processing Big Data.

This is accomplished through two major operations: Map and Reduce. During the Map, each block file is processed resulting in one key / value pair. Then these values go through sort and shuffle, and are forwarded to the next step. And finally,

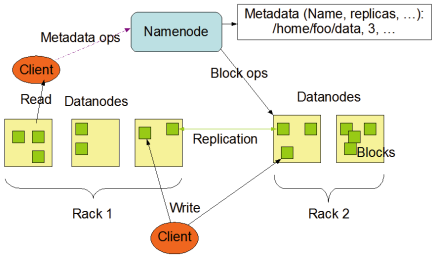


Fig. 3 Distributed Data Storage [17].

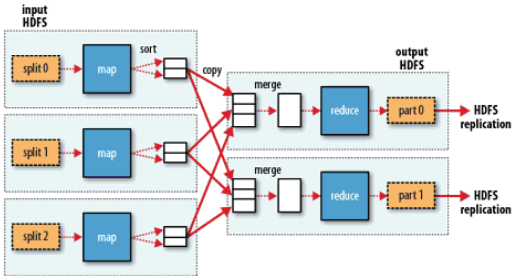


Fig. 4 Map / Reduce data flow [17].

the Reduce performs calculations according to the parameters key / value sorted and rearranged [17, 18, 19, 20].

2.2 Programming Language Python

Python is a free programming language and is open source. Was developed by Guido van Rossum in 1989 and is currently managed by the Python Software Foundation [21, 22, 23].

It is a high-level language, classified in the literature as third generation language. Interpreted, multi-platform, imperative, object-oriented, dynamic and strong typing [21, 22, 23].

She is considered to be easy to learn. Simple language that has no delimiters and performs the control blocks by indentation. With a focus on code readability and in a concise and clear syntax [21, 22, 23].

Because the implied facilities in its purpose, is a language that prioritizes productivity over maintenance. At the same time offer powerful features of its standard library and its modules and frameworks that can be coupled to the environment, in addition to a high-level data structure [21, 22, 23].

It is an interpreted language. As she does not perform compile it provides a gain of time for developers during the development phase of the software, but loses in performance when compared to compiled languages such as C and FORTRAN [21, 22, 23].

2.3 Apache Hive

The Apache Hive began as a subproject of Apache Hadoop, but due to its importance, currently has its own scope. It is a project open and collaborative code, managed by the Apache Software Foundation [24, 25].

Is a software for Data Wherehouse that facilitates management, manipulation and extraction of information from a large volume of data [24, 25].

The Hive is a module that can be coupled to Hadoop and make use of its technology for distributed storage HDFS and its processing Map / Reduce, focusing on abstracting the complexity. Its query language is HiveQL, is simple and similar to SQL [24, 25].

Even with similar features and functionality, Hive should not be compared with other Database Management Systems (DBMS), such as MySQL, PostgreSQL, Oracle, etc. Because it has high latency even when working with a reduced data set, but this time interval is relatively short when it comes to a large volume of data [24, 25].

Thus Apache Hive is designed for working with big data, where traditional DBMSs do not provide results at the expected time [24, 25].

3 Case Study

The experiments implemented in research related and reported in this paper address a problem in supercomputing environment Cray XT-6. Where users of the system generate a large number of files. So it is essential that management identify timely data volume and number of files belonging to each system user.

For these experiments a Log File generated on the Cray XT-6 system were used, which has a size of 38 GB. This Log file contains structured data by columns, as: inode, block size, permissions, number of links, owner name, owner group, size, month, day, year, file name.

The current method employed by CPTEC / INPE consists of a program implemented in Python. It computes each line sequentially and monoprocessed, running on a node operating environment Cray XT-6. This node has 4 CPUs Quad-Core AMD Opteron totaling 16 cores, 128 GB memory, 320 GB of local disk, 866 TB of primary storage, 3.84 PB of secondary storage and 6 PB of tertiary storage.

The second method addresses concepts of computing and big data using Hadoop framework. The programming language will also be Python, however processing will parallel and distributed. And will be processed in Hadoop Cluster [16, 17, 18, 19].

The third experiment also uses techniques of working with Big Data and can be considered a supplement to the second experiment. Instead of programming in Python, uses the Hive, which abstracts the use of HDFS and Map / Reduce using a high level and fourth-generation language [24, 25].

The HPC Cluster used in this research was acquired by CPTEC / INPE in 2007, initially used to improve the efficiency of numerical models for weather, climate and climate change [16].

This Cluster has scalar architecture besides being classified in the literature as Beowulf or MIMD. He has two central computers that manage the system, 275 slave nodes, 1100 processors, 5.7 TFlops, 24 TB of RAM and 72 TB of hard drive [16, 26, 27, 28, 29, 30].

The next section shows the comparison between the two methods employed in this case study, as well as answers to questions that arise in the analysis of the results process.

4 Results

To build the application compatible with the concepts of Big Data and to use the framework Hadoop, the authors of this paper have implemented two scripts in Python language. The first has the function to perform the mapper operation, while the second performing the reducer operation.

An example of Python code, that performs the function mapper can be seen in Figure 5.

```
#!/usr/bin/env python import sys
for line in sys.stdin:
    line = line.strip()
    words = line.split()
    print '%s\t%s' % (words[4]+"--"+words[5], "1")
```

Fig. 5 Mapper function.

Summing the total lines of code of both, has 54 lines as a result. And excepting blank lines and comments remain useful 26 lines of code. The current method used by CPTEC / INPE uses a single script developed in Python, with 149 lines of code and 91 useful lines.

Analyzing the values shown in Table 1, the current method has higher values in: total lines of code, useful lines, and processing time. The only factor that has lower value is the number of files used in each application.

Although the solution for Big Data is numerically better, by implementing the code, handle files, handle exceptions in code, does the Hadoop be considered difficult usability. Because users are accustomed with tools that address this issue in a more abstract level.

Table 1 Metrics of experiment 1

Factor	Current Methodology	Second Experiment
Place of execution	Tupa Cluster	UNA
Total lines of code	149	54
Useful lines	91	26
Files	1	2
Processing time	19 minutos	2m21s
Language	Python	Python

In the execution of both applications, the Haddop obtained the best result in processing time because it was 8.6 times faster. In addition to running on a specific cluster processing Big Data. In the current model work is executed on the server in a node of the supercomputer Cray XT-6. This can lead to competition with the operational activities of the machine and encumber the system.

A question that may arise among the readers of this paper is: if to code the algorithm currently used by CPTEC / INPE for parallel execution and perform processing on 16 CPUs, would be faster than Hadoop? The answer would be "No". In this case it would not be possible in the basement optimal parallelization, when the cost has the same order of sequential processing: $p.T_p = \theta(T_s)$. Because the hard drive is a device of I / O considered slow, being the bottleneck in the system. Hadoop has better performance in this respect to perform distributed scheduling I / O, by HDFS [31, 32].

The Figure 6 illustrates a speed disk write test performed on the local disk of the Cluster UNA where Hadoop is working.

```
root@una1:~# time dd if=/dev/zero of=testeSpeedy bs=64k count=16k conv=fdatasync
16384+0 registros de entrada
16384+0 registros de saída
1073741824 bytes (1,1 GB) copiados, 12,7898 s, 84,0 MB/s

real    0m13.032s
user    0m0.016s
sys     0m2.880s
root@una1:~#
```

Fig. 6 Speed disk write test.

Following this same analogy, it is possible to formulate another question: in possession of the parallel program, of 16 CPUs for execution, and a volume of 9 drives configured in a RAID 0, where I / O is done in parallel (striping), the speed would be greater than the Hadoop? The Figure 7 illustrates the architecture of RAID 0.

In this configuration would be faster. But Hadoop is even better, since it is Big Data. With nine hard drives of 200GB capacity in RAID 0, for example, would have a total of 1.8 TB capacity. And when it comes to Big Data units are: tera, peta, exa, etc.

Thus the volume of data used in this second experiment is below the capacity of processing and storage of Hadoop, yet he obtained satisfactory results.

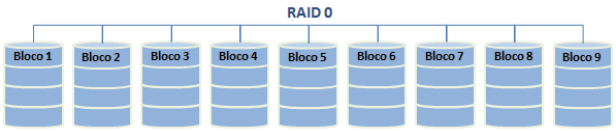


Fig. 7 Architecture of RAID 0 (Striping) [26].

The third experiment can be regarded as a complement of the second experiment. In it we used the Apache Hive, a module that can be coupled to Hadoop and make use of its technology for distributed storage HDFS and its processing Map / Reduce.

The Hive abstracts the complexity when using queries HiveQL, thus it is not necessary to program Map / Reduce for every need of the query based Big Data. Therefore, offers a way to work more friendly and at the same time best performance by efficient use of resources of Hadoop.

As can be seen in HiveQL query (Figure 8), in just one line of code you can select a list of users and groups, well as the number of files and total disk space used.

```
SELECT user, group, COUNT(file), SUM(size) FROM tupadisk GROUP BY user, group;
```

Fig. 8 A select HiveQL query.

Comparing the second and third experiments, the query HiveQL had time to 1m18s, being 44.68% faster.

5 Conclusion

The Haddop has a relatively simple architecture, operating as client / server, but does not offer the features that users are accustomed. Therefore, has a use evaluated as complex.

Another highlight was the relationship among Big Data and the terms I / O to disk, storage capacity and a possibility of using unstructured files. Thus performance on Big Data makes it possible to analyze the data in its entirety and to generate information in a timely manner.

The experiment with Apache Hive showed better results due to efficient use of Hadoop resources. And about usability, this architecture had low difficulty and complexity. The HiveQL offered greater dynamics and flexibility to extract information from a large database.

The authors of this research recommend the implementation of new case studies, involving larger data volume and the study of modules compatible with Hadoop, that working in a higher layer, abstracting the complexity and making its simplified use.

Among the tools which could be evaluated by performance and easiness are: Pig (a scripting language for MapReduce), Hbase (database Hadoop), Flume (export system logs) and Sqoop (export of DBMSs data to Hadoop).

Acknowledgments The authors of this paper thank the Center for Weather Forecasting and Climate Studies linked to Brazilian Institute for Space Research and Brazilian Aeronautics Institute of Technology and its Computer Engineering Undergraduate Program and Electronic and Computer Engineering Graduate Program for supporting the development of this work and accepting the challenges of innovation in this research.

References

1. Kaisler, S., Armour, F., Espinosa, J.A., Money, W.: Big Data: issues and challenges moving forward. In: 46th Hawaii International Conference on System Sciences. IEEE (2013)
2. Katal, A., Wazid, M., Goudar, R.H.: Big data: issues, challenges, tools and good practices. In: 2013 Sixth International Conference on Contemporary Computing (IC3) (2013)
3. Sagiroglu, S., Sinanc, D.: Big data: a review. In: 2013 International Conference on Collaboration Technologies and System (CTS) (2013)
4. Xiong, W., Yu, Z., Bei, Z., Zhao, J., Zhang, F., Zou, Y., Bai, X., Li, Y., Xu, C.: A characterization of big data benchmarks. In: 2013 IEEE International Conference on Big Data (2013)
5. Zhou, R., Liu, M., Li, T.: Characterizing the efficiency of data deduplication for big data storage management. In: 2013 IEEE International Symposium on Workload Characterization (IISWC) (2013)
6. Tekiner, F., Keane, J.A.: Big data framework. In: 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (2013)
7. Gunther, N.J., Puglia, P., Tomasette, K.: Hadoop superlinear scalability. Communications of the ACM **58**(4), 46–55 (2015)
8. Hingave, H., Ingle, R.: An approach for MapReduce based log analysis using Hadoop. In: 2015 2nd International Conference on Electronics and Communication Systems (ICECS), pp. 1264–1268. IEEE (2015)
9. Joshi, S.B.: Apache hadoop performance-tuning methodologies and best practices. In: Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering, pp. 241–242. ACM (2012)
10. Kamal, K.C., Hsu, C.J., Freeh, V.W.: Evaluation of mapreduce in a large cluster. In: 2015 IEEE 8th International Conference on Cloud Computing (CLOUD), pp. 461–468. IEEE (2015)
11. Herodotou, H.: Hadoop performance models (2011). arXiv preprint arXiv:1106.0940
12. Heger, D.: Hadoop performance tuning-a pragmatic & iterative approach. CMG Journal **4**, 97–113 (2013)
13. Hameed, S., Ali, U.: On the Efficacy of Live DDoS Detection with Hadoop (2015). arXiv preprint arXiv:1506.08953
14. Kim, J., Kumar, T.K.A., George, K.M., Park, N.: Performance evaluation and tuning for MapReduce computing in Hadoop distributed file system. In: 2015 IEEE 13th International Conference on Industrial Informatics (INDIN), pp. 62–68. IEEE (2015)
15. Prabhu, S., Rodrigues, A. P., Prasad, G., Nagesh, H.R.: Performance enhancement of hadoop mapreduce framework for analyzing bigdata. In: 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), pp. 1–8. IEEE (2015)
16. Centro de Previsão de Tempo e Estudos Climáticos. http://www.cptec.inpe.br
17. The Apache Software Foundation. http://hadoop.apache.org
18. White, T.: Hadoop: The Definitive Guide, 3rd edn. O'Reilly Media (2012)

19. Shvachko, K., Kuang, H., Radia, S., Chansler, R.: The hadoop distributed file system. In: Mass Storage Systems and Technologies (2010)
20. Pandey, S., Tokekar, V.: Prominence of mapreduce in BIG DATA processing. In: 2014 Fourth International Conference on Communication Systems and Network Technologies (CSNT) (2014)
21. Lutz, M.: Learning Python, 3rd edn. O'Reilly Media, Inc (2008)
22. Gift, N., Jones, J. M.: Python for Unix and Linux Systems Administration, 1st Edition, O'Reilly Media, Inc (2008)
23. Python Software Foundation. https://www.python.org
24. Apache Hive TM. https://hive.apache.org
25. Hive Tutorial. https://cwiki.apache.org/confluence/display/Hive/Tutorial
26. Tanenbaum, A.S.: Organização Estruturada de Computadores, 5th edn. Prentice-Hall, São Paulo (2007)
27. Tanenbaum, A. S.: Redes de Computadores, 4th Edition, Elsevier (2003)
28. Tanenbaum, A.S.: Modern Operating Systems, 3rd edn. Pearson Education, Inc (2008)
29. Dantas, M.: Computação Distribuída de alto desempenho: Redes, Cluster e Grids Computacionais, 1st edn. Axcel Books do Brasil Editora, Rio de Janeiro (2005)
30. Pitanga, M.: Construindo Supercomputadores com Linux, 2nd edn. Rio de Janeiro (2004)
31. Mattson, T., Sanders, B., Massingill, B.: Patterns for Parallel Programming, 1st edn. Addison-Wesley Professional (2004)
32. Pacheco, P.S.: An Introduction to Parallel Programming, 1st edn. Elsevier (2011)

Towards Earlier Fault Detection by Value-Driven Prioritization of Test Cases Using Fuzzy TOPSIS

Sahar Tahvili, Wasif Afzal, Mehrdad Saadatmand, Markus Bohlin, Daniel Sundmark and Stig Larsson

Abstract Software testing in industrial projects typically requires large test suites. Executing them is commonly expensive in terms of effort and wall-clock time. Indiscriminately executing all available test cases leads to sub-optimal exploitation of testing resources. Selecting too few test cases for execution on the other hand might leave a large number of faults undiscovered. Limiting factors such as allocated budget and time constraints for testing further emphasizes the importance of test case prioritization in order to identify test cases that enable earlier detection of faults while respecting such constraints. This paper introduces a novel method prioritizing test cases to detect faults earlier. The method combines TOPSIS decision making with fuzzy principles. The method is based on multi-criteria like fault detection probability, execution time, or complexity. Applying the method in an industrial context for testing a train control management subsystem from Bombardier Transportation in Sweden shows its practical benefit.

Keywords Software testing · Fault detection · Test Cases Prioritization · Optimization · Fuzzy logic · MCDM · TOPSIS · Failure rate

1 Introduction

Value-based software engineering [1] emphasizes the importance of integrating value considerations in software development. In software testing, prioritizing by value has immediate benefits, given that a lot of time and effort is spent on testing, rework and

S. Tahvili(✉) · M. Saadatmand · M. Bohlin · S. Larsson
SICS Swedish ICT, Västerås, Sweden
e-mail: {sahart,mehrdad,markus.bohlin,stig.larsson}@sics.se

S. Tahvili · W. Afzal · M. Saadatmand · D. Sundmark
Mälardalen University, Västerås, Sweden
e-mail: {wasif.afzal,daniel.sundmark}@mdh.se

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_65

fixing of faults. In this paper, we use a value-based method of test case prioritization whereby we combine a multi-criteria decision making (MCDM) technique with fuzzy logic. The MCDM technique used is called TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) [2] that is combined with fuzzy principles and hence called as FTOPSIS [3]. To optimize for value, we have considered multiple factors but start with factors most important from a customer perspective: quality and time-to-market. In order to maximize quality, we want to execute test cases having higher fault detection probability while in order to reduce time-to-market, we need to minimize time in terms of reduced number of test executions. Still, we want to satisfy additional criteria such as cost and requirements coverage but after having optimized for the most important ones. The advantages of using TOPSIS include its ease of use, simplicity and ability to keep constant steps regardless of the problem size [4]. Using fuzzy principles with TOPSIS suits the problem of test case prioritization which is characterized by interplay of complex factors and difficulties in gathering precise data. Considering the time and cost required to execute each test case from a test suite, the ideal situation is to detect the same number of faults (as detected by the whole test suite) by executing as few test cases as possible and under minimum amount of time. To achieve this goal, prioritization of test cases based on their fault detection probability and execution time is required. However, other test case properties such as the number of requirements that they cover may also serve as additional prioritization criteria. Properties such as these are important for testers and thus need to be considered as well. Therefore, in our optimization problem to find the closest solution (i.e., a set of test cases) to the aforementioned ideal situation, such other criteria play a role and can affect the result. It is naturally beneficial to use a multi-criteria technique for solving such a prioritization problem. This paper, proposes such an approach for prioritizing a set of test cases for integration testing by using TOPSIS. Our approach enables to identify a set of test cases which are closest to the ideal situation, and therefore, contribute towards the ultimate goal of identifying faults earlier and under a shorter time. To facilitate the application of our approach in real world scenarios where precise specification of quantified values for different criteria may not always be possible, we apply and combine fuzzy concepts with TOPSIS to enable the specification of values using linguistic variables (i.e., high, low, etc.). Finally, to demonstrate and evaluate the applicability of our approach, we have applied it on a train control management subsystem from BT.

2 Background and Preliminaries

Today, there are several aspects of optimization in the testing process with different goals, such as increasing fault detection rate [5], decreasing the use of redundant test cases, or use resources more efficiently. By prioritizing test cases, we are able to propose an order for executing test cases at different levels of testing process. In this order, every single test case will be ranked with a value, which can either be a “sharp” value, or, in the context of fuzzy logic, a linguistic value (e.g., low, high, etc.) As there are

more than one criteria which can affect the ranking of test cases, a sensible approach is to consider several of the aforementioned objectives at the same time as a multi-objective optimization problem (e.g. maximizing the rate of fault detection and minimizing cost). To solve this kind of problem, several multi criteria decision making (MCDM) analysis techniques have been developed. MCDM techniques formulate the problems as the interdependency between various criteria and alternatives [4]. TOPSIS, originally developed by Hwang and Yoon (see [2]), is one such MCDM technique which has been shown useful when combined with a fuzzy approach, and when there are some conflicting and no commensurable criteria in the initial problem.

2.1 Motivating Example

As mentioned earlier, for using TOPSIS we need to identify the criteria which have effect on the alternatives. Some standard criteria for software testing might be time, cost, requirement coverage, etc. In reality, there are some additional limitations during the testing process, for example due to budget, time, or resources constraints. By proposing an optimal order for execution, we are able to execute test cases according to such constraints. Figure 1 illustrate a MCDM problem where 5 different criteria (C_1 to C_5) have a direct effect on every single alternative (test cases), of which there are 10 in total. As we can see in Figure 1, there are limitations on budget and time in the problem. In fact, testers can only execute a limited number of test cases before the deadline or budget limit is reached. 'Goal achieved' in Figure 1 represents that testers detect the expected number of faults after executing 8 test cases in this particular order. The last two test cases (T_9 , T_{10}) are thus considered as redundant.

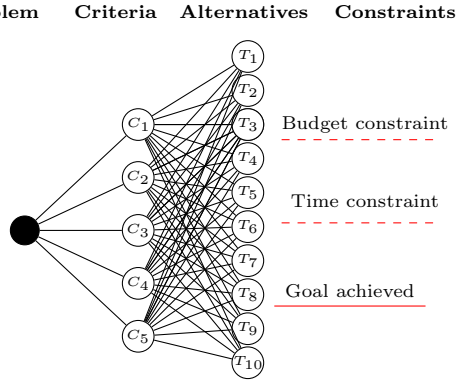


Fig. 1 Illustration of a MCDM problem with constraints.

3 Proposed Approach

The TOPSIS method is based on identification of two ideal solutions; the *positive ideal solution* (PIS) and the *negative ideal solution* (NIS). The PIS consists of the best value for each criteria (taken from the set of alternatives), and similarly, the NIS consists of the worst values for each criteria. The PIS, in the concept of prioritizing test cases, will then be an ideal test case which satisfies all the identified criteria properly, such as a fast (execution) and cheap (implementation cost) test case with very high probability of fault detection which tests a set of complex requirements. However, the negative ideal solution comprises slow and expensive test cases with very low probability of detecting the faults and just tests a simple single requirement. In TOPSIS, alternatives are then ranked according to their distance from the PIS (in increasing order of preference) and the NIS (in decreasing order of preference) simultaneously [6].

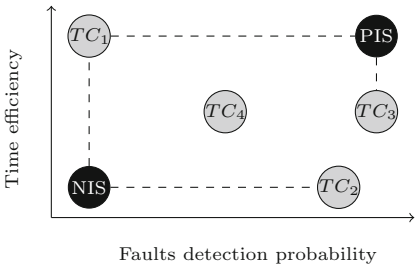


Fig. 2 Positive (PIS) and negative (NIS) ideal solutions in a bi-criteria prioritization problem for four test cases.

Figure 2 illustrates TOPSIS for two different criteria; time efficiency and fault detection probability. Measuring the effect of the criteria on the test cases is not an easy task as some criteria such as execution time, can be measured by a sharp value and some criteria like cost, can be written as a function of time or lines of code (or other measurable criteria). For estimating some other criteria (such as fault detection probability), software developers and testers can provide estimations using fuzzy linguistic variables by comparing test cases with each other. For this purpose, we later extend the initial problem to the fuzzy environment and use linguistic variables to capture the effect of weighted values with some uncertainty. It should be noted that this approach is not limited to any particular set of criteria and some other criteria can well be applied. After identifying a set of test cases that are closest to the positive ideal solution, we can calculate the fault failure rate (explained in Section 3.3) for this set which enables us then to compare them against other (sub-)set of test cases from the test suite or the whole test suite. Moreover, by using the fault failure rate, we can compare the efficiency of different set of test cases with respect to detecting faults and the number of test cases to execute to detect them. In short, our proposed

approach consists of the following steps to enable earlier detection of faults through test case prioritization:

1. Identify the set of criteria for prioritization of test cases (besides fault detection probability and time efficiency).
2. Determine the criteria value for each test case using fuzzy linguistic variables. This can be done, for instance, by sending a questionnaire to testers.
3. Apply fuzzy TOPSIS in order to prioritize test cases and find a set of test cases closest to the positive ideal solution. During this process, consider the fault detection probability and time efficiency criteria as the most important among the set of criteria.

As an additional step and in order to compare and enable the evaluation of the identified set resulting from the application of FTOPSIS, fault failure rate can be calculated for the set of test cases. In the following sections, we describe the fuzzy TOPSIS method by using basic concept of fuzzy logic provided by Yang [7], Baets and Kerre [8] and Yun Shi [9].

3.1 Intuitionistic Fuzzy Sets (IFS)

Fuzzy set theory was proposed by Zadeh [10] for solving MCDM problems based on the inclusion degrees of Intuitionistic Fuzzy Sets (IFS) and a membership function. In this section, we define these formally.

Definition 1. *A fuzzy set is a pair* (A, m_A) *where* A *is a set and* $m_A : A \to [0, 1]$*; for each* $x \in A$*,* $m_A(x)$ *is called the grade of membership of* x *in* (A, m_A)*. Let* $x \in A$*, then* x *is fully included in the fuzzy set* (A, μ_A) *if* $\mu_A(x) = 1$ *and is fully excluded if* $\mu_A(x) = 0$ *where* x *is a fuzzy member if* $0 < \mu_A(x) < 1$ *(see[7]).*

The membership function can be illustrated by different shapes which helps interpreting the values appropriately. Triangular, bell-shaped, Gaussian and trapezoidal membership functions are the most common. In this paper, we use bell-shaped membership functions (see Figure 3) which is defined using the following formula:

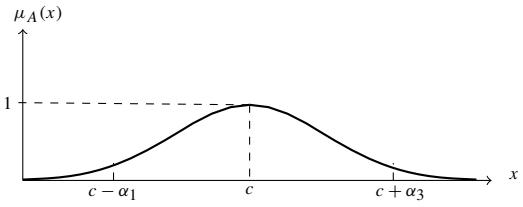


Fig. 3 Bell-shaped fuzzy membership function.

The membership of an element to a fuzzy set is a single value between 0 and 1 and can be obtained by:

$$\mu_{\tilde{A}}(x) = \frac{1}{1 + (x - \frac{c}{\alpha})^{2b}} \tag{1}$$

Since in reality, there is always a hesitation degree for calculating functional and non-functional degree, a generalization of fuzzy sets had been proposed as intuitionistic fuzzy sets (IFS) which incorporated the degree of hesitation called hesitation margin [11].

Definition 2. *An Intuitionistic Fuzzy Set (IFS) A on a universe U is defined as the following form:* $A = \{(u, \mu_A(u), \nu_A(u)) \mid u \in U\}$, *where the functions* $u_A : U \rightarrow [0, 1]$ *and* $v_A : U \rightarrow [0, 1]$ *define the degree of membership and the degree of non-membership of the element* $u \in U$ *in A, respectively, and for every* $u \in U$ *we have* $0 \leq \mu_A(u) + \nu_A(u) \leq 1$, *furthermore,* $\pi_A(u) = 1 - \mu_A(u) - \nu_A(u)$ *called the intuitionistic fuzzy set index or hesitation margin of u in A (see [11, 12]).*

Thus, a fuzzy set can be written as:

$$\{(u, \mu_A(u), 1 - \mu_A(u)) \mid u \in U\} \tag{2}$$

IFS distributes fuzzy sets for every membership function μ and non-membership functions ν where $\nu = 1 - \mu$ (see [13]).

Definition 3. *A function* $I : [0, 1]^2 \rightharpoonup [0, 1]$ *is called a fuzzy implication if the following conditions are satisfied for all* $x, y, z \in [0, 1]$*: if* $x \leq y$ *then* $I(x, z) \geq I(y, z)$ *and if* $y \leq z$ *then* $I(x, y) \geq I(x, z)$ *also the boundary conditions are:* $I(0, 0) = I(0, 1) = I(1, 1) = 1$ *and* $I(1, 0) = 0$ *(see [9]).*

A fuzzy implication can be calculated by using various formula such as Residuated implications (R-implications), Strong implications (S-implications) and Quantum Logic implications (QL-implications), in this work we use Residuated implications, proposed by Goguen [14]:

$$R_\pi(a, b) = \begin{cases} 1, & if\ a = 0, \\ \min\left(\frac{b}{a}, 1\right) & if\ a > 0 \end{cases} \tag{3}$$

Let U be a finite universe and R is an implication:

Definition 4. *Inclusion degree function of IFS denoted by* I_{IFS}, *if R satisfies the following conditions (see [9]):*

- $\forall a, b \in [0, 1]$ and $a \leq b \Rightarrow R(a, b) = 1$
- $R(a, b)$ is non-decreasing with respect to b and non-increasing with respect to a.

thus:

$$I_{IFS}(A, B) = \frac{1}{|U|} \sum_{u \in U} [\lambda R_\pi(\mu_A(u), \mu_B(u)) + (1 - \lambda) R_\pi(\nu_B(u), \nu_A(u))], \qquad \lambda \in [0, 1] \tag{4}$$

where $|U|$ is the cardinality of U and calculated as (see [15]):

$$|U| = \sum_{u \in U} \frac{1 + \mu_A(u) - \nu_A(u)}{2} \tag{5}$$

3.2 *Fuzzy TOPSIS*

As previously mentioned, TOPSIS is based on positive (PIS) and negative (NIS) ideal solutions. The best alternative per time, is an alternative which has the farthest distance from NIS and the shortest distance from PIS.

Let $A = \{A_1, A_2, ..., A_n\}$ be a set of alternatives and $C = \{C_1, C_2, ..., C_m\}$ be a set of identified criteria. By using IFS, we are able to represent A and C as:

$$\begin{aligned} A_1 &= \{(C_1, \mu_{1,1}, \nu_{1,1}), (C_2, \mu_{1,2}, \nu_{1,2}), ..., (C_m, \mu_{1,m}, \nu_{1,m})\} \\ A_2 &= \{(C_1, \mu_{2,1}, \nu_{2,1}), (C_2, \mu_{2,2}, \nu_{2,2}), ..., (C_m, \mu_{2,m}, \nu_{2,m})\} \\ &\vdots \\ A_n &= \{(C_1, \mu_{n,1}, \nu_{n,1}), (C_2, \mu_{n,2}, \nu_{n,2}), ..., (C_m, \mu_{n,m}, \nu_{n,m})\} \end{aligned} \tag{6}$$

where $\mu_{i,j}$ indicates the degree by which the alternative A_i satisfies criterion C_j and $\nu_{i,j}$ indicates the degree by which the alternative A_i does not satisfy criterion C_j [13]. In FTOPSIS, the sum of $\mu_{i,j}$ and $\nu_{i,j}$ does not exceed 1.

Definition 5. *A fuzzy positive ideal solution is defined as:*

$$PIS_f = \{(C_1, \max\{\mu_{i,1}\}, \min\{\nu_{i,1}\}), \ldots (C_m, \max\{\mu_{i,m}\}, \min\{\nu_{i,m}\})\} \tag{7}$$

Definition 6. *A fuzzy negative ideal solution is defined as:*

$$NIS_f = \{(C_1, \min\{\mu_{i,1}\}, \max\{\nu_{i,1}\}), \ldots (C_m, \min\{\mu_{i,m}\}, \max\{\nu_{i,m}\})\} \tag{8}$$

The distance between the alternatives to PIS_f and NIS_f can be measured by inclusion degrees:

Definition 7. *The inclusion degree $D^+(A_i)$ of the positively ideal solution in alternative A_i is calculated by:*

$$D^+(A_i) = \max(I(PIS_f, A_i)), \tag{9}$$

and the inclusion degree $d^{-}(A_i)$ of the negatively ideal solution in alternative A_i is respectively measured as:

$$d^{-}(A_i) = \min(I(A_i, NIS_f)) \tag{10}$$

where I represents the inclusion degree function, calculated by Eq. (4).

Definition 8. *The ranking index of alternative A_i is defined as:*

$$P_i = \frac{D^{+}(A_i)}{d^{-}(A_i) + D^{+}(A_i)} \tag{11}$$

where $0 \leq P_i \leq 1$.

If there exists $i_0 \in \{1, 2, ..., n\}$ where $P_{i_0} = \max\{P_1, P_2, ..., P_n\}$, then A_{i_0} is the best alternative [13]. In fact, by selecting the maximum value of ranked index of alternatives (P_i) per mentioned criteria, we propose a set of alternatives which have satisfied the criteria properly. The ranking index of alternatives in the concept of prioritizing test cases, is a set of test cases (alternatives) which has maximum probability of detecting faults and also has a high time efficiency.

3.3 Fault Failure Rate

Definition 9. *Assume T test cases which are available for use, let F be the number of test cases that fail during the testing a system under test. Then the failure rate λ_{f_n} can be defined as the proportion of the failed test cases on the total number of executed test cases:*

$$\lambda_{f_n} = \frac{F}{T} \tag{12}$$

A high value for λ_{f_n} indicates that the faults are overall easy to detect, which implies that, the executed test cases have detected relatively more faults. However, a low value for λ_{f_n} shows that the faults in the system under test are harder to detect [16]. To calculate the failure rate, we have checked the initial version (the faulty version) of a set of test cases in integration testing level at BT.

4 Application OF FTOPSIS

As mentioned in Section 3, for prioritizing test cases using FTOPSIS, we need to identify a set of criteria having a direct effect on test cases. In consultations with our industrial partner (BT), the following criteria have been identified as the most effective ones:

– *Fault detection probability:* The probability of fault detection of test cases.
– *Time efficiency:* The sum of setup, implementation and execution time of each test case.
– *Cost:* The sum of cost incurred in implementation, hardware setup and configuration for each test case.
– *Requirement coverage:* The requirement(s) tested by each test case.

A set of 86 test cases, which test *Drive-Brake Control* and *Auxiliary Control* sub-level function groups in the train control management system at BT at the integration testing level, has been chosen as alternatives. To measure the effect of above-mentioned criteria on the test cases, a questionnaire was filled by a test expert at BT, using linguistic variables such as low, medium and high. To avoid lengthy calculations, we present a subset of test cases with ratings for different criteria in Table 1:

Table 1 The effect of criteria on test cases, with values very low (VL), low (L), medium (M), high (H) and very high (VH)

Test Case ID	Fault Detection	Time Efficiency	Cost	Requirement Coverage
HVAC-007	M	L	L	M
AirSupply-036	H	L	M	L
Drive-S-046	M	H	H	VH
Speed-S-IVV-005	M	M	VL	M
AirSupply-IVV-047	M	VH	M	M
SyTs-ExtDoors-S-IVV-011	H	M	L	M
⋮	⋮	⋮	⋮	⋮
Brake-041	VH	M	M	H

The specifications of the alternatives (test cases) and criteria for IFS based on Eq. (6) are calculated as follows (again we show it for a subset of test cases):

$$\begin{aligned}
\text{HVAC-007} &= \{(C_1, 0.5, 0.5), (C_2, 0.1, 0.9), (C_3, 0.1, 0.9), (C_4, 0.5, 0.5)\} \\
\text{Airsupply-036} &= \{(C_1, 0, 9, 0.1), (C_2, 0.5, 0.5), (C_3, 0.5, 0.5), (C_4, 0.1, 0.9)\} \\
&\vdots \\
\text{Brake-041} &= \{(C_1, 0.9, 0.1), (C_2, 0.6, 0.4), (C_3, 0.5, 0.5), (C_4, 0.5, 0.5)\}
\end{aligned}$$

By using Eqs. (7) & (8), we calculate the positive and negative ideal solutions:

$$\begin{aligned}
PIS_f &= \{(C_1, 1, 0.1), (C_2, 0.8, 0), (C_3, 0.7, 0.3), (C_4, 0.9, 0.1)\} \\
NIS_f &= \{(C_1, 0.1, 0.9), (C_2, 0, 1), (C_3, 0.1, 0.9), (C_4, 0.1, 0.8)\}
\end{aligned}$$

The inclusion degrees of PIS and NIS are calculated by Eqs. (9) & (10) and the results have been summarized in Table 2.

Table 2 The inclusion degrees of PIS_f and NIS_f in A_i

	HVAC-007	Airsupply-036	Drive-S-046	...	Brake-041
$I(PIS_f, A_i)$	0.35	0.45	0.51		0.61
$I(A_i, NIS_f)$	0.78	0.41	0.48		0.45

Table 3 show the ranking indices (P_i) of test cases that are obtained by using Eq. (11). As we can see in Table 3, the test cases with ID number Brake-041 has the

Table 3 The ranking index of test cases

Test Case ID	HVAC-007	Airsupply-036	Drive-S-046	...	Brake-041
P_i	0.30	0.52	0.50		0.58

maximum value ($P = 0.58$), which means that this test case should be executed first. As is evident from Table 1, fault detection probability for it is high and the mentioned test case has a higher time efficiency in comparison with HVAC-007, Airsupply-036 and Drive-S-046. Based on Table 3, we propose the following order of execution: {Brake-041 → Airsupply-036→Drive-S-046 → ⋯ → HVAC-007}. By calculating the ranking index for every single test case in Table 1, we are able to get a set of best candidates for execution that satisfy first fault detection probability and time efficiency and then rest of other criteria. We propose a subset of 48 test cases for execution (48 test cases of among total 86 test cases) which are are time efficient test cases and also have a high ability to detect the faults.

4.1 Industrial Evaluation

To evaluate our approach using an industrial case study, we have monitored the result of executing 86 test cases at the initial level of integration testing at BT. In this level, 2 sub-level function groups, which are *Drive-Brake Control* and *Auxiliary Control*, have been tested by 86 test cases. Table 4 presents the results of monitoring the test effort.

Table 4 Integration test result at BT

Number of sub-level function groups	2
Executed test cases	86
Total passed test cases	69
Total failed test cases	7
Not set test cases	10
Fault failure Rate λ_{f_0}	0.092

As is shown in Table 4, to detect 7 faults in the initial level of integration testing at BT, 86 test cases have been executed, which implies that 69 test cases could not find any fault during the testing process. Total of 10 test cases have not been set up for execution, which indicates that there were errors in the test specifications and these 10 test cases could not even get started. According to BTs method for executing test cases, ?Not set test cases? (10 mentioned test cases in Table 4) will be tested in the next level of testing, which is system testing at BT. Table 4 also show the fault failure rate for the testing effort, as described in Section 3.3. The rate is obviously low since very few test cases failed.

Figure 4 shows the correlation observed for the time efficiency and fault detection probability for 86 test cases. Not all 86 test cases are distinctly visible in this figure as some of them are overlapping. The X-axis (horizontal axis) represents the probability of detecting fault per test case and the Y-axis (vertical axis) represents the time efficiency. Since probability is quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty), X-axis is built up on a scale of 0 to 1. As explained earlier, the truth value of the fuzzy number lies in between 0 and 1, which have been used to show the scale of Y-axis in Figure 4.

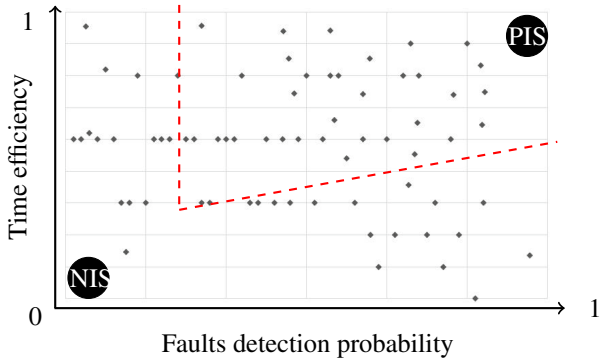


Fig. 4 Correlation for fault detection probability and time efficiency

We can see the positive ideal solution (PIS) and also the negative ideal solution (NIS), along with other test cases. As explained in Section 3, PIS represents a test case (or a set of test cases) that optimally satisfies the most important criteria. In Figure 4, 'fault detection probability' and 'time efficiency' have been illustrated as two critical criteria. In this situation, PIS is a test case (or a set of test cases) having highest probability of detecting faults and also being highly time efficient. As Figure 4 shows, some test cases have shorter distance to PIS and longer distance to NIS in comparison with other test cases and FTOPSIS identifies such test cases. The red dashed lines in Figure 4 represent the feasible set [1] of proposed test cases for execution. In other words, the test cases in the feasible set are satisfying both time efficiency and fault detection probability simultaneously. Other test cases which

[1] Is the set of all possible points of an optimization problem that satisfy the problem's constraints

are not located in the feasible set have still a chance to detect some faults but they don't satisfy other criteria. In fact, by adding a new test cases in the proposed set for execution (48 test cases of among total 86 test cases) we are exiting from the feasible set, which means that, the newly joined test case (in the proposed set) is not able to satisfy at least on of the criteria. To evaluate the performance of our prioritized test cases, we need to measure the fault failure rate of prioritized test cases at regular intervals of time and compare it with current execution of test cases as shown in Table 4. For example, if we take 48 test cases to execute, using the prioritized set of test cases, we achieve a fault failure rate of $\lambda_{f_1} = 0.145$ (where 7 failed test cases are divided by 48 executed test cases). This is a significant improvement over the initial fault failure rate of $\lambda_{f_0} = 0.092$. It needs to be mentioned that some of the ?Not set test cases? (10 mentioned test cases in Table 4) are also in our proposed set for execution, but because these 10 test cases could not even get started, we are not able to use them for measuring fault failure rate. This result signifies that prioritization of test cases has immediate benefits in detecting faults early and hence improving software quality. We aspire to continue monitoring the fault failure rate at more regular time instances to better quantify the gains from prioritizing test cases.

5 Related Work

Prioritization of test cases in a test suite is about ordering their execution that increases test effectiveness, typically through maximizing early fault detection. This means that certain test cases are more valuable than others and prioritizing their execution will add value to the testing effort. Much of the literature on test case prioritization is in the context of regression testing, see e.g., the review paper by Yoo and Harman [17] and rely on code coverage based strategies. The motivation for using code coverage based strategies is that maximizing structural coverage will improve early maximization of fault detection. These strategies are of course not applicable for cases where such coverage information is either not readily available (e.g., in real-time embedded systems [18]) or is hard to trace (e.g., for testing at higher levels than unit). Other authors have used a requirements-based approach [19, 20] where test cases are prioritized by properties such as customer-assigned priority and implementation complexity. A criticism of such approaches is that such properties are subjective [17]. This has opened opportunities for using fuzzy approaches for prioritizing test cases since these approaches can better simulate experts' reasoning (see e.g., [21, 22, 23, 24, 25, 26]). A fuzzy reasoning approach is also suited for a complex ranking problem as test case prioritization which is impacted by different criteria having uncertain, subjective and imprecise data. Fuzzy-based TOPSIS has previously been used in other domains (see e.g., [27]); in this paper we seek to evaluate its applicability in the context of test case prioritization. In this paper, we have used a fuzzy-based multi-criteria decision-making method known as FTOPSIS. Fuzzy-based TOPSIS has previously been used in other domains (see e.g., [27]); in this paper we seek to evaluate its applicability in the context of test case prioritization.

6 Conclusion and Future Work

Considering that the execution time for running various test cases are different, and also that each test case can have a different fault detection rate, it makes more sense to select the test cases for execution which have higher probability to detect faults and at the same time take shorter execution time. By prioritizing test cases based on such an order, it becomes possible to detect the faults in a system earlier. It should be noted that by prioritizing the test cases we are not able to detect more faults, yet the hidden faults in the system under test will be detected earlier. Using this method can be useful, for instance, in exploratory testing when time is severely limited to select and execute some random test cases to detect the hidden faults. In this paper, we introduced an approach based on the combination of fuzzy logic and TOPSIS multi-criteria decision making technique towards identifying a set of test cases that can detect faults with higher detection probability while at the same time result in a shorter overall execution time; hence earlier fault detection. We did this particularly by considering fault detection probability and time efficiency properties of each test case as two important criteria that are used in the decision making process. Through an example, we showed how the approach can be applied in practice. To enable the efficiency evaluation of the identified set of test cases, we used the concept of fault failure rate as an indicator to compare the fault detection capability of different sets of test cases. As a future work, we plan to perform more industrial case studies in order to evaluate more precisely how the identified set of test cases based on our approach and considering the estimated fault detection probability of test cases compares in practice against the test cases that have failed after executing the complete test suite. In other words, we are interested to perform more evaluations to see how our prediction in terms of suggesting a set of test cases matches the result of test suite after execution. From this perspective, one factor that can affect the outcome of our proposed approach is the estimation accuracy of fault detection probably and execution time of test cases which are provided by testers in a subjective manner, and thus are prone to human judgment errors. Investigation of methods and techniques that help in providing more accurate estimation of fault detection probability of test cases would be another interesting direction of this work. One observation that we had as part of this work was that the test cases that had higher fault detection probability, also targeted and covered more requirements as well. It would be interesting to investigate whether such correlation can also be observed in other systems and domains, and how existence of such correlation can help with better design of test cases to achieve a higher fault detection rate in a more efficient way. Another advantage of our proposed method is that the decision making part can be performed in an automatic and systematic way. However, the part that requires manual work and intervention is when values for various test criteria are to be specified. This can also be a limiting factor with respect to the scalability of our approach when the number of test cases for which various criteria values need to specified grows more and more. Identification of techniques and methods for more

precise estimation of criteria values might help with the scalability issue but deserves a more thorough study and investigation in a separate work.

Acknowledgements. This work was supported by VINNOVA grant 2014-03397 through the IMPRINT project and the Swedish Knowledge Foundation (KKS) grant 20130085 through the TOCSYC project, and the ITS-EASY industrial research school.

References

1. Boehm, B.: Value-based software engineering. SIGSOFT Software Engineering Notes **28**(2), 4 (2003)
2. Hwang, C.-L., Yoon, K.: Methods for multiple attribute decision making. In: Multiple Attribute Decision Making, pp. 58–191. Springer (1981)
3. Chen, C.-T.: Extensions of the {TOPSIS} for group decision-making under fuzzy environment. Fuzzy Sets and Systems **114**(1), 1–9 (2000)
4. Velasquez, M., Hester, P.T.: An analysis of multi-criteria decision making methods. International Journal of Operations Research **10**(2), 56–66 (2013)
5. Rothermel, G., Untch, R.H., Harrold, M.J.: Prioritizing test cases for regression testing. IEEE Transactions on Software Engineering **27**(10), 929–948 (2001)
6. Amir-Aref, M., Javadian, N., Kazemi, M.: A new fuzzy positive and negative ideal solution for fuzzy TOPSIS. WSEAS Transactions on Circuits and Systems **11**, 92–103 (2012)
7. Yang, J., Watada, J.: Fuzzy clustering analysis of data mining: Application to an accident mining system. International Journal of Innovative Computing, Information and Control **8**(8), 5715–5724 (2012)
8. De Baets, B., Kerre, E.: Fuzzy relational compositions. Fuzzy Sets and Systems **60**(1), 109–120 (1993)
9. Michał, B., Balasubramaniam, J.: Fuzzy implications. Springer-Verlag, Heidelberg (2008)
10. Zadeh, L.A.: Fuzzy sets. Information and Control **8**(3), 338–353 (1965)
11. Ejegwa, P.A., Akowe, S.O., Otene, P.M., Ikyule, J.M.: An overview on intuitionistic fuzzy sets. International Journal of Scientific and Technology Research **3**(3), 142–145 (2014)
12. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets and Systems **20**(1), 87–96 (1986)
13. Yu, C., Luo, Y.: A fuzzy optimization method for multi-criteria decision-making problem based on the inclusion degrees of intuitionistic fuzzy sets. In: Huang, D.-S., Wunsch II, D.C., Levine, D.S., Jo, K.-H. (eds.) Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. LNCS, vol. 5227, pp. 332–339. Springer, Heidelberg (2008)
14. Goguen, J.A.: The logic of inexact concepts. Synthese **19**(3), 325–373 (1969)
15. Fan, J., Xie, W., Pei, J.: Subsethood measure: new definitions. Fuzzy Sets and Systems **106**(2), 201–209 (1999)
16. Debroy, V., Wong, W.E.: On the estimation of adequate test set size using fault failure rates. The Journal of Systems and Software, 587–602 (2011)
17. Yoo, S., Harman, M.: Regression testing minimization, selection and prioritization: A survey. Software Testing, Verification and Reliability **22**(2), 67–120 (2012)
18. Wikstrand, G., Feldt, R., Gorantla, J.K., Wang, Z., White, C.: Dynamic regression test selection based on a file cache: an industrial evaluation. In: Proceedings of the 2009 International Conference on Software Testing, Verification and Validation (ICST 2009) (2009)
19. Srikanth, H., Williams, L., Osborne, J.: System test case prioritization of new and regression test cases. In: Proceedings of the 2005 International Symposium on Empirical Software Engineering (ESE 2005) (2005)
20. Krishnamoorthi, R., Sahaaya Arul Mary, S.A.: Factor oriented requirement coverage based system test case prioritization of new and regression test cases. Information and Software Technology **51**(4), 799–808 (2009)

21. Tahvili, S., Saadatmand, M., Bohlin, M.: Multi-criteria test case prioritization using fuzzy analytic hierarchy process. In: Proceedings of the 10th International Conference on Software Engineering Advances (ICSEA 2015) (2015)
22. Alakeel, A.M.: Using fuzzy logic in test case prioritization for regression testing programs with assertions. The Scientific World Journal **2014**, 1–9 (2014)
23. Malz, C., Jazdi, N., Gohner, P.: Prioritization of test cases using software agents and fuzzy logic. In: Proceedings of the 2012 IEEE 5th International Conference on Software Testing, Verification and Validation (ICST 2012) (2012)
24. Zhiwei, X., Gao, K., Khoshgoftaar, T.M., Seliya, N.: System regression test planning with a fuzzy expert system. Information Sciences **259**, 532–543 (2014)
25. Hettiarachchi, C., Do, H., Choi, B.: Risk-based test case prioritization using a fuzzy expert system. Information and Software Technology **69**, 1–15 (2016)
26. Schwartz, A., Do, H.: A fuzzy expert system for cost-effective regression testing strategies. In: Proceedings of the 2013 29th IEEE International Conference on Software Maintenance (ICSM 2013) (2013)
27. Kim, Y., Chung, E.-S., Jun, S.-M., Kim, S.U.: Prioritizing the best sites for treated wastewater instream use in an urban watershed using fuzzy TOPSIS. Resources, Conservation and Recycling **73**, 23–32 (2013)

Part VII
Model-Driven Engineering for Cyber-Physical Systems

Experimenting with a Load-Aware Communication Middleware for CPS Domains

Luis Cappa-Banda and Marisol García-Valls

Abstract This paper describes the design of a middleware that is able to perform hybrid load balancing in a distributed systfarm, ensuring timely load balancing and client service. The described middleware is based on a distributed, modular software architecture design that allows to interface with sensor nodes that operate on a physical environment. The core entity is a balancer module that can run different algorithms in order to obtain different results for parameters such as client response times, response dispatching at the servers, or interaction strategy with the sensor nodes. A software implementation over a specific object oriented communication infrastructure is achieved, and a validation scenario is also presented to show the suitability of our design. We test it in an actual multi-process and multi-threaded scenario, analyzing its latency on several work-load conditions.

Keywords Load balancing · Server farms · Middleware · Cyber-physical systems

1 Introduction

Cyber-Physical Systems (CPS) are network intensive systems with an increasing trend to export the physical monitored data to backend centers and later provided to external clients upon request. Processing sensor data in a more powerful location (as a data center or simply a server farm) can apply more intelligent processing to the physical data in order to obtain an accurate picture on the state of the environment and on how to operate on it. Due to the increasing client demand for remote services that are executed in the cloud, it is extremely important to balance the load of the different

L. Cappa-Banda
Bugoroo, Alcobendas, Spain
e-mail: luiscappa@gmail.com

L. Cappa-Banda · M. García-Valls(✉)
Universidad Carlos III de Madrid, Av. de la Universidad 30, 28911 Leganés, Spain
e-mail: mvalls@it.uc3m.es

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_66

communication sessions that are, in the end, equivalent to user/device requests. To ease the programmability of these complex systems, communication middleware is a powerful alternative as it is an intermediate layer that facilitates the programming of distributed applications by abstracting the underlying details of the hardware and the specific protocols. It allows programmers to design their applications in high level programming paradigms (such as object-oriented or SOA-based) hiding the details of the underlying transport and routing to the server. However, middleware is often seen as a black box that provides a set of communication primitives, therefore introducing new hidden logic that increases the complexity of software systems and their analysis. Precisely, the CPS community tends to look at communication middleware as a source of unpredictability that may threaten the essential CPS requirements of *timeliness* and fully *predictable* behavior.

Large scale deployments require that middleware is able to handle the myriad of interactions that happen at any time and in complex interaction topologies. In these contexts, the heavy load penalties are paid by the servers that have to handle large numbers of requests. With the appearance of virtualization technology, servers are prone to receiving higher numbers of requests due to the node consolidation that takes place at data centers. Despite some solutions for systems with powerful computation nodes, it must be considered that more an more sensors of a CPS are used as service providers or/and data input to the functionality executed in a server farm or a data center. Therefore, it is needed to explore the benefits of middleware to be applied to distributed embedded devices in a CPS context to balance the load of the server farms, i.e., ensuring timely operation and bounded message delivery under different conditions.

A server farm is consisting of a collection of a number of computing nodes (also named hosts or servers) with a front-end high-speed dispatcher. Incoming request are immediately dispatched to one of the specific servers. Advantages of using server farms are price and flexibility of the computational capacity. In order to support the request bulks that may occur at given points in time, balancing mechanisms are considered in order to properly assign request services to different nodes/threads/entities servicing the clients.

In this work, we design a solution for balancing the load of a distributed server that communicates with clients using middleware. There are some widely adopted communication solutions such as DDS [23][24] for distributed real-time systems, ICE [29] for distributed embedded systems, or even Java-based technology such as RMI for mainstream distribution in environments with few (or no) resource limitations and timeliness requirements.

The proposed solution presents a modular, technology-independant architecture design that makes intensive use of multi-threading to balance incoming requests over a server farm, scaling horizontally under high overload conditions, and could be used in several environments and contexts: real-time or not. This will be the first step in a long march of research and work that will lead to a scalable and intelligent software architecture for server balancing, adaptable for many existing middleware solutions, that will be evaluated under CPS requirements. As CPS are intensive in the use of wireless nodes that sense and actuate on the environment, the middleware design

contains a specific entity (*servicing unit*) that encapsulates the provided functionality in general, whether local to a server or provided by a remote sensor. For the latter, this unit can be used as a gateway to communicate with sensor nodes.

The structure of the paper is as follows. Section 1 motivates the work. Section 2 presents the related work. Section 3 describes and discusses the proposed load-aware middleware architecture design . Section 4 explains the laboratory tests done and its results. 5 presents the conclusions and further steps in the work. Finally, 5 includes references and acknowledgments related to this investigation.

2 Load Balancing

2.1 Mainstream Techniques to Load Balancing

Load balancing is required in a number of domains such as popular websites that host content of interest or cloud hosting sites. These can handle heavy and continued request bursts from (typically) remote clients. For distributing incoming jobs on the several servers, dispatchers are used. Dispatchers use different scheduling algorithms whose role is extremely important in the achieved efficiency of the system. In [27], scheduling alternatives are proposed for increasing server farms performance. Typically, distributed load balancing for cloud computing environments as classified in [20] do not rely on a centralized scheduling facility, but distribute the decision among selected nodes.

Directory-based solutions such as LDAP (Light weigh Directory Access Protocol) or now OUD [22] (Oracle Unified Directory) are also related; however, these are clearly focused solely on enabling authorized access to the computation facility, including storage, synchronization, proxy and virtualization.

In general, server load balancing solutions are of two main types [25] [26]: (i) transport-level load balancing, such as the DNS-based approach or TCP/IP-level load balancing that acts independently of the application payload, and (ii) application-level load balancing that uses the application payload to make load balancing decisions. TCP/IP server load balancers operate on low-level layer switching. This is the case of Linux Virtual Server (LVS) [31] that is a specific OS service for balancing. It is highly scalable and highly available servers built on a cluster of actual server nodes that appear to the outside world as a single virtual server. The load balancer forwards the incomming requests on a TCP connection to the real servers, typically using a round robin scheme. One disadvantage is that the balancer runs a modified kernel, patched to include IP Virtual Server (IPVS) code. This is avoided in our solution.

DNS (Domain Name Service)-based load balancing is one of the early server load balancing approaches. It associates an IP addresses with a host name and allows several real addresses to be associated to a host name. It uses a round robin strategy to rotate the assignment of actual servers. Although it is simple to implement, the DNS approach has serious drawbacks. As clients tend to cache the DNS queries, if

a server crashes, the client cache as well as the DNS server will continue to contain a useless server address.

Our solution differs as it is not transport-based nor application-level as defined above, but a hybrid approach. On the one side, it is application-level as we use the infrastructure provided by the communication backbone (an object oriented middleware) to provide information on the requesting entity that may allow to guide the selection process. On the other side, we use information about the available resources (including the transport-level) in order to make efficient use of the computing resources across servers.

Scheduling policies are needed in a balancer design to determine the order of request queuing and dequeuing, and the dispatching parameters. Their goal is to minimize the mean service time of requests, preserving fairness (if priority is not used) and avoiding starvation. Static algorithms are the fastest solution as no overload due to decision logic is present; however, they are bound to the same dispatching decision, regardless of how optimum or poor it may become. FCFS (fist-come-first-served) is an example of a non-preemptive static algorithm that ensures fairness, but penalizes short processes. Dynamic algorithms have the potential to outperform static algorithms as they use some state information, at the cost of incurring in overhead due to the additional logic. Examples of dynamic algorithms are RR (round robin) and WRR (weighted round robin). These are essentially non-preemptive algorithms, but time-sensitive domains such as real-time may require to enforce preemption to handle high priority requests.

2.2 Resource Constrained Systems

Load balancing has been applied since the early years of multiprocessors by providing fixed redirections of communications to specific cores. In embedded systems, the software architecture influences the temporal behavior [4]. Precisely, in embedded multimedia consumer electronics, balancing has been ensured due to the existance of resource managers such as mobile nodes [18] built with the basic operating system primitives to account processor usage, memory consumption, or network usage [2], among others. This is the case of resource manager architectures such as [9] that apply task priorization [6, 10, 17]. Other contributions deal with the dynamic execution [3, 11] such as scheduling of services for reconfiguration in distributed real-time systems [5, 12, 13] or dynamic adaptation in CPS [1, 8].

In the current days, with the push of the cloud computing paradigm where powerful and highly populated data centers have to be managed different issues appear that are especially challenging for specific environments such as real-time [7]. The current middleware solutions[14] concentrate on providing the low level communication facilities to transmit data and messages from source to destination in different ways (e.g. client/server [29] or data-centric [23]) but they do not integrate specific balancing logic.

Our approach overcomes this drawback by considering the presence of a repository with information about the specific servicing units and their deployment in the system. This repository has application-level semantics, e.g., the operations/services provided to clients. Consequently, it is not information that should be integrated in the operating system for abstraction and performance reasons. Our approach builds the balancing logic as part of the communication middleware and outside of the operating system.

3 Design of the Load-Aware Middleware

This section presents a hybrid approach to balancing as it combines application- and transport-level mechanisms in a communication middleware. The design based on a middleware is located at application-level as we use the infrastructure provided by the communication backbone to feed requests with client information; this information can be used by the balancer to decide on how to forward the request. The transport-level part of our solution relates to the presence of a resource monitor on the nodes that periodically sends information to the load balancer node. Resource information includes statistical processor usage of the nodes and physical resources such as transport information.

3.1 Overview and Main Entities

The high level view of the middleware design is described in Figure 1. It presents a representative example showing the different types of middleware entities and an example deployment in different nodes of the server farm. The middleware design contains the following components software entities.

The **Clients** is a middleware component located at the client side. It allows the interaction with the remote server farm by enabling requests for service. In practice, it will be implemented as a simple driver integrated in the side of the client application. It will contain the stubs and proxies needed to reach the server side of the middleware that is installed in the nodes of the server farm.

The **Servicing units** are autonomous units that receive the operation requests from the clients. These requests trigger the execution of some code at the server farm that produces a response that will be sent back to the requesting client. *Servicing units* are software logic units that can be physically located on a single machine or distributed over different server nodes in the farm over the same network. They can operate autonomously and independently from others. Selected servicing units can also act as gateways to wireless nodes that sense and actuate on the environment. Servicing units contain a *resource monitor* that periodically informs the balancer node of the usage of the physical resources such as the processor, memory, thread pools, and transport connections.

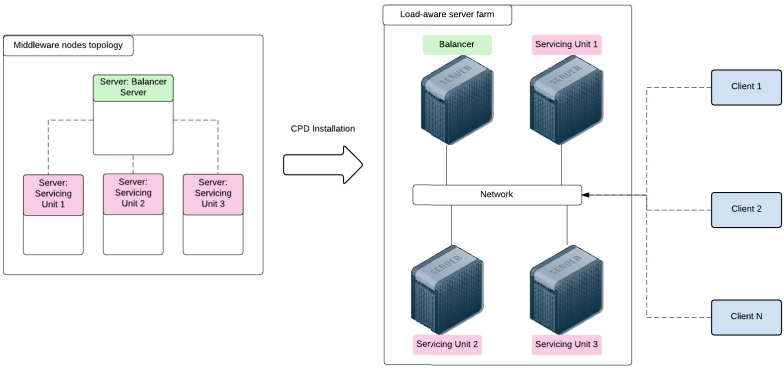


Fig. 1 Middleware node types. Example of topology installed in a Server farm

The **Balancer** is an autonomous unit that contains the logic for deciding on how to handle a specific client request. The decision will consist of forwarding the request among the different servicing units that are available in such a way that the load that these requests generate is balanced among the farm nodes. A *balancer* acts like a facade over all the available servers being a black-box from an external client point of view. It is aware of all the operations implemented by the servicing units, therefore as the entry point of all requests, it forwards them to the indicated servicing units at specific physical nodes.

3.2 The Balancer

Upon start up, a servicing unit informs the balancer of its presence in the system and of the operations that it provides to clients. From this point on, the servicing unit will receive requests from clients that will be forwarded by the balancer. Every server will create a local *registry* as a directory that stores all the required information necessary to set up a communication between the remote entities as the server and the balancer.

The balancer executes a periodic task running every δ units of time (typically in the order of hundred milisecons) on its default mode, and it can be configurated at initialization given the specific needs of applications. This task checks the health-status of each servicing unit that is tracked by the balancer by sending *keep-alive* messages. If a servicing unit becomens unavailable, this will be detected by the balancer after the health status monitoring period, and no further requests will be forwarded to it. The balancer will create a local registry to store all communication stubs with the necessary information to establish communication with the clients.

The balancer applies a *round-robin* policy to forward client requests just to those servicing units running. When a new server is started it immediately sends a request

to the balancer to mark itself as an alive server available to execute new incoming requests. Thereafter, if one of the balancer *keep-alive* fails, the failing servicing unit is removed from the list of available units.

The balancer chooses an available servicing unit to dispatch the incoming requests with a *round-robin* algorithm . The Balancer will store a list with the available servers in the cluster. Plus, it will store also a numeric variable that will point to a given position from the server list. When receiving a new request from a client, the balancer will increase this numeric variable value and then will get the servicing unit stored in that list position. The balancer will then dispatch the request to that selected server, returning its response and forwarding it from the balancer to the requesting client.

3.3 Software Design

This section describes a specific software design of the middleware containing the balancing logic and using Java language and middleware technology based on RMI (remote method invocation). Following the main classes are given in a Java based design.

MonitoringServer expresses the functionality of a servicing unit. It is the basic interface that specifies the available operations that are implemented in a servicing unit. A `MonitoringServer` implementation includes a `getId` remote method to return its server identifier, and a `ping` remote method necessary for balancers to execute keep-alive policies over all tracked servicing units. It is provided below:

```
public interface monitoringServer extends Remote{
    String getID();
    String ping();
    long changeRecording(int, int);
    long sendLog(int);
}
```

Operation `changeRecording` is given as an example of a specific operation provided by a servicing unit that will be exemplified in the validation section.

MonitoringServerImpl. Implementation of a `MonitoringServer` interface. This implementation uses Java RMI libraries to interact with the balancer (BalancerServer). As the design aims at being flexible, the reference to a specific balancer is contained. It is important to mention that this modular and flexible design should be separated from the specific implementation that, in the present paper, is a centralized one. The existance of multiple balancers would allow to implement complex balancing algorithms with distributed information (outside the scope of this paper).

BalancerServer. It is the basic interface for the balancer entity that is also implemented as a remote object. This interface extends the `MonitoringServer` interface by adding two methods to address both (*i*) the start of the monitoring of servicing units (`addMonitoringServer()`) and (*ii*) stopping this monitoring

(`removeMonitoringServer`) in a given balancer. When a new servicing unit is available, it executes a remote call to register itself to the BalancerServer to be tracked. From that moment on, it is available to clients and ready to receive incoming requests. When the servicing unit is shutdown correctly it will execute a remote call to be removed from the list of tracked servicing units.

BalancerServerImpl. It is the RMI implementation of a `BalancerServer` that interacts with servicing units. It will receive incoming requests from clients, select an active servicing unit forward the requested operation to it with a new remote call over the chosen servicing unit.

4 Experimental Validation

This section validates the proposed middleware design by implementing the above design in Java for a specific video-surveillance example. Even if Java probably is not the best solution to solve programmatically a time-sensitive application, RMI is however an interesting alternative for architectural validation do to the relative easy construction of fast prototypes; also, it has been used in many research investigations such as [19], concluding that the performance of RMI in multi-threaded environments suited well for a theoretical CPS scenario. Most of other relevant video applications have used DDS such as [15, 16].

Servicing units of the middleware are connected to the wireless nodes (the IP cameras). Our deployment has `n` distributed cameras connected by a wireless link camera servicing units that interface with the infrastructure. Servicing units interfacing to the cameras include operations to control their execution (i.e., video data retrieval, selection of a new video camera to view its recording in real-time, and camera status retrieval). In addition, each video servicing unit exposes this mentioned functionality to video clients. Clients are able to interact with the video-cameras ecosystem using the server farm as a kind of gateways through the servicing units.

For our tests we have deployed a video-surveillance topology composed by **one balancer** and **two servicing units**. For client operations we have developed a simple Java client class named `ClientOperator` that interacts with the balancer using Java RMI. This client executes `send_log` and `chg_recording` remote methods into `BalancerServer`.

Tests and measurements were done in a lab setting on a core i7 3.0 Ghz processor with 16GB RAM DDR3.Results from the tests are summarized now. For **uni-processor**, several executions were carried out in a uni-processor environment running `ClientOperator` side. The test yielded an average response time of 102ms for the servicing units, and an average response time of 104ms for the balancer. The average response time for clients was 105ms. For **uni-processor multithreaded**, running the client side `ClientOperator` was done in the presence of interfering threads that performed similar operations over the balancer. Several executions were performed and results obtained an average response time for service unites of 102ms, and an average time of time of 104ms for the balancer. The client response times were 105ms on average terms.

5 Conclusions

We have presented a design for a middleware that performs load balancing in distributed systems that involve clients requesting the services contained in a server farm. The middleware addresses server farms that service requests from clients with respect to both (a) functionality contained exclusively in the server farm, and (b) functionality offered by sensor nodes that monitor the physical environment. The architecture is based on a hybrid load balancing model that can be used in different scenarios and application domains. The design has demonstrated to be flexible by support N servicing units balanced by a single balancer, supporting the system to scale horizontally, avoiding single points of failure from the servicing units side. Further investigations with ellaborate on the design and implementation of a Cluster system with various balancers in a master-slave or a replica-set role-status synchronization model. Furthermore, we have tested this design in laboratory with successful results over multi-threaded and multi-process plus multi-threaded scenarios, with reduced response times.

Acknowledgment This work has been partly funded by projects REM4VSS (TIN2011-28339) and Open and Dynamic CPS (EUIN2013-51179), both granted by the Spanish Ministry of Science.

References

1. Bersani, M.M., García-Valls, M.: The cost of formal verification in adaptive CPS. An example of a virtualized server node. In: Proc. of 17th IEEE High Assurance Systems Engineering Symposium (HASE), January 2016
2. Breuer, P.T., García-Valls, M.: Raiding the Noosphere: the open development of networked RAID support for the Linux kernel. Software Practice and Experience **36**(4), 365–395 (2006)
3. Romero, J.C., García-Valls, M.: Scheduling component replacement for timely execution in dynamic systems. Software: Practice and Experience **44**(8), 889–910 (2014)
4. Duenas, J., Alonso, A., Lopes Oliveira, W., Garcia, M., Leon, G.: Software architecture assessment. In: Software Architecture for Product Families: Principles and Practice. Addison-Wesley (2000)
5. García-Valls, M., Fernández Villar, L., Rodríguez López, I.: iLAND: An enhanced middleware for real-time reconfiguration of service oriented distributed real-time Systems. IEEE Transactions on Industrial Informatics **9**(1), 228–236 (2013)
6. García-Valls, M., Alonso, A., de la Puente, J.A.: A dual priority assignment mechanism for dynamic QoS resource management. Future Generation Computer Systems **28**(6), 902–911 (2012)
7. García-Valls, M., Cucinotta, T., Lu, C.: Challenges in real-time virtualization and predictable cloud computing. Journal of Systems Architecture **60**(9), 726–740 (2014)
8. García-Valls, M., Perez-Palacin, D., Mirandola, R.: Time sensitive adaptation in CPS through run-time configuration generation and verification. In: Proc. of 38th IEEE Annual Computer Software and Applications Conference (COMPSAC), pp. 332–337, July 2014
9. García-Valls, M., Alonso, A., de la Puente, J.A.: An architecture of a quality of service resource manager middleware for flexible multimedia systems. In: Proc. of 3rd Int'l Workshop on Software Engineering and Middleware (SEM 2002), Orlando (Florida), USA. Lecture Notes in Computer Science, vol. 2596, pp. 36–55 (2002)

10. García-Valls, M., Alonso, A., de la Puente, J.A.: Mode change protocols for predictable contract-based resource management in embedded multimedia systems. In: Proc. of IEEE Int'l Conference on Embedded Software and Systems (ICESS), May 2009
11. García-Valls, M., Basanta-Val, P.: A real-time perspective of service composition: key concepts and some contributions. Journal of Systems Architecture **59**(10), 1414–1423 (2013)
12. García-Valls, M., Basanta-Val, P.: Comparative analysis of two different middleware approaches for reconfiguration of distributed real-time systems. Journal of Systems Architecture **60**(2), 221–233 (2014)
13. García-Valls, M., Uriol-Resuela, P., Ibánez-Vázquez, F., Basanta-Val, P.: Low complexity reconfiguration for real-time data-intensive service-oriented applications. Future Generation Computer Systems **37**, 191–200 (2014)
14. García-Valls, M., Baldoni, R.: Adaptive middleware design for CPS: considerations on the OS, resource managers, and the network run-time. In: Proc. of 14th ACM Workshop on Adaptive and Reflective Middleware, December 2015
15. García-Valls, M., Basanta-Val, P.: Usage of DDS Data-Centric Middleware for Remote Monitoring and Control Laboratories. IEEE Transactions on Industrial Informatics **9**(1), 567–574 (2013)
16. García-Valls, M., Basanta-Val, P., Estévez-Ayres, I.: Adaptive real-time video transmission over DDS. In: Proc. of 8th IEEE International Conference on Industrial Informatics (INDIN), July 2010
17. García-Valls, M., Basanta-Val, P., Estévez-Ayres, I.: Real-time reconfiguration in multimedia embedded systems. IEEE Transactions on Embedded Consumer Electronics **57**(3), 1280–1287 (2011)
18. García-Valls, M., Crespo, A., Vila, J.: Resource management for mobile operating systems based on the active object model. International Journal of Computer Systems Science & Engineering **28**(4), 195–205 (2013)
19. Halash, E.: Offloading CPU Intensive Applications to the Cloud Using Java RMI. Wayne State University 2009. Technical report (2003)
20. Katyal, M., Mishra, A.: A Comparative Study of Load Balancing Algorithms in Cloud Computing Environments. International Journal on Distributed and Cloud Computing **1**(2), December 2013
21. The Linux Virtual Server Project, January 2015. http://www.linuxvirtualserver.org
22. Oracle: Oracle Unified Directory. White Paper, December 2013
23. Object Management Group: Data Distribution Service for Real-Time Systems. OMG, 1.2 formal/07-01-01 edition, January 2007
24. Pardo Castellote, G.: OMG data-distribution service: architectural overview. In: 23rd International Conference Distributed Computing Systems Workshops, May 2003
25. Roth, G.: Server load balancing architectures, Part 1: Transport-level load balancing JavaWorld, October 2008
26. Roth, G.: Server load balancing architectures, Part 2: Application-level load balancing JavaWorld, October 2008
27. Saboori, E., Mohammadi, S., Parsazad, S.: A new scheduling algorithm for server farms load balancing. In: 2nd International Conference on Industrial and Information Systems (IIS), pp. 417–420, July 2010
28. Wickramasuriya, J., Venkatasubramanian, N.: A directory enabled middleware framework for distributed systems. In: IEEE International Workshop on Objectoriented Real-time Dependable Systems (WORDS 2003), Guadalajara, Mexico (2003)
29. ZeroC: Internet Communications Engine. Ice Manual (2015). http://docs.zeroc.org/
30. Wolf, W.: The Good News and the Bad News (Embedded Computing Column). IEEE Computer **40**(11), November 2007
31. Zhang, W., Zhang, W.: Linux Virtual Server Clusters: Build highly-scalable and highly-available network services at low cost. Linux Magazine (2003)

Adaptive Message Restructuring Using Model-Driven Engineering

Hang Yin, Federico Giaimo, Hugo Andrade, Christian Berger and Ivica Crnkovic

Abstract Message exchange between distributed software components in cyber-physical systems is a frequent and resource-demanding activity. Existing data description languages simply map user-specified messages literally to the system implementation creating the data stream that is exchanged between the software components; however, our research shows that the exchanged information is often redundant and would allow for runtime optimization. In this paper, we propose a model-based approach for adaptive message restructuring. Taking both design-time properties and runtime properties into account, we propose to dynamically restructure user-specified messages to achieve better resource usage (e.g., reduced latency). Our model-based workflow also includes formal verification of adaptive message restructuring in the presence of complex data flow. This is demonstrated by an automotive example.

Keywords Domain-specific language · Model-based engineering · Automotive software · Verification · UPPAAL

1 Introduction

Self-driving vehicular technology is a promising evolutionary trend in the automotive industry. Along with the recently released Volvo XC90, an interdisciplinary research project has been launched, aiming to develop a self-driving vehicle platform to tackle complex inter-city traffic scenarios. It is envisioned to eventually realize

H. Yin(✉) · F. Giaimo · H. Andrade · I. Crnkovic
Chalmers University of Technology, Gothenburg, Sweden
e-mail: {yhang,giaimo,sica,crnkovic}@chalmers.se

C. Berger
University of Gothenburg, Gothenburg, Sweden
e-mail: christian.berger@gu.se

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_67

"CampusShuttle", a self-driving vehicle connecting the two Chalmers campuses. This project involves an iterative and agile collaboration between both academic and industrial partners for defining the required sensor setup and identifying requirements for the algorithmic data processing. Due to the several expertise and competencies involved in the project it is important to find a flexible yet robust way to specify and maintain the data model from the onboard vehicle interface and the different sensors being integrated over time.

As a general requirement in cyber-physical systems, software components distributed in the vehicle must be able to exchange data by transmitting user-specified messages. Existing popular frameworks such as AUTOSAR [7], Automotive Data and Time-Triggered Framework (ADTF) [6], and ROS [3] realize a direct mapping between messages in the user layer and messages in the system layer. Such a direct mapping may be inefficient due to redundant information during transmission, thus suffering from a number of limitations:

- No flexibility of switching different message marshalling schemes [8, 11] for better resource usage upon changing requirements. Hence, there is little room for runtime adaptation;
- Complex messages, which carry for example many data fields originating from various data resources, might slow down communication between software components, since a sender simply waits before sending a complete message;
- This does not facilitate flexible software modularization and reuse, as a message structure defined at design-time assumes a specific interfacing to software components that also results in challenges for runtime adaptation.

To overcome these limitations adopting a Model-Driven Engineering (MDE) [10] approach, we propose a model-based technique to specify and enable adaptive message restructuring at runtime to achieve better resource usage. Adaptive message restructuring involves non-trivial mechanisms to reach a certain Quality of Service level regarding communication latency, bandwidth resource usage, or other metrics. However, we must ensure that the restructuring process is correct with respect to chosen properties, that may be for example its timeliness concerning the specified timing boundaries. Hence, we use model checking as formal verification of these mechanisms before the actual implementation. This boils down to the following two research questions:

RQ1: *How would a model-based approach enable adaptive message restructuring?*
RQ2: *How can rigorous verification guarantee a correct adaptive message restructuring scheme at an early design phase?*

Our adaptive message restructuring scheme not only takes advantage from design-time properties, but also considers runtime properties. Both types of properties can be either domain-independent or domain-specific. Design-time properties are static properties known at design time, e.g., the composition of fields in a message, as the data type in each field of a message is fixed and does not change during the execution of a system. Based on the number and type of fields of each message to be

transmitted, we could utilize our previous research results [8] to select the optimistic data marshalling approach to facilitate data exchange between software components and save bandwidth resource.

In contrast to design-time properties, runtime properties are not known until runtime, such as message transmission latency, timeout, and message arrival/drop rate. Runtime properties can be used to further improve the performance of message restructuring. For instance, timeouts could be introduced to transmit incomplete messages to meet stringent timing requirements.

In addressing the two research questions, the contributions of this work are:

- A model-based workflow for our open source software environment OpenDaVINCI [9] to enable adaptive message restructuring taking both design-time properties and runtime properties into account.
- The application of a model checking approach to formally verify the adaptive message restructuring at runtime to ensure the production of a result in accordance to the desired properties, e.g. timing properties.

The remainder of the paper is structured as follows: Section 2 discusses related work. Section 3 sketches our model-based workflow for achieving adaptive message restructuring. Our domain-specific language for model transformation is described in Section 4. Section 5 describes our model checking approach for formal verification. Section 6 revisits the identified research questions and discusses threats to validity before the paper is concluded in Section 7.

2 Related Work

Data marshalling plays a key role in complex distributed software systems. It enables the communication between different threads and processing units, providing means to standardize message passing even when different technologies are being used on the sender and on the receiver sides.

Different marshalling approaches have been proposed, such as Google Protobuf [1], which provides support for several popular languages. In Protobuf, data structures are specified in a textual representation, from which concrete language bindings are generated to enable communication between software components. Another widely used approach is LCM [2], which is targeted at real-time systems and thus, provides lightweight solutions to message composing. LCM has shown to be especially efficient when there are several fields of double data types and integral data types in a single message [8].

Particularly in the domain of self-driving vehicles, a large amount of data is captured by sensors, processed, and transmitted between different software components, resulting in high bandwidth consumption. However, our previous work shows that the actual differences between two consecutive messages is rather small, thus contributing to data redundancy. This observation motivated our self-adaptive

marshalling approach Delta [11], which has been realized in the OpenDaVINCI framework [9] to improve resource utilization in aforementioned scenarios by sending only the differences between consecutive messages.

Other modeling approaches in the automotive domain (e.g., AUTOSAR [7] and ADTF [6]) and in the robotics domain (e.g., ROS [3]) only give 1:1 mapping between the user-defined specifications and the actual messages in use by the system. For example, ADTF enables the modeling of behavior and specification of the data to be exchanged between two entities, but it does not support different message formats in the code generation phase.

3 The Model-Based Workflow

Specifying messages to be exchanged between distributed software components is in practice realized with a Domain-Specific Language (DSL). This modeling approach abstracts from concrete implementation details at low-level and enables system designers to focus on the semantic relations between communicating entities. In vehicular research projects, the CAN bus [12] is still a widely used network providing an established communication means. System architects define signals representing for example measurements from the vehicle's sensorial apparatus.

As the generic structure of a CAN message allows the specification of a few of such signals into one enclosed byte sequence, there is apparently a need to compose several of these low-level CAN messages to semantically more expressive high-level messages to be exchanged in the middleware. For this purpose, we have defined a textual DSL to describe the mapping that binds several CAN signals of interest to a high-level message. The definitions of the signals of interest and the relative mappings comprise messages in the user layer.

As outlined earlier, in communication frameworks currently in use, there is no difference between the messages defined in the user layer from the system architect and the ones in the system layer, i.e. the messages actually exchanged between software components. This literal structure in the system layer might result in an inefficient resource usage like bandwidth consumption due to information redundancy as described in our earlier work [11].

In this paper we introduce a model-based approach that enables adaptive restructuring of user-defined messages during the transformation from model descriptions to a corresponding implementation. Shown in Fig. 1, the adaptive message restructuring process exploits both design-time properties and runtime properties to improve resource usage at runtime. For example, the proposed restructuring might (1) merge two smaller messages into a larger message; (2) split a complex message into several smaller messages; or (3) exchange some data fields among multiple messages. A "restoring" phase will transparently take place after the new message is received to restore the reconstructed message.

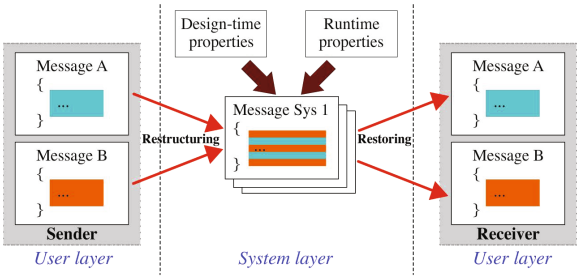


Fig. 1 Model-based workflow for adaptive message restructuring

4 DSL and Model Transformation

In order to define the needed CAN message mappings that are of interest, a textual representation using our DSL is made available at design time; then, a code generator will use this file's content to generate the classes implementing the mapping from CAN messages to high-level messages. The DSL has been implemented using the Xtext framework [5] that allows the definition of customized grammars and the automatic generation of a corresponding parser. The code generator is realized with Xtend [4], a functional dialect of Java, that will access the tokens from Xtext and generate the appropriate code to handle the signals' parameters. The results of the code generator are state machines that are triggered by incoming CAN messages to create higher-level messages.

A small snippet from a user-supplied .can file is shown in Fig. 2. The file starts with a global scope where information about the signals of interest is collected, followed by a set of mapping specifications that collect these signals together in semantically significant units. The signal specifications include details such as bit displacement in the physical CAN message, length of the field, linear operation parameters to be performed on the "raw" value (e.g. coefficient or the addition of a constant). The mapping specifications comprise a given name and a list of interested signal names, each followed by a signal ID to identify the signal within the mapping.

```
1    wheelspeed.frontleft    in 0x123 at bit 0 for 16 bits is big endian multiply by 0.01 add 0 within range [0, 200];
2    wheelspeed.frontright   in 0x123 at bit 16 for 16 bits is big endian multiply by 0.01 add 0 within range [0, 200];
3    wheelspeed.rearleft     in 0x124 at bit 32 for 16 bits is big endian multiply by 0.01 add 0 within range [0, 200];
4    wheelspeed.rearright    in 0x124 at bit 48 for 16 bits is big endian multiply by 0.01 add 0 within range [0, 200];
5
6    mapping WheelSpeed {
7        wheelspeed.frontleft : 1;
8        wheelspeed.frontright : 2;
9        wheelspeed.rearleft : 3;
10       wheelspeed.rearright : 4;
11   }
```

Fig. 2 Example for a user-supplied CAN message mapping specified in a ".can" file

On the code generator side, the Xtend environment transforms tokens from Xtext to generate the resulting C++ code such that for each mapping a state machine is

derived to (1) collect the CAN messages containing signals of interest, (2) extract them from the raw content of the CAN message payload, (3) transform them if necessary via linear mathematical operations, and (4) create a high-level message reflecting the desired mapping. These state machines will implement adaptive message restructuring by exploiting design-time and runtime properties. The implementation is generating the messages-related classes and test suites based on the user specifications expressed via our DSL language. The source code can be accessed at http://code.opendavinci.org/.

5 Modeling and Formal Verification of Adaptive Message Restructuring

Our model-based workflow includes a model checking approach using UPPAAL [13], which allows modeling and formal verification of complex algorithms. The purpose of model checking is twofold: (1) to ensure correct message restructuring at an early design phase before the actual implementation, and (2) to reuse our UPPAAL model as a design sketch guiding the C++ implementation in OpenDaVINCI. In this section, we use a simple automotive example to elaborate on our model checking approach.

5.1 Domain Analysis

Restructuring messages from low-level CAN messages to high-level automotive messages is of importance to realize communication between distributed software components. In principle, the restructuring of each automotive message can be described as a state machine that maps one or multiple CAN messages to a single automotive message as outlined in Section 4. Concurrent restructuring of multiple automotive messages by multiple state machines is often desired to facilitate the communication between software components.

Fig. 3 illustrates three automotive messages A, B, C composed by CAN messages from two CAN busses CAN 0 and CAN 1. These automotive messages should be transmitted to other relevant software components. Message A contains wheel speed information provided by CAN messages with ID 0x10 from CAN 0, while Message C contains GPS time provided by CAN messages with ID 0x13 from CAN 1. Message B has a more complex structure, as it includes both vehicle motion data (0x11 and 0x14 from CAN 0) and vehicle location data (0x12 and 0x15 from CAN 1). S_A, S_B, and S_C in the figure are state machines for composing messages A, B, C, respectively.

Suppose another software component expects information of both vehicle motion and position in message B every 5 ms. Since S_B expects CAN messages from both CAN busses, the transmission of vehicle motion data may be delayed by position data or the other way round. Practical scenarios that require precise control of the

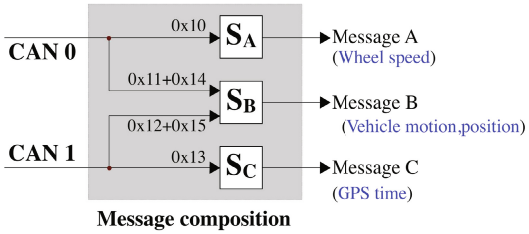


Fig. 3 An example of concurrent message restructuring from multiple CAN busses

system requires low message transmission latency, so receiving either vehicle motion data or position data in time is better than receiving both late. Then a timeout could be introduced in S_B to reduce transmission latency. For instance, S_B can start a 5ms timer after receiving 0x11 and 0x14. If S_B cannot receive both 0x12 and 0x15 before the timeout, it composes an incomplete message B with only vehicle motion data to be sent to the recipient, which can construct a complete message B after receiving position data at a later time.

5.2 Model Overview

We have modeled the example in Fig. 3 using UPPAAL. The model consists of a set of timed automata, which jointly implement the message stream of each CAN bus and the behavior of each state machine.

Both CAN busses are modeled in the same fashion. Fig. 4(a) shows the model of CAN 0 that periodically generates a CAN message by a channel *message[0]*. A clock x is used to express timing constraints. The guard $x >= 1$ of transition 1 and the invariant $x <= 1$ of the initial state **Ready** (marked with a double-circle) define a message arrival rate of 1 time unit from CAN 0. The model maintains a list of CAN IDs {10,11,14,0}, where 10 represents 0x10 for S_A, 11 and 14 are expected by S_B, and 0 represents a message not used for composing any automotive message in the example. Before generating a CAN message by transition 2, the model randomly selects an ID from the list by transition 1.

Taking S_B as an example, we model its behavior by three automata. Apart from S_B itself, there are two internal state machines S_B^0 and S_B^1 for collecting messages from two CAN busses separately. Fig. 4(b) shows the model of S_B^0, which expects 0x11 and 0x14 in the right order. The arrival of the first 0x11 enables transition 1 and a subsequent arrival of 0x14 enables transition 4. Thereafter, a function *updateLocalBuffer()* sets a boolean variable *allReceived[0]* to true, indicating that vehicle motion data is ready for restructuring. This further leads to transition 1 of the model of S_B in Fig. 4(c), as a timer of 5 time units is started, indicated by the invariant $x <= 5$. If another boolean variable *allReceived[1]* becomes true before timeout, i.e., S_B^1 has received both 0x12 and 0x15 in the expected order, the restructuring of message is

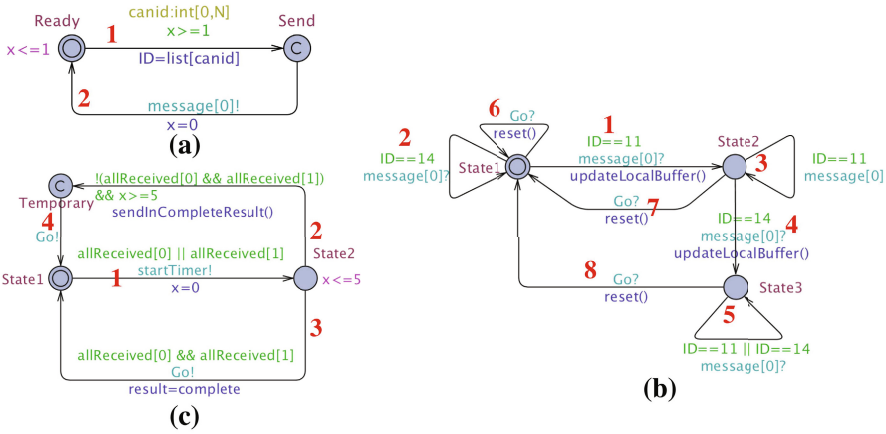


Fig. 4 The UPPAAL models of (a) CAN 0, (b) S_B^0, (c) S_B

complete by transition 3. Otherwise, an incomplete message B only containing vehicle motion data is transmitted upon timeout by transition 2. The complete UPPAAL model is available at https://github.com/se-research-studies/static-dynamic-views.

5.3 Formal Verification

Formal verification in UPPAAL is achieved by checking if a set of properties (formulated in the UPPAAL query language) are satisfied with regard to a particular model. UPPAAL is able to automatically return the verification result of each property. If a property is not satisfied, a counter example will be provided, potentially implying a design flaw that needs to be fixed. Based on the model described in Section 5.2, we formulate the following three types of properties:

P1. *A[] not deadlock*: The model is deadlock-free.
P2. *!ReceiverB.State1 && (forall (i:int[0,1]) allReceived[i]) –>ReceiverB.State1 && ReceiverB.result==complete*: If S_B receives all the expected CAN messages from CAN 0 and CAN 1 before timeout, the restructuring of Message B will be completed in time.
P3. *E<> (ReceiverB.Temporary && !allReceived[i])*: There exists such a scenario where not all the expected messages from a CAN bus arrive upon timeout.

P1 and P3 jointly imply that the message restructuring can be completed even if an expected message from a CAN bus never arrives. All the three properties above are satisfied by the verification. Furthermore, the verification result of P3 also includes a possible execution trace attesting the satisfaction of P3.

6 Discussion

There are a number of benefits that can be achieved by utilizing model-based strategies to a message restructuring solution. First, we improve the workflow by restructuring the user-defined mappings according to two different aspects: (1) design-time properties (e.g., for determining which marshalling approaches to use); and (2) runtime properties (e.g., latency restrictions implemented with concurrency). These properties allow the recomposition of messages to occur in a more effective way, thus improving utilization of resources such as bandwidth. Further, by adding a transparent solution at system level, the user is not required to sacrifice semantic modeling preferences in order to achieve savings in bandwidth consumption. Instead, the model-based workflow solution can be implemented utilizing context-aware models which can dynamically generate restructured messages that are suitable for different scenarios. In addition, by using a model-based approach, the overall solution can be modeled and formally verified already during a system's architecture phase.

Marshalling approaches are commonly implemented in the context of cyber-physical systems, where inter-process communication is a key aspect of the overall system's performance. These systems are often subjected to restrictions to satisfy specific requirements, such as time limits to deliver a computation result, or certain level of correctness to be achieved due to the nature of the application domains. For some domains such as health care or self-driving vehicles, there are safety and security requirements that must be satisfied to avoid severe consequences. This makes thorough system verification and validation desirable. Model checking is a valuable verification technique because flaws can be detected even before the implementation phase. This plays a strategic role for the success of the system, especially in the case of big and complex projects. In addition, a robust and tool-supported process of verification and validation enables automated testing of complex scenarios.

Our UPPAAL model described in Section 5.2 is based on a small but practical automotive example. We could foresee that both model complexity and verification time will grow when the number of CAN busses or the number of expected messages of a state machine increases. However, the additional effort for model checking should be marginal, as the model is generic and preserves the same pattern as long as the adaptive message restructuring scheme remains unchanged. Our Xtext/Xtend environment is already able to automatically generate UPPAAL models for messages from a single source based on the message specification in our DSL. It is our future work to extend our Xtext/Xtend environment further to cover messages from multiple sources described in Section 5.

We are aware of the potential threats to validity to our study [14]. Regarding *construction validity*, the properties for our UPPAAL model are based on our experience from previous automotive projects; though, an independent confirmation would be recommended. Furthermore, we only verified the properties for our UPPAAL model, as a code generation thereof towards C++ is missing. Concerning *internal validity*, we followed the "design science" research methodology [15] as the development of the self-driving vehicle platform is currently ongoing. Thus, alternative solutions

might have been overlooked due to the nature of this method. Regarding *external validity*, we only considered the automotive domain as evaluation context; as the concept for considering design-time properties is generalizable as shown in our previous study [8], only runtime properties need to be further evaluated using case studies from further domains.

7 Conclusion and Future Work

In this paper, we have introduced a model-based approach for adaptive message restructuring at runtime using Model-Driven Engineering. Our message restructuring scheme exploits both design-time properties and runtime properties, thereby improving resource usage during the communication between distributed software components in cyber-physical systems. We have embedded our previous results in our model-based workflow to utilize design-time properties. Meanwhile, runtime properties from the concrete message exchange are extracted for further adaptation of message restructuring. The correctness of our message restructuring scheme is verified by model checking using UPPAAL.

Future work is required to capture properties from the additional computation required to enable adaptive message restructuring. These constraints are in addition dependent on a concrete execution platform and thus, influence a distributed system's communication as well. Also, another direction for future efforts would be the definition of message restructuring "patterns" to address the general problem of resource optimization.

Acknowledgments This work is conducted in the *RALF3* project funded by the Swedish Foundation for Strategic Research and the *CampusShuttle* project supported by Chalmers University of Technology.

References

1. Google protocol buffers. https://github.com/google/protobuf (accessed April 28, 2015)
2. LCM: Lightweight Communications and Marshalling. http://lcm-proj.github.io (accessed April 28, 2015)
3. ROS - Powering the world's robots. http://ros.org (accessed April 28, 2015)
4. Xtend - Modernized Java. http://eclipse.org/xtend/ (accessed October 23, 2015)
5. Xtext - Language Development Made Easy! http://eclipse.org/xtext (accessed October 23, 2015)
6. EB Assist ADTF (2001). https://www.elektrobit.com/products/eb-assist/adtf/ (accessed October 21, 2015)
7. AUTOSAR. http://autosar.org (accessed October 21, 2015)
8. Andrade, H., Giaimo, F., Berger, C., Crnkovic, I.: Systematic evaluation of three data marshalling approaches for distributed software systems. In: Proceedings of the 15th Workshop on Domain-Specific Modeling (DSM), Pittsburgh, Pennsylvania, USA, p. 5, October 2015

9. Berger, C.: From a Competition for Self-Driving Miniature Cars to a Standardized Experimental Platform: Concept, Models, Architecture, and Evaluation. Journal of Software Engineering for Robotics **5**(1), 63–79 (2014). http://arxiv.org/abs/1406.7768
10. Bézivin, J.: On the Unification Power of Models. Software & Systems Modeling **4**(2), 171–188 (2005)
11. Giaimo, F., Andrade, H., Berger, C., Crnkovic, I.: Improving bandwidth efficiency with self-adaptation for data marshalling on the example of a self-driving miniature car. In: Proceedings of the 1st International Workshop on Software Architectures for Next-generation Cyber-physical Systems (SANCS), Dubrovnik, Croatia, p. 6, September 2015
12. ISO: Road vehicles - Controller area network (CAN). ISO 11898-1:2003, International Organization for Standardization, Geneva, Switzerland (2003)
13. Larsen, K.G., Pettersson, P., Yi, W.: UPPAAL in a Nutshell. International Journal on Software Tools for Technology Transfer **1**(1–2), 134–152 (1997)
14. Runeson, P., Höst, M.: Guidelines for Conducting and Reporting Case Study Research in Software Engineering. Empirical Software Engineering **14**(2), 131–164 (2008). http://link.springer.com/10.1007/s10664-008-9102-8
15. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action Design Research. MIS Quaterly **35**(1), 37–56 (2011)

Towards Modular Language Design Using Language Fragments: The Hybrid Systems Case Study

Sadaf Mustafiz, Bruno Barroca, Claudio Gomes and Hans Vangheluwe

Abstract Cyber-physical systems can be best represented using hybrid models that contain specifications of both continuous and discrete event abstractions. The syntax and semantics of such hybrid languages should ideally be defined by reusing the syntax and semantics of each components' formalisms. In language composition, semantic adaptation is needed to ensure correct realization of the concepts that are part of the intricacies of the hybrid language.

In this paper, we present a technique for the composition of heterogeneous languages by explicitly modelling the semantic adaptation between them. Each modelling language is represented as a language specification fragment (LSF): a modular representation of the syntax and semantics. The basis of our technique is to reuse the operational semantics as defined in existing simulators. Our approach is demonstrated by means of a hybrid language composed of timed finite state machines (TFSA) and causal block diagrams (CBD).

1 Introduction

A typical cyber-physical system (CPS) exhibits complex behaviour that can only be best represented using a mix of continuous (i.e., mathematical differential equations) and discrete (logical computation) models. In particular, to engineer such systems, we need to first study how existing physical laws and quantities, co-exist and interact with logical controllers, which themselves are also bound to the same laws. The engineering of a CPS involves the modelling and simulation of hybrid

S. Mustafiz(✉) · B. Barroca · H. Vangheluwe
McGill University, Montreal, QC, Canada
e-mail: {sadaf,bbarroca,hv}@cs.mcgill.ca

C. Gomes · H. Vangheluwe
University of Antwerp, Antwerp, Belgium
e-mail: claudio.gomes@uantwerpen.be

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_68

models (e.g., combinations of sets of piece-wise continuous functions alternating with events). Moreover, due to the extreme complexity of the CPS, it is often the case that we need to break the models into several orthogonal views, aspects or quantities (e.g., thermal, electric, power), in order to be able to study the behaviour of each of them over time, and only afterwards study their integration and interaction.

It is therefore obvious that the success of building such complex CPS can only be maximized if one first identifies the most appropriate level of abstraction to break the complexity in each of these orthogonal views, and then uses the most appropriate formalisms that somehow realize these abstractions, while bringing to the CPS engineers a valuable engineering toolbox that integrates several CPS modelling and simulation environments. However, we advocate that the CPS engineer should (instead of being yet another language engineer) become a language integrator, by reusing and integrating existing languages into a new language that fits the expressiveness needs of a particular CPS. Unfortunately, the existing techniques for language modularity are still not adequate to fully realize this objective. From our recent work [12], the need for techniques to modularize the design of modelling languages has become apparent to us.

In this paper, we underline new techniques for language modularity based on the concept of Language Specification Fragments (LSF). Such fragments include the syntax (both abstract and concrete), the semantics, and the user-interface behaviour, along with the various interleaving of the elements involved. The LSFs form the basis of reuse and abstraction in language design, and allows the development of new formalisms by re-using and merging existing fragments, leading to reduced efforts on the language designers part. For example, an existing platform-dependent DEVS formalism can be extended by integrating an existing Neutral Action Language in it to compose a neutral DEVS formalism [2]. Such compositions become even more useful when integrating UML formalisms, e.g. Statechart and Class Diagrams with Action Code.

Following the initial ideas of meta-modeling hybrid formalisms [10], we first observe and meta-model a particular hybrid formalism that combines a language for Timed Finite State Automata (TFSA) with a language for Causal Block Diagrams (CBD). We then perform a conceptual de-construction of the hybrid formalism into two (reusable) LSFs, while identifying the composition operations that are required in order to re-construct it back to the original formalism. Contrasting to the co-simulation domain, where the operational semantics of each language is considered as a black-box during language composition (otherwise referred to as semantic adaptation) [7], we follow a white-box approach that allows us to the explicitly model the composition of LSFs by taking into consideration the concepts that rule each of the languages.

This paper is structured as follows: Section 2 gives an introduction to the case study used, Section 3 introduces language specification fragments, and the LSFs used in our case study. Section 4 describes our language composition technique by means of the case study, and Section 5 discusses possible means to generalize and automate the fragmentization and composition process. Finally, Section 6 presents related work and Section 7 concludes with future work.

2 Hybrid Language Case Study

We introduce here the case study, the Hybrid TFSA-CBD language, that is used throughout this paper to demonstrate the composition of language fragments.

2.1 The Hybrid TFSA-CBD Language

The hybrid case study is essentially the hybrid language case that can effectively model hybrid systems, i.e., systems that exhibit both discrete and continuous behaviour in the form of piece-wise continuous interleaved with discrete events. We have taken a basic language from each of the domains, TFSA (discrete) and CBD (continuous), to study the interleaving of the behaviour of the two languages and the two time domains.

Timed Finite State Automata (TFSA) are used for describing behaviour of reactive systems. A TFSA consists of the following: a set of states including a start stat; a set of transitions between the defined states, that include triggers/events (and/or guards) that can be either an event name or an `after` indicating some delay in time; and finally, a set of input events tagged in time.

Causal Block Diagrams (CBDs) is a visual modelling language commonly used for embedded control design that models systems with differential equations. CBD is the basic language of Mathworks Simulink®. It consists of blocks and connections between blocks. Each block has (optional) inputs and one output[1], and it can either represent an algebraic mathematical operation (such as summation, multiplication) or a time sensitive operation (e.g., delay the input). In our work, we focus on continuous-time models represented as CBDs. The simulation of such CBDs are however usually carried out in digital computers using a discrete-time approximation (i.e., by discretizing the differential equations and translating them into difference equations, which are represented as discrete-time CBDs) [6].

A TFSA-CBD composition is the weaving of a TFSA and CBD together in possibly different ways: 1) TFSA composed of a CBD; 2) CBD composed of a TFSA; or 3) a hierarchical composition with a TFSA composed of a CBD which in turn is composed of a TFSA (TFSA-CBD-TFSA or even CBD-TFSA-CBD). In this paper, our case study specifically focuses on the first kind of composition, that is a CBD embedded within a TFSA state where TFSA is the parent language and CBD is the child language. This is similar in concept to that of having Stateflow® models within Simulink® models.

When simulating dynamical systems modelled as a CBD, the output of the difference equations (described in the CBD) becomes not only a function of the input, but also a function of the whole history of inputs and the initial conditions, due to the use of time sensitive blocks (e.g., delay). The history of inputs is stored in the form of state variables and so the output is calculated as a function of the inputs and the state.

[1] Without loss of generality, we assume here that blocks in a CBD only have one output.

This not only implies that the detection of enabled transitions must be done at the end of each simulation step of the CBDs, but also means that a mechanism must be in place to ensure that after a transition occurs to a state with a CBD contained, proper initial conditions are provided to that CBD [13]. The transitions are usually triggered via events or guards that are usually implemented using *if* statements. In our case however, we want to know *when* the transition occurs as opposed to *if* the transition has occurred. For this purpose, *monitoring functions* need to be defined that allow for zero-crossing detection. The detection of the exact zero-crossing time is non-trivial and requires the use of state-event location techniques involving rollbacks to find the exact point. The syntax and semantics of the hybrid language needs to address this requirement. Additionally, modellers should also be given support to create hybrid models with the appropriate syntax to model such situations.

2.2 A Hybrid TFSA-CBD Example

The bouncing ball is a classic example of a hybrid system displaying both continuous and discrete behaviour. A ball is in free-fall motion when dropped. The dynamics change when the ball collides with the ground and bounces up again with reduced energy. An external event, such as a kick, is represented here, for illustrative purposes, as a constant increment in the velocity of the ball.

Fig. 1 shows the hybrid model of a bouncing ball modelled using our composed TFSA-CBD language. It includes the *freefall*, *collision*, and *kicked* states with state changes triggered by discrete events. Within each state, the continuous dynamics of the ball (essentially modelled with its height, x, and the velocity, v) evolves with time. During the free fall, gravity acts on the ball until it collides with the floor (x_0 crosses 0). At that instant, a collision occurs (detected by the `when +-` event), and the new v_0 is given by $v_0 = -0.8v$. In the same instance, the ball goes back to free falling again.

3 Hybrid TFSA-CBD Language Fragments

3.1 Language Specification Fragments

Language fragments can be defined as reusable components in the language engineering process. The fragments allow both reuse and modular language design. We define a language specification fragment (LSF) as a description of the specification of a modelling language in terms of its abstract syntax (AS) , concrete syntax (CS), operational semantics (OS), and the user-interface (UI) behaviour. As presented in Fig. 2, such a specification involves the definition of several models (referred to as m1-m9).

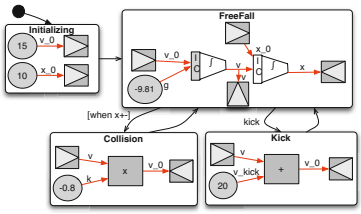


Fig. 1 Bouncing Ball Hybrid TFSA-CBD Model

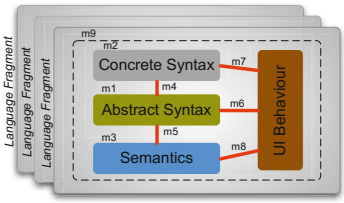


Fig. 2 Language Specification Fragments (LSF)

- m1: The AS model (at times referred to as the linguistic type model) describes the essence of the language (its structure and elements) by means of a meta-model usually modelled as a class diagram or an entity-relationship model.
- m2: The model instances of the language needs to be defined using some CS: a visual syntax, a textual syntax, or a combination of both. A single AS may be associated with one or more CS models.
- m3: The semantic model describes the language OS. In the case of a modelling and simulation environment, this refers to the simulator semantics defined with an algorithm outlining the computation steps.
- m4: The mapping model defining the mapping and rendering links between the AS and one CS.
- m5: The semantic mapping model which specifies how the AS elements are treated in the OS.
- m6-m8: These models specify the mappings between the UI behaviour of the interactive modelling environment and the AS (m6), CS (m7), and OS (m8).
- m9: A *glue model* that binds together models m1-m8.

This set of models together define one LSF, which in turn describes a part of the overall semantics of a hybrid language. Our focus here is on the AS and the semantics (models m1, m3, and m5), and we intend to look at the interfaces and a composition mechanism to compose the AS as well as the OS.

3.2 TFSA and CBD Fragments

We present here the TFSA and CBD LSFs. Fig. 3 presents the AS model of the TFSA language as a class diagram. Algorithm 1 outlines the behaviour for the FSA simulator. Besides the traditional notions of states and transitions, where on each transition we have the event triggers, it also includes the `after` construct that is implemented with the implicit notion of *elapsed time*.

Fig. 4 presents the AS model of the CBD language as a class diagram. Algorithm 2 outlines the behaviour of the CBD simulator in the *main loop* of the algorithm. For now, we are assuming fixed time-step simulation. This simplifies the zero-crossing detection without compromising the strength of our contribution because we are interested in coming up with the possible weavings assuming a fixed set of capabilities in the simulators.

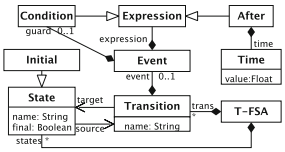


Fig. 3 TFSA Abstract Syntax Model

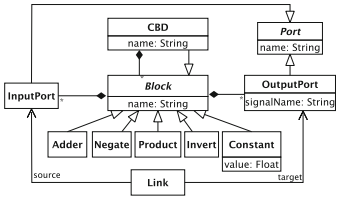


Fig. 4 CBD Abstract Syntax Model

Algorithm 1. TFSA Operational Semantics

$logicalTime, elapsedTime \leftarrow 0$; $currentState \leftarrow initialState$
while not endCondition **do**
  $E \leftarrow getInputEventAt(logicalTime)$
  if out-transition T from currentState has E **then**
    $currentState \leftarrow currentState.T.destination$;
    $removeInputEventFromInputList()$; $elapsedTime \leftarrow 0$
  end if
  if out-transition T from currentState has after(time) & time $\leq$ elapsedTime **then**
    $currentState \leftarrow currentState.T.destination$; $elapsedTime \leftarrow 0$
  end if
  $logicalTime \leftarrow logicalTime + \Delta t$; $elapsedTime \leftarrow elapsedTime + \Delta t$ ▷ Δt is a parameter to this algorithm
end while

Algorithm 2. CBD Operational Semantics (Adapted from [17])

$logicalTime \leftarrow 0$
while not end_condition **do**
  $schedule \leftarrow LOOPDETECT(DEPGRAPH(cbd))$
  for gblock **in** schedule **do**
    $COMPUTE(gblock)$
  end for
  $logicalTime \leftarrow logicalTime + \Delta t$ ▷ Δt is a parameter to this algorithm
end while

4 Hybrid TFSA-CBD Language Composition

In this section, we describe the composition of the hybrid TFSA-CBD language with regards to the abstract syntax and the semantics. In our case study, the hybrid system requirement is to model the continuous dynamics of a system when the system is in certain states. This entails TFSA states to be *embedded* with CBDs. Hence, a single state in a TFSA can be a simple state or a CBD.

Weaving the language descriptions defined as fragments as per our requirements gives us a new language definition (which can then again be referred to as a re-usable fragment) with a new AS model (class diagram) and a new automaton with new action code and a bigger store.

4.1 Composition of the Abstract Syntax Models

The AS models (or meta-models) modelled as class diagrams (shown in Fig. 3 and Fig. 4) are composed, resulting in the AS model of the hybrid language. The composition process involves the use of rule-based graph transformations that matches the pre-defined parent and child classes and joins them in a containment relationship within the class diagram. Due to space reasons, Fig. 5 only presents part of the composed class diagram - the complete details of the `Block` class and `Expression` class are shown in the original class diagrams. The `Transition` in the hybrid AS model is adapted to include a special kind of guard, specified (in the concrete syntax) as `[when +-]` to define an event for the zero-crossing detection in the source state dynamics. The *when* condition (both `[when +-]` and `[when -+]`) is added as a specialization of `Expression`.

Composing the AS models along with the CS models in our meta-modelling and model transformation tool, AToMPM (A Tool for Multi-Paradigm Modelling) [16], allows us to generate a working visual editor for the hybrid language. The AS data structures are also composed at the textual model level using SCCDs (described in the next subsection).

4.2 Composition of the Operational Semantics Models

Semantic adaptation involves the composition of time bases, and the interleaving of the control flow and data flow [7]. Fig. 6 shows how the semantics of the two languages, TFSA and CBD, are joined and adapted using operations to allow the two simulators to work in conjunction. The semantics behind the transitions labelled 1 to 6 are discussed below.

– **Time Adaptation**: This entails computing the new hybrid time step based on the outcome of a maximal common divisor (MCD) of the original time steps (see label 1). The CBD clock is initialized to zero at every simulation step when processing control is handed over to the CBD from the FSA (see label 5). The TFSA-CBD clock is incremented when all enabled transitions are processed and the FSA is ready to advance to the next state (see label 4).

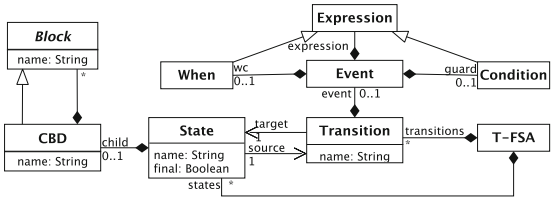


Fig. 5 TFSA-CBD Composed Abstract Syntax Model (Partial Model)

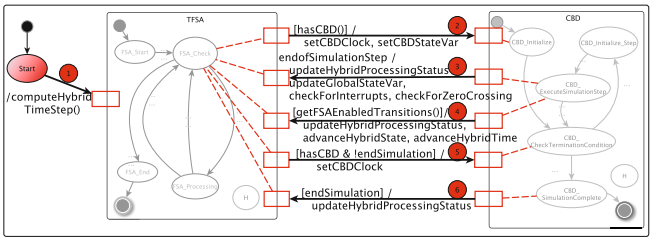


Fig. 6 Hybrid TFSA-CBD Semantic Adaptation

- **Control Adaptation**: The parent language, TFSA, has initial control. When a CBD is detected inside a FSA state, control is passed on to the CBD (see label 2). Following the execution of a CBD step, the simulator checks for zero-crossings and for any external events in the TFSA (see label 3). If a zero-crossing is detected, the CBD is reinitialized and execution continues (see label 2). An external event being triggered might involve a change in state and possibly the execution of a new embedded CBD. It should be noted that the switching of control is rather conceptual in our case, since our hybrid simulator is one combined simulator and does not require the passing of control from one simulator to another.
- **Data Adaptation**: The global state of the model is updated at every iteration at the child level and at the parent level (see labels 2 and 3).

The OS in algorithm form (presented in Section 3) is mapped to an automaton (SCCD model) with details in action code for each fragment involved. SCCD is a formalism that combines Class diagrams and Statecharts to define classes, their behaviour and interactions. The weaving of the simulation algorithm automatons (SCCDs) is based on a syntactic and semantic *glue* model, similar to the model shown in Fig. 6. The use of the SCCD language allows us to generate SCCD XML or SCCD HUTN (human-usable textual notation) from a complete behavioural model of a simulator instead of hard coding the simulator and carrying out the adaptation within wrappers at the code level.

The SCCD models describing the OS of the TFSA and the CBD languages are shown as part of Fig. 7, which presents the composed SCCD model for hybrid TFSA-CBD. In the figure, the TFSA and CBD composite states represent the SCCD models of the TFSA and CBD simulators respectively. For space reasons, the internal details of the simulator classes are not presented. These classes include the definition of the data structures, runtime variables, and methods used in the Statechart part of the SCCD model. The visual SCCD is first transformed to a SCCD HUTN model which is then compiled to Python source. This results in a fully automatically generated simulator from our description of the OS.

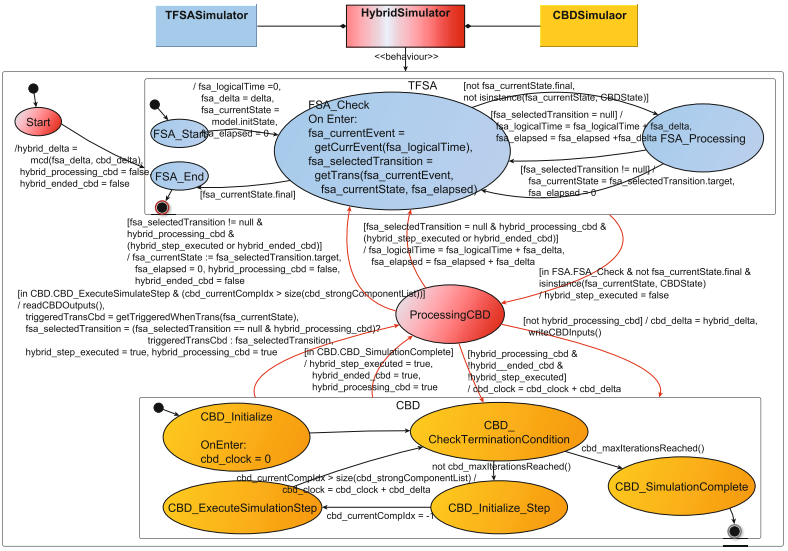


Fig. 7 Hybrid TFSA-CBD Operational Semantics

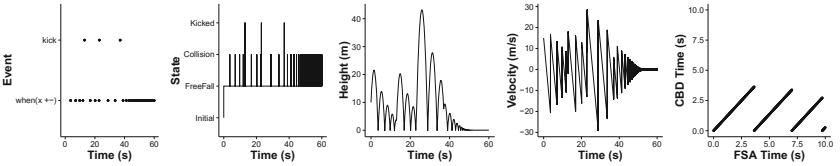


Fig. 8 Bouncing Ball Hybrid Model Simulation Results

4.3 Simulation using the Composed Simulator

The bouncing ball model (shown in Fig. 1) was simulated using our woven simulator. Figure 8 presents five graphs all plotted against the *time* of the parent formalism, TFSA in this case. The simulation results shows the zeno behaviour of the bouncing ball. It also shows how the nested child (CBD) clock advances with the hybrid (parent) clock.

5 Modular Language Design of Hybrid Languages

Based on the insights gained from the hybrid case study, we are looking more deeply into the fragmentization and composition process. Ongoing work is discussed here.

5.1 Fragmentization

We propose using a network language (with input/output ports) as a way to package each LSF to enable interleavings with other LSFs via pre-defined ports (similar to what is shown in Fig. 6). The network language provides an interface for each fragment SCCD which can be used to define explicit communication between the LSFs involved and to carry out semantic adaptation. This model can then be automatically transformed to the woven SCCD for the new language. These LSFs ultimately can be used to build a library of reusable fragments.

In this paper, we have defined LSFs to be complete and meaningful languages. This constraint can be further relaxed to allow LSFs to be underspecified or to include *holes*. These model elements or *holes* can then be referenced and linked with special ports in the network language, which can then be replaced or extended with another complex model (defined as a LSF) in any part of the AS or OS model. For instance, a statement can be replaced by a block of statements or an association replaced by an entire meta-model. The nested LSF can itself have *holes* leading to a composed language that can further be replaced. These *holes* must be identified and replaced in order to build valid and meaningful languages. At the moment, we are looking into techniques to allow automatic checking for inconsistencies and incompleteness in LSFs.

5.2 Composition

Based on the studies carried out, we propose five patterns for language composition: *embed*, *weave*, *build*, *replace*, *extend*, and *slice*. Fig. 9 gives an example of each pattern. In the hybrid TFSA-CBD case study, we implicitly used the *embed* pattern to compose the two languages. Embedding leads to a parent-child structure which needs to be addressed in a specific manner in the weaving. The addition of the `when` construct was implicitly done using the *extend* pattern. Having a model transformation defined for each pattern would allow modellers to set the initial parameters for the composition and select the required pattern to build a composed syntax or semantics model. A higher-order transformation could identify the *holes* in the language(s) and define the order in which each pattern should be applied, leading to an automated weaving process. AToMPM will be used to build and compose the AS models using graph transformations.

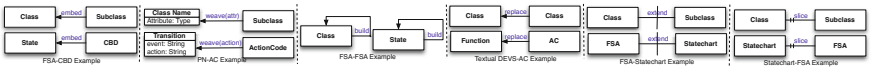


Fig. 9 Language Composition Patterns

5.3 Automating the Semantic Adaption Process

As part of the fragmentization and composition process, the definition and composition of the SCCD models need to be developed further. We begin by generalizing the OS specification by defining a fixed structure that every simulator design has to adhere to. We propose adapting the simulator design to the form of a Statechart (as part of the SCCD) where every simulator has the following (basic or composite) states: *Start*, *Prepare*, *Process*, *CheckTermination*, *EndSimulation*, and a history state. The execution follows in sequence with a loop that returns to *Prepare* following *CheckTermination* to continue with the remaining iterations. To make the simulator a reusable LSF, we next instrument the Statechart with a *ProcessChild* (which now acts as a *hole*) state with transitions to and from *Process*. The *ProcessChild* is then connected to a network port which allows it to be referenced by other LSFs. Once we have a generic structure for all simulators, we can easily use graph transformations to compose the simulators. The semantic adaptation can be carried out by defining interleavings and adding special constructs with the aid of the patterns discussed above.

We envision having a composition process that is closed-under-composition: the woven simulator is a LSF that can be composed again with other languages, to provide arbitrarily complex modelling languages for the development of ever more complex CPS.

6 Related Work

In modular language engineering, the disparate models of computation employed in the semantics of domain-specific modelling languages bring about many challenges in what matters to their reuse. Reuse is the main focus of CORE [15], which allows specification of reusable aspects as *concerns* and provides support for automated weaving of requirements models. In [1], structural models expressed as instances of UML class diagrams, can be seen as graphs and therefore reused across languages. The reuse mechanism is expressed and ruled by the notion of fragments, which are reusable and composable as new fragments.

Some existing approaches focus on reusing models of modelling languages to build new modelling languages, and on trying to identify the most convenient abstractions to convey the reuse of such kinds of models of languages. In [4], several composition operations such as merge, aggregation, and deletion are introduced to allow new languages to be built from more simpler ones. The concepts for language reuse were further developed and studied in language workbenches, such as Spoofax [8] and MPS [18]. Moreover, the problem of language composition from other language components was extensively explored in [9].

In [19], the structure of languages (i.e., abstract syntax and semantics) is taken into account as different roles, in the composition of languages here considered as language components that can be reused. Similar to [19] and [1], the work in [20] introduces a language for the definition of language fragments, which treats languages as components that can be reused and composed.

The reuse of the language semantics mostly remains unaddressed in the current state of the art, and is seen as one of the main challenges in modular language engineering. If we approach this problem by considering only the semantics that are given by translations, then it might be solved using parameterization [14]. This solution though only glosses over the problem of reusing the semantics of languages, and the meaning and behaviour of the translated models in the target language with regards to composability and compatibility are not addressed. However, the work in [5] does tackle semantics composition by means of aspect-oriented concepts, but can be considered to be a code-level solution. A framework for composing semantics for the purpose of simulating multi-formalism models is proposed in ModHelX [3], which involves an explicit representation of the model of computation used in each of the languages. In [11][7], hybrid formalisms are created as the result of composing OS by assuming a common interface abided by black-boxed simulators. The interface ensures that they can be composed while preserving important semantic properties such as a common time-base thereby ensuring that the composition of the language semantics is meaningful.

7 Conclusion

This paper presents a language composition technique for hybrid systems, demonstrated using the hybrid TFSA-CBD case study. We introduce language specification fragments (LSF) to model the specification of languages as a set of syntax and semantics models. The composition technique focuses on fragmentization and composition of the LSFs (specifically, the AS and the OS). The semantics (i.e., the behaviour of the simulators) are explicitly modelled as SCCDs (Statechart+Class Diagram), which are then woven together by applying semantic adaptations which address the problems that arises in hybrid languages (such as, the need for zero-crossing detection). From the composed semantics model, we are able to generate a fully functional hybrid simulator for the TFSA-CBD language. We plan on working on several case studies involving compositions of discrete-discrete and continuous-continuous formalisms, in addition to other discrete-continuous formalisms to further validate our claims.

As future work, we also intend to adapt the composition technique to include the interleaving of the concrete syntax (along with the UI behaviour) in the LSFs.

Acknowledgments This work was partly funded by the Automotive Partnership Canada (APC) in the NECSIS project as well as by Flanders Make.

References

1. Amalio, N., de Lara, J., Guerra, E.: FRAGMENTA: a theory of fragmentation for MDE. In: MODELS, pp. 106–115. IEEE (2015)
2. Barroca, B., Mustafiz, S., Mierlo, S.V., Vangheluwe, H.: Integrating a neutral action language in a DEVS modelling environment. In: SIMUTOOLS, pp. 19–28. ACM (2015)
3. Boulanger, F., Hardebolle, C., Jacquet, C., Marcadet, D.: Semantic adaptation for models of computations. In: ACSD, pp. 153–162. IEEE (2011)
4. de Lara, J., Guerra, E., J. Sánchez-Cuadrado, J.: Abstracting modelling languages: a reutilization approach. In: CAiSE, pp. 127–143. Springer (2012)
5. Degueule, T., Combemale, B., Blouin, A., Barais, O., Jézéquel, J.-M.: Melange: A meta-language for modular and reusable development of DSLs. In: SLE 2015, pp. 25–36. ACM (2015)
6. Denckla, B., Mosterman, P.: Formalizing causal block diagrams for modeling a class of hybrid dynamic systems. In: CDC-ECC 2005, pp. 4193–4198 (2005)
7. Denil, J., Meyers, B., De Meuleneare, P., Vangheluwe, H.: Explicit semantic adaptation of hybrid formalisms for FMI co-simulation. In: TMS/DEVS 2015, pp. 852–859. SCS International (2015)
8. Kats, L.C., Visser, E.: The spoofax language workbench: Rules for declarative specification of languages and IDEs. SIGPLAN Not. **45**(10), 444–463 (2010)
9. Krahn, H., Rumpe, B., Völkel, S.: Monticore: a framework for compositional development of domain specific languages. Journal on STTT **12**(5), 353–372 (2010)
10. Lacoste-Julien, S., Vangheluwe, H., Lara, J.D., Mosterman, P.J.: Meta-modelling hybrid formalisms. In: CACSD, pp. 65–70. IEEE (2004)
11. Meyers, B., Denil, J., Boulanger, F., Hardebolle, C., Jacquet, C., Vangheluwe, H.: A DSL for explicit semantic adaptation. In: 7th Workshop on Multi-Paradigm Modelling, MoDELS 2013, pp. 47–56 (2013)
12. Mierlo, S.V., Barroca, B., Vangheluwe, H., Syriani, E., Kühne, T.: Multi-level modelling in the modelverse. In: Workshop on Multi-Level Modelling, MoDELS, pp. 83–92 (2014)
13. Mosterman, P.J.: An overview of hybrid simulation phenomena and their support by simulation packages. In: Hybrid Systems: Computation and Control. LNCS, vol. 1569, pp. 165–177. Springer (1999)
14. Pedro, L., Amaral, V., Buchs, D.: Foundations for a domain specific modeling language prototyping environment: a compositional approach. In: 8th OOPSLA Workshop on Domain-Specific Modeling (DSM), October 2008
15. Schöttle, M., Alam, O., Ayed, A., Kienzle, J.: Concern-oriented software design with TouchRAM. In: MODELS 2013, pp. 51–55 (2013)
16. Syriani, E., Vangheluwe, H., Mannadiar, R., Hansen, C., Mierlo, S.V., Ergin, H.: AToMPM: a web-based modeling environment. In: MODELS 2013 Demonstrations (2013)
17. Vangheluwe, H., Denil, J., Mustafiz, S., Riegelhaupt, D., Van Mierlo, S.: Explicit modelling of a CBD experimentation environment. In: TMS/DEVS 2014, pp. 13:1–13:8. SCS International (2014)
18. Völter, M., Visser, E.: Language extension and composition with language workbenches. In: OOPSLA 2010 Companion, pp. 301–304. ACM (2010)
19. Wende, C., Thieme, N., Zschaler, S.: A role-based approach towards modular language engineering. In: SLE 2010, pp. 254–273. Springer (2010)
20. Živković, S., Karagiannis, D.: Towards metamodelling-in-the-large: interface-based composition for modular metamodel development. In: Enterprise, Business-Process and Information Systems Modeling, pp. 413–428. Springer (2015)

Part VIII
Data Mining

Data Clustering Using Improved Fire Fly Algorithm

Mehdi Sadeghzadeh

Abstract Clustering is considered as one of the most important techniques for data mining that is used for data analysis in some areas such as text identification, image processing, economic science, and spatial data analysis. Several algorithms have been proposed for solving the problem clustering. These algorithms are using different techniques. Firefly algorithm was inspired by the process of producing twinkle lights of this insect, and is considered as one of the designed base on a collective behavior of insects. In this paper, the evolutionary algorithm of Firefly is used for solving clustering problem. The proposed algorithm is compared with firefly algorithm, differential evolution algorithm and k-means algorithm on some important data sets from database UCI. According to the results, this algorithm is more appropriate for better clustering rather than firefly algorithm.

Keywords Clustering · Firefly algorithm · Data mining · Meta-heuristic algorithm · Data clustering

1 Introduction

Data clustering is an NP-hard problem, which is regarded as a challenge due to an increase in the size of data sets, solution in a logical and acceptable time and its optical solution. This problem has been the subject of many research articles. K-means algorithm is one of the significant algorithms in solving data clustering problem. Its implementation is simple and has high speed for finding the local optimum. This algorithm has been successful in some clustering problems. However, high dependency of the final results on the initial state in this algorithm and convergence possibility of the problem to local optimum instead of global

M. Sadeghzadeh(✉)
Department of Computer Engineering, Mahshahr Branch,
Islamic Azad University, Mahshahr, Iran
e-mail: Sadegh_1999@yahoo.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_69

optimum caused this algorithm not to be able to solve many problems [1]. Therefore, for solving data clustering problem, researchers tend to alternative optimization methods such as genetic algorithm, Firefly algorithm, imperialist competitive algorithm and differential evolution algorithm [2-5]. In this paper, the improved firefly algorithm was used for solving data clustering problem.

2 The Clustering Problem

Clustering is used for organizing models in homogenous clusters for discovering intra-relations and inter-collections of samples and templates. Clustering can be utilized in some areas such as a decrease in volume and data dimensions for data pre-processing and analysis of similarity or dissimilarity. Many factors should be considered for solving clustering problems such as initial conditions or similarity standards. One of the common standards for measuring similarity among the existing samples is the euclidean distance, defined in (1).

$$d(X,c_j)=\sqrt{\sum_{i=1}^{m}(x_i-c_j)^2} \tag{1}$$

Where m is the number of data, X is a data vector, and c_j is a center of cluster j.

In fact, we consider each n sample as a point in an N-dimensional space, and then we assign these points to k cluster based on pre-defined standards. One of the important algorithms in solving data clustering problem is the k-means algorithm, used to obtain the local optimal solution due to a relatively high speed and simplicity. The main objective of this algorithm is to find the k center of the clusters. However, despite its success in solving most of the clustering problems, this algorithm cannot solve many problems due to the dependency of the final output on an initial value of centers and new convergence and being in a local optimum.

3 Firefly Algorithm

Fireflies produce twinkle light rhythmically in nature, which is used for absorbing mate and hunting. Firefly algorithm is an evolutionary model resulting from nature that is based on collective intelligent algorithms. For the first time, this algorithm proposed by Mr. Yang in Cambridge University in 2008 [6].

There are two main problems in firefly algorithm including a difference in light intensity, and formulating the amount of absorbency. For simplicity, we can consider that Firefly absorbency with its luminosity depends on target function. Since Firefly absorbency is in proportion to light intensity of adjacent Firefly, β absorbency from Firefly is defined as follows:

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{2}$$

β_0 is absorbency in $r = 0$.

The distance between both fireflies i, j in points x_i and x_j is shown as follows regarding the Cartesian distance:

$$r_{ij} = \left\| x_i - x_j \right\| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \tag{3}$$

In which $x_{i,k}$ is k the portion of spatial coordination x_i of firefly i in two-dimensional state, and r_{ij} is shown as follows:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4}$$

Also, Firefly movement and its absorption to Firefly j that is move luminous is determined as follows.

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_i - x_j) + \alpha(rand - \frac{1}{2}) \tag{5}$$

In which the second statement shows absorbency, and the third statement is chance maker that is done by chance maker parameter α. Rand is the random number that is obtained by uniform distribution within [0,1]. Parameter β_0 shows the amount of absorption in the light source. Parameter γ is very effective in determining convergence speed and behavior of firefly algorithm. This parameter is determined based on absorbency changes which are within [0, ∞) theoretically.

4 Differential Evolution Algorithm

Differential evolution algorithm was proposed by Storn and Price in 1997 [7]. This algorithm is a random algorithm based on population and is considered as one of the evolutionary algorithms. The method for producing a new solution in differential evolution algorithm is a unique method. In order to solve the optimization problem, sampling target function is randomly used in multiple particular points. The predetermined parameters constraints specify some regions in which initial population are produced. This algorithm produces a new solution in a d-dimensional space and therefore this new solution results from the difference of available points. This algorithm needs determining three parameters, including parameter NP as population size, parameter CR as the probability of crossover, and parameter F as mutation weigh which is multiplied by the difference of two vectors and added to the third vector.

For mutation, three vectors are randomly selected (two by two). The mutated vector is generated by (6) for each vector within population, which is the first

variant of DE [7]. This design increases the greedy degree of algorithm movement toward optimization by using optimum vector. Convergence is appropriate for increasing the speed, especially for problems which their global optimum is found simply.

$$V_{i,G+1} = X_{r1,G} + F * (X_{r2,G} - X_{r3,G}) \tag{6}$$

In the crossover stage, each of mutated vector components is transferred to candidate vector by probability CR, means crossover constant belongs to [0,1]. Otherwise, the corresponding component is substituted for the main vector.

$$u_{ji,G+1} = \begin{cases} v_{ji,G+1} & if \quad rand(j) \le CR \quad or \quad j = randb() \\ x_{ji,G} & Else \end{cases} \tag{7}$$

In Equation (7), rand (j) is j th call of random function that is produced a number within [0, 1]. However, for ensuring that at least one component is transferred to primary vector, one component is transferred randomly from mutated vector to primary vector without regard to CR. Therefore, one component is randomly selected for each candidate vector by using function randb() for transferring to the next generation.

The greedy method is used for selecting survivors. Each vector is compared with similar candidate vector and those which are more competent are transferred to the next generation. Figure1 shows the DE algorithm.

$$x_{i,G+1} = FitnessVector(u_{i,G+!}, x_{i,G}) \tag{8}$$

Fig. 1 shows differential evolution algorithm [7].

Differential Evolution Algorithm
Generate $P=(x_1,x_2,\ldots,x_n)$*; N(point in D)*
Repeat
  For *i=1* ***to*** *N* ***do***
    Compute a mutant ***u****;*
    Create ***y*** *by the crossover of* ***u*** *and* $\mathbf{x_i}$*;*
    If $f(\mathbf{y}) < f(\mathbf{x_i})$ ***then*** *insert* ***y*** *into Q*
      Else *insert* $\mathbf{x_i}$ *into Q*
  End *if;*
End *for;*
P:= Q;
Until *stopping condition;*

Fig. 1 The outline of differential evolution algorithm [7]

Differential evolution algorithm experiences a differential operator for producing new solutions and this operator exchange information among population members. One of the advantages of this algorithm is having a memory that keeps information of suitable solution in the new population. Another benefit

of this algorithm is related to the operator of its selection. In this algorithm, all members of one population can have an equal chance of being selected as one of the parents. In this algorithm, instead of each member, a gene donor is produced which is responsible for exchanging by that member. In the proposed method, at first, minimum distances among data from centers, are computed and maximum distances among the centers of clusters are obtained. Then we consider the total distance with some coefficients for each of these clusters as target function. By this change, clustering producer is done more accurately. Further, clustering centers are determined with higher precision.

5 Proposed Algorithm

The purpose of this paper is to find the centers of clusters as optimum. Various standards can be introduced for optimal clustering. One of them is intra-cluster distance that is finding the least distance of each data to the cluster. In this paper, algorithm HEFA was used as a combination of methods of Firefly and differential evolution. In this algorithm, initial population were sorted based on fitness and then divided into two subpopulations. The best values were placed in the first subpopulation, and the worse values were put in the second subpopulation. In the first subpopulation, absorption of each firefly was determined based on formula (2) and Firefly algorithm. Moreover the movement of each firefly toward the best (the most absorbent) Firefly was done by an equation (5). Then in the second subpopulation, the best values were produced by the weakest ones by using differential evolution algorithm. Member's generation was done by using mutation operator that has been shown in Eq. (9).

$$V_{i(t)} = \{v_{it(1)}, v_{it(2)}, \ldots, v_{it(d)}\} \tag{9}$$

$$v_{i(t)} = x_{best(t)} + F.(x_{r_1(t)} - x_{r_2(t)}) \tag{10}$$

Where, $x_{best(t)}$ is the vector of the current best solution, F is the mutation vector, x_{r_1} and x_{r_2} are randomly chosen vectors from the neighboring solutions. Afterwards, a new member is generated by intersecting between generated member from the previous stage and the new member from X_i. You can observe the way of member generation in (11).

$$Y_{i(t)} = \{y_{it(1)}, y_{it(2)}, \ldots, y_{it(d)}\} \tag{11}$$

$$y_{i(t)} = \begin{cases} V_{i(t)} & if \quad R<CR \\ X_{i(t)} & otherwise \end{cases} \tag{12}$$

Where R is a uniformly distributed random value between 0 and 1 and CR is the predefined crossover constant. In this stage, if the newly generated member is more competent than the previous member, then it is replaced.

$$x_{i(t+1)} = \begin{cases} Y_i(t) & if \quad f(Y_{i(t)}) \le f(X_{i(t)}) \\ X_i(t) & if \quad f(Y_{i(t)}) > f(X_{i(t)}) \end{cases} \tag{13}$$

As a result, the best values are selected from each subpopulation, and then the best value is obtained among data by their integration. In fact, this algorithm considers both local and general optimum. It causes the space to be searched better. Fig. 1 shows the outline of the proposed HEFA method.

Hybrid Evolutionary Firefly Algorithm (HEFA)

Input: *Randomly initialized position of d dimension problem:* X_i
Output: *position of the approximate global optima:* X_G
Begin
Initialize population; Evaluate fitness value;
X_G ⟵ *Select current best solution;*
For *t* ⟵ *1 to max*
Sort population based on the fitness value;
X_{good} ⟵ *first _half(X);* X_{worst} ⟵ *second_ half(X);*
For *i* ⟵ *0 to number of* X_{good} *solutions*
For *j* ⟵ *0 to number of* X_{good} *solutions*
If *(f(Xi)>f(Xj))* ***then***
Calculate distance and attractiveness;
Update position;
End if
End for
End for
For *i* ⟵ *0 to number of* X_{worst} *solutions*
Create trivial solution, $V_{i(t)}$ *;*
Perform crossover, $Y_{i(t)}$ *;*
Perform selection, $X_{i(t)}$ *;*
End for
X ⟵ *combine(*X_{good} *,*X_{worst}*);*
X_G ⟵ *Select current best solution;*
t ⟵ $t+1$
End for
End Begin

Fig. 2 The outline of proposed HEFA method

In this paper, we used three distance metrics as fitness function in proposed HEFA method. One of the metrics used to evaluate the clustering problem is the distance to within cluster, used as follows (14):

$$\text{Function1}=\frac{1}{M}\sum_{i=1}^{M}\left\|Data_i - Cl_i\right\| \tag{14}$$

Where, Cl_i is a cluster center in $Data_i, M$ is the number of data and the maximum distances between centers of clusters are obtained.

The second distance metric is the average distance between the clusters centers are also considered to be as (15).

$$\text{Function2}=\frac{1}{c*(c-1)}\sum_{i=1}^{c}\sum_{i=1}^{c}\left\|Cl_i - Cl_j\right\| \tag{15}$$

The third metrics is obtained from both the first and second as follows:

Function3=Function1/Function2

The objective is minimizing Function1 and Function3 and maximizing Function2. These three metrics are not accurate. The best metrics is error rate of clustering as (16):

$$Err=\left(\frac{1}{N}\sum_{i=1}^{N}(if\,(class(i)=cluster(i))then \quad 0 \quad else \quad 1\right)*100 \tag{16}$$

6 Simulation Results

In this section, the results from proposed algorithm are compared with algorithms k-means, Firefly, and differential evolution for solving clustering problems. For DE algorithm, Firefly algorithm and proposed algorithm we consider $N_{pop}=50$ and a number of iteration equal 100 times. The other parameters initialization are $\gamma=0.01,\alpha=0.01$, $\beta=2,F=0.01,CR=0.9$, mutation factor is 0.01 and number of clusters are set to 3.

Implementation of the algorithm was done by MATLAB software package. The data sets used included Pima, Glass, Wine, Breast cancer, Iris which were extracted from the popular database UCI. Table 1 shows the mean of inter-class distance related to data set: Breast cancer, Iris, Glass, and Sonar. Table 2 shows the average of extra-cluster distance. Table 3 shows the division of inter-class to Extra class distance. Finally Table 4 shows clustering error rate. As it is clear from these tables, evolutionary Firefly algorithm did clustering more accurately than the other mentioned algorithms. According to the results, this algorithm can be used in clustering problem for finding optimal or semi-optimal solution in order to divide N objects in K cluster.

Table 1 average of inter-cluster distance

Dataset/ algorithm	k-means	DE	FA	HEFA
BCW	0.5438	0.4494	0.4529	**0.4488**
Glass	0.5517	0.3045	0.3180	**0.2900**
Iris	0.4650	0.1954	0.1945	**0.1943**
Sonar	**1.3513**	1.6177	1.6104	1.5992

Table 2 average of extra-cluster distance

Dataset/ algorithm	DE	FA	HEFA
BCW	**0.4122**	0.4092	0.4083
Glass	**0.4109**	0.4082	0.4082
Iris	**0.6150**	**0.6150**	**0.6150**
Sonar	**0.1894**	0.1634	0.1677

Table 3 final distance based on average of inter-cluster distance and extra-cluster distance

Dataset /algorithm	DE	FA	HEFA
BCW	**1.0902**	1.1068	1.0992
Glass	0.7411	0.7790	**0.7104**
Iris	0.3177	0.3163	**0.3159**
Sonar	**8.5412**	9.8555	9.5361

Table 4 clustering error rate

Dataset/ algorithm	k-means	FA	HEFA
BCW	19.12	12.42	**5.98**
Glass	47.66	45.54	**44.39**
Iris	11.33	10.45	**9.33**
Sonar	30.22	35.57	**28.06**

7 Conclusions

In the present study, a solution for solving data clustering problem was proposed by using evolutionary firefly algorithm. In this algorithm, the least distance between the available data to the centers was computed. Also, the farthest distance from centers of clusters was obtained. Then the ratio of inter-distance to extra distance was considered. The obtained results were compared with algorithms k-means, Firefly and differential evolution. The results indicated that proposed algorithm presents more acceptable result than algorithm k-means, Firefly, and differential evolution.

References

1. Liu, Y., Wu, X., Shen, Y.: Automatic clustering using genetic algorithms. Applied Mathematics and Computation **218**, 1267–1279 (2011). Elsevier
2. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surveys **31**(3), 264–323 (1999)
3. Senthilnath, J., Omkar, S.N.: Clustering using firefly algorithm: performance study. Swarm and Evolutionary Computation **1**(3), 164–171 (2011)
4. Yi, T.K., Zahara, E., I-wei, K.: A hybridized approach to data clustering. Expert System with Applications **34**(3), 1754–1762 (2008)
5. Hosseini, M., Sadeghzadeh, M., Nourmandipour, R.: An efficient approach based on differential evolution algorithm for data clustering. Decision Science Letters **3**, 319–324 (2014)
6. Yang, X.S.: Firefly algorithms for multimodal optimization. In: Watanabe, O., Zeugmann, T. (eds.) SAGA. LNCS, vol. 5792, pp. 169–178. Springer, Heidelberg (2009)
7. Storn, R., Price, K.: Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization **11**(4), 341–359 (1997)

Intelligent Mobile App for You-Tube Video Selection

Mark Smith

Abstract The usage of mobile devices has increased dramatically in recent years. These devices serve us in many practical ways and provide us with many services – many of them in real-time and "on demand". The delivery of streaming audio, streaming video and internet content to these devices has become common place. One of the most popular sources of video/audio content is You-Tube – where billions of videos are uploaded and accessed each day. An increasing challenge is to locate the desired video from among many dozens of possibilities. This paper introduces an intelligent mobile application that utilizes the You-Tube Application Programming Interface (API) in developing a novel algorithm for selecting the most appropriate video associated with a song title and artist. The test case used in this application is the domain of all the top 40 popular songs (as provided by Billboard Inc.) from 1970 to present – resulting in approximately 14,000 possible songs. The application described in this work invokes the You-Tube API based on 3 different criteria – most popular video, most relevant video and a key word search of the video's comments. These criteria are then merged in a voting algorithm thus providing the best possible video pertaining to a song title and artist. This system, implemented for iOS 9 using XCode and Swift, allows the user to choose from thousands of song titles and provide a "music on demand" system therefore playing the best possible video associated with the song title and artist. Results for the app consist of randomly selecting 50 songs from each year (1867 total) and verifying that the most appropriate video was selected.

1 Introduction

Mobile computing has become one of the fastest evolving areas of computer science. Consumer demand and interest in mobile devices has exploded, as exemplified by the introduction of smart phones such as Apple's iPhone [4], T-Mobile's

M. Smith(✉)
University of Central Arkansas, Conway, USA
e-mail: marks@uca.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_70

Android [2,12], and Microsoft Windows Phone[1]. Indeed, many programmers have found new opportunities developing applications (better known as Apps) for these smart phones as exhibited by almost 2 billion apps downloaded for the iPhone to date. Many of these apps involve accessing internet content consisting of streaming audio and video from various websites. One of the most popular websites for storing and accessing video content is You-Tube – considered by many as a cultural phenomenon. But much searching often must be performed when locating the most appropriate video for a given topic. Perhaps no genre of video requires more scrutiny when searching as music videos stored at You-Tube. The challenge for selecting the best music videos is due to many cover videos, many user-created videos and unrelated videos due to common titles. These videos currently must be filtered manually before finally accessing the desired video of choice. The system introduced by this paper, implemented in the form of an iOS App, proposes a way of automatically filtering the sometimes dozens of additional videos and locating the most appropriate video for the user to play on demand. The versatile You-Tube API is given keywords consisting of the song title and the artists along with the following options:

- Most viewed (i.e., popular) video
- Most relevant (based on a rating system) video
- Most frequently occurring keywords (i.e., song title and artist) search performed on video's comment section

These options result in approximately 50 possible videos returned for each of the three criterion. A comparison is then made between the videos returned from each option in the form of a voting algorithm. The voting algorithm determines which videos occur most frequently between the three options with priority given to the most viewed (i.e., popular) video. The primary motivation behind the project presented in this paper is to improve the video selection information currently available at You-Tube by combining and filtering these sources by one application. The voting algorithm is a key component of the algorithm and is utilized in identifying and filtering less valid videos thus providing and playing the most appropriate and best fitting video. The smartphone utilized in this work is Apple's iPhone, along with its Xcode development platform [5,6,7]. The system developed from this research and presented in this paper is separated into the following sections:

- Integration of the iPhone with You-Tube using the three API options
- Selection of the most relevant You-Tube options
- Extraction of individual video data using JSON and HTML
- Consolidation of data by voting algorithm
- Classification of one video as the most appropriate video to be displayed
- Comparison of consolidated results with individual results.

This paper will explore each of these items in and the subsequent algorithm implementation in detail.

2 Extraction of You-Tube Data

An important first step is to integrate this iOS App with the You-Tube website using the three different criterion described in Section 1. The criterion are implemented using the You-Tube API that allows keyword searches based on the desired option – Most Popular, Highest Rated, and Comments Evaluation [14,15,16]. The Application Programming Interface is essentially a set of library functions allowing data to be extracted from the You-Tube website programmatically as opposed to accessing the data via a basic website URL using HTML A system diagram showing a portion of this system for the 3 criterion is shown below in Fig. 1. The diagram outlines the different ways of utilizing the You-Tube API [3].

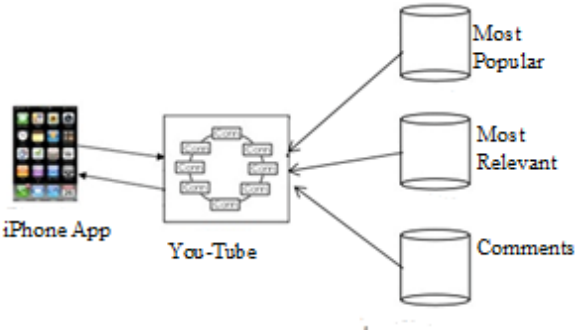


Fig. 1 You-Tube System Diagram

The You-Tube API is first invoked by using a URL specifying the keywords pertaining to the song title and artists and whatever options are desired. An example of invoking the You-Tube API for a song title of "Close to You" by the artists "The Carpenters" for the "Most Popular" video option is shown below:

```
https://gdata.youtube.com/feeds/api/videos?
    q="The+Carpenters+Close+To+You"
    &orderby=viewCount
    &start-index=1
    &max-results=50
    &v=2
```

The option for "Most Popular" is specified as "viewCount". Modifying this option to "rating" specifies the "Most Relevant" video category. Separate URLs must be constructed for different options that are mutually exclusive such as viewCount and rating.

The data is returned from You-Tube to the iPhone in the form of Java Script Object Notation (JSON). This notation is a standard notation most commonly used for transferring data between websites. In this case, the data exchange is between a mobile device and a website. An example of the format of a JSON data stream is shown below [11,12]:

```
"items": [
  {
   "kind": "youtube#searchResult",
   "etag": "\"fpJ9onbY0Rl_LqYLG6rOCJ9h9N8/-
4Ost0cp1yxnCNDYWZ39R7O67Tk\"",
   "id": {
    "kind": "youtube#video",
    "videoId": "pVlr4g5-r18"
   },
   "snippet": {
    "publishedAt": "2008-03-31T23:06:50.000Z",
    "channelId": "UCL5M-4S7q_X0ArlBxxdjK3Q",
    "title": "The Carpenters- Close To You(Official
Video)",
```


(…more fields and more video results follow)

The field names are delimited with a colon while the actual values for the fields follow the colon. Only a small portion of the entire JSON data exchanged between You-Tube and the iPhone is shown. The description and comments section comprise a large segment of all data exchanged.

Unfortunately, the search string format used with the You-Tube API returns many videos that are unrelated to the intended target as specified by the user [8,9,10]. In addition, the first video retrieved for a given option is rarely the most appropriate video to be displayed to the user. Clearly a filtering process is needed to narrow the focus to only those videos that best fit the video desired by the user. This filtering process is described next in section 3

3 Video Filtering Process

The next step in this process is to remove unrelated videos from the set returned from the You-Tube API. This filtering process consists of the following three criterion:

1. Filter based on song keywords found in the title
2. Filter based on keywords found in the comments
3. Filter based on a required number of comments

The first criteria simply assures that the keywords pertaining to the song title actually exist in the title of the video. Most of the videos retrieved from the You-Tube API pass this filtering step, although approximately 10% (about 15 videos) are removed from consideration due to this criteria [13].

The next step performs a search of all keywords – consisting of both the song title and the artist – throughout all of the comments posted and pertaining to the video. Our analysis has shown that this step assures that unrelated videos due to

special circumstances and coincidences are almost always eliminated by performing this step. The special circumstances generally occur when the artists or songs have brief names that are commonly used words in everyday speech. This step has shown to remove videos pertaining to these special circumstances from consideration.

The last step examines the total number of comments posted by viewers of the given video. The central idea is that only those videos that have received an adequate number of comments should be considered as a candidate for the best fitting and most appropriate video to be displayed to the user. The required number of comments is computed adaptively and is based on the group of videos returned from the You-Tube API. First the average number of comments is computed over the videos that pass the first two tests as given in Equation (1);

$$\mu_c = \frac{\sum_{i=0}^{N-1} C_i}{N} \tag{1}$$

Where C_i is the number of comments for the *ith* video, N is the number of videos that satisfy tests 1 and 2 specified above, and μ_c is the average number of comments for the set. Next, the standard deviation for the number of comments occurring in the videos is computed in Equation (2). This standard deviation will be used for adaptively determining the required number of comments that should be computed based on the overall results.

$$\sigma_c = \sqrt{\frac{\sum_{i=0}^{N-1} (C_i - \mu_c)^2}{N-1}} \tag{2}$$

Our analysis has shown that the required number of comments for a given set of vides should exceed the average number of comments plus twice the standard deviation. This adaptive threshold is computed as showing in Equation (3) below:

$$T_{thresh} = \mu_c + 2\sigma_c \tag{3}$$

In other words, if the number of comments for a given video is greater than T_{thresh} specified above, the video remains as a candidate for most appropriate video to display to the user. This step is given by equation (4) specified below:

$$N_c > T_{thresh} \tag{4}$$

Where N_c is the number of comments for a given video and T_{thresh} is the adaptive computed threshold given in Equation (3). This step filters out many

user-created videos that generate little interest and have a smaller viewership. This step also filters out certain cover videos done by other artists that filtering based simply on the views does not detect. Filtering based on the comments provide more accurate results than filtering based on viewership alone.

The video filtering process is repeated for all of the videos ordered by views (i.e., Most Popular) and by rating (i.e., Most Relevant). The remaining videos must then be matched between the two groups leaving the most appropriate or best fitting video remaining to be displayed to the user [17].

4 Video Matching and Classification

The next step in the system process is to match the filtered videos remaining from step 3 between the two existing groups – most popular (highest viewed) and most relevant (highest rating). Usage of the videoID is a key component in this step of the process. The videoID is a field returned from the JSON video data extraction process described in section 2. The videoID must be unique for every video posted at You-Tube, thus making it the primary candidate for matching videos between different You-Tube extractions. The videoID is extracted for each of the remaining videos and is compared by using a set intersection performed between the two independent groups – highest viewed and highest rated. This matching can be described by Equation (5) as:

$$\overrightarrow{M_{ij}} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} V_{v_i} \bigcap V_{r_j} \tag{5}$$

Where V_{v_i} represents the *ith* video from the viewed category, V_{r_j} represents the *jth* video from the rating category, and $\overrightarrow{M_{ij}}$ is a 2-dimensional vector indicating whether a match occurred between the *ith* video of the viewed group and the *jth* video of the ratings group. A one is placed in the given row/column if a match occurs. The two-dimensional vector now contains all the candidate videos to be considered for selection as best fitting and most appropriate video to be displayed to the user.

The possible candidates are now filtered to only one video based on the following rules:

1. Most Views
2. Most Comments
3. Highest Rating

Clearly the views should be weighted with more importance than comments or ratings. But in situations where the number of views is relatively small (less than 100,000), the comments and ratings provide additional information that should not

be ignored. This work determined that that Equation (6) provides the best results for all videos tested in the large subset of popular music songs:

$$V_m = max(0.75 * V_{v_i} + 0.25 * V_{c_i} + 0.25 * V_{r_i}) \quad (6)$$

Where V_{v_i} is the number of views for the *ith* video, V_{c_i} is the number of comments for the *ith* video, and V_{r_i} is the rating for the *ith* video. The operation described in Equation (6) is performed over all candidate videos selected from Equation (5) shown above. V_m is the video which maximizes Equation (6) for all possible candidate videos and is the video selected as most appropriate to display to the user.

5 App Implementation

All concepts outlined in the previous four sections were implemented in the form of an iOS version 9 App using XCode 7.0 and the Swift programming language. The App utilized all the top 40 popular songs from January 1st, 1970 to the present date in 2015. The App would use the song title along with the artists to select the most appropriate video based on the algorithm provided by this work. The initial screen of the app is shown below using a TableView illustrating the years covered by the App in Fig. 2 (all other years are displayed beneath 1972):

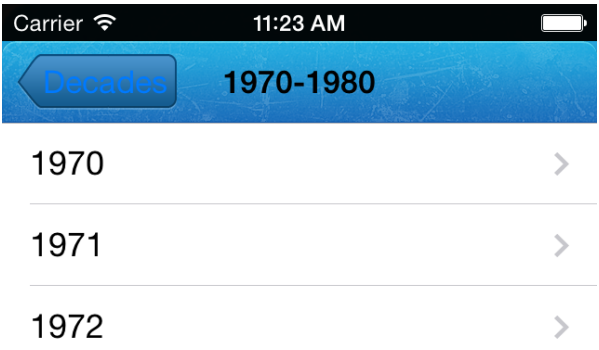


Fig. 2 iPhone App Main Screen

Pressing one of the cells pertaining to the years allows to user to select the songs pertaining to the specified year in the decade as shown below in Fig. 3. This list corresponds to the top 40 songs extracted from year 1970

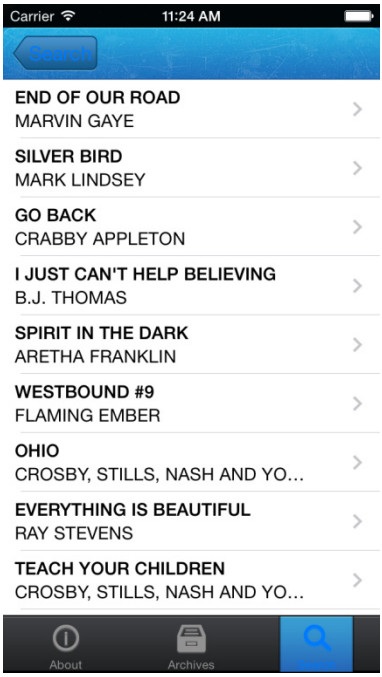


Fig. 3 iPhone App Song List

Note there are 30 more songs that can be accessed from the table illustrated below:

Pressing one of the cells of the TableView shown above invokes the algorithm described in this paper resulting in the most appropriate video displayed to the user.

6 Testing and Results

A fully function iPhone App implemented in section 5 was fully tested over 4 decades of top 40 songs. Approximately 500 samples from each decade was selected. The assumption is that a video was posted at You-Tube for each sampled video, and the objective of our test is to verify that the most appropriate video was displayed for the user. Table 1 shown below illustrates the results of the system

Table 1 Results

Years	Total	Correct	False	Percent Correct
1970-1979	421	384	37	91.2%
1980-1989	478	451	27	94.3%
1990-1999	507	488	19	96.3%
2000-2015	461	458	3	99.3%

Observe that the percent correct steadily increases as the decade year progresses. This is de the fact that there are far fewer cover videos for the more recent songs with most songs performed by the original artists. The results appear very promising illustrating the accuracy of this system. The error rate is well within bounds and provides users with a very current and up-music on-demand system.

References

1. Deng, Y., Manjunath, B.S.: Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. on Pattern Analysis and Machine Intelligence **22**(6), 939–954 (2001)
2. Air Pressure: Why IT Must Sort Out App Mobilization Challenges. InformationWeek, 5 (2009)
3. Gelasca, E.D., Salvador, E., Ebrahimi, T.: Intuitive strategy for parameter setting in video segmentation. In: Proceedings of IEEE Workshop on Video Analysis, pp. 221–225 (2000)
4. MPEG-4: Testing and evaluation procedures document. ISO/TEC JTC1/SC29/WG11, N999 (1995)
5. Mech, R., Wollborn, M.: A noise robust method for segmentation of moving objects in video sequences. In: ICASSP 1997 Proceedings, pp. 2657–2660 (1997)
6. Aach, T., Kaup, A., Mester, R.: Statistical model-based change detection in moving video. IEEE Trans. on Signal Processing **31**(2), 165–180 (1993)
7. Chiariglione-Convenor, L.: Technical specification MPEG-1 ISO/IEC JTC1/SC29/ WG11 NMPEG 96, pp. 34–82 (1996)
8. MPEG-7: ISO/IEC JTC1/SC29/WG211, N2207, Context and objectives (1998)
9. Deitel, P.: iPhone Programming. Prentice Hall, pp. 190–194 (2009)
10. Zhan, C., Duan, X., Xu, S., Song, Z., Luo, M.: An improved moving object detection algorithm based on frame difference and edge detection. In: 4th International Conference on Image and Graphics (ICIG) (2007)
11. Cucchiara, R., Grana, C., Piccardi, M.A., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(10), 1337–1342 (2003)
12. Rothganger, F., Lazebnik, S., Schmid, C., Ponce, J.: Segmenting, modeling, and matching video clips containing multiple moving objects. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(3), 477–491 (2007)
13. Day, N., Martinez, J.M.: Introduction to MPEG-7, ISO/IEC/SC29/WG11 N4325 (2001)
14. Ghanbari, M.: Video Coding an Introduction to standard codecs, Institution of Electrical Engineers (IEE), pp. 87–116 (1999)
15. Davis, L.: An empirical evaluation of generalized cooccurrence matrices. IEEE Trans. on Pattern Analysis and Machine Intelligence **2**, 214–221 (1981)
16. Gonzalez, R.: Digital Image Processing, Prentice Hall, 2nd edn, pp. 326–327 (2002)
17. Castelman, K.: Digital Image Processing, Prentice Hall, pp. 452–454 (1996)

An Application of GEP Algorithm for Prime Time Detection in Online Social Network

Hsiao-Wei Hu, Wen-Shiu Lin and I-Hsun Chen

Abstract Online Social network services, a web-based information sharing service which is utilized for public relationship and advertising promotion by brands and organizations. Since online social network is always altering, how to influence as much members as possible in community at appropriate moment will be a thoughtful problem. And strategy of article publishing decision will be important to every kind of online social network community operator.

In this paper, we focus on developing an innovative GEP algorithm, which we called Social network Gene Expression Programming (SGEP), to detect the prime time of information spreading on online social network.

Keywords Online social network · Pattern detection · Time series forecasting · Gene expression programming

1 Introduction

Facebook, among many online social network service providers, had done a remarkable work by raise its active user number and expand to the worldwide rapidly. They provide various technological affordances to their users for interacting and content sharing. Which means users' actions and links between their friends is traceable. Information which hides underneath these data is also analyzable and could be used on various fields. Fan Pages was established by Facebook, allowing users to interact and affiliate with business and organization through their

H.-W. Hu(✉)
School of Big Data Management, SooChow University, Suzhou, Taiwan, ROC
e-mail: camihu@gmail.com

W.-S. Lin · I-H. Chen
Department of Information Management, Fu-Jen Catholic University,
New Taipei City, Taiwan, ROC
e-mail: wslin1949@gmail.com, chenesing@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_71

Facebook profiles [1]. It is a specific function like themed community on Facebook and can be set up for a brand, entity or public figure. Fan Page is mainly used to establish positive impression or marketing strategy promotion, especially customer relationship management [2]. Brands and organizations usually utilize it to update news and forthcoming events, yet users are required to subscribe their channel. Fan Page's owners can publish post concerned with their business information or personal narrative through text, photos, videos, etc. User who had liked their pages will be notified the update and capable to interact with them. In 2011, more than 50% of social media users follow brands on social media [3] and companies are increasingly investing in social media.

Brand fans can have positive or negative comments under page's posts; consumer's discussions may do positive affecting value perceptions of a product [4]. However, although there are couple researches have analyzed the Fan Page [2] and few studies about predicting in social networking [5] [6], but rare of them have evaluate the influence of social network and evolve with forecasting. Fan Page's effectiveness is already been demonstrated with many practical example [7] [8], it provides an easier way to people to receive more information they interest in, regardless of business or entertainment dimension.

We apply Gene Expression Programming (GEP) algorithm to forecast future movement of Fan Page on Facebook, adapting the chromosome from determined feature factors which represent the user activities on social network. This adapted algorithm can be contributed to assist the research of online social network and business use. Also, it could improve the efficiency of brand promoting and make a better posting strategy. We will first introduce Fan Page of Facebook as our target for analysis and how to implement Prime Time Detection method, Influence Score formula and GEP algorithm. The proposed Social network Gene Expression Programming (SGEP) methodology and strategy design will then be drawn out. Experiment design, results and comparison with other methods are discussed.

2 Preliminaries

2.1 Prime Time Detection

We define "right people" and "right time" to characterize Prime Time. "right people" means people have better influence that can spread the information rapidly, and they could directly or indirectly affect other users in very short time. The "right time", which means the prime time that Fan Page community can reach the maximum amount of right people.

2.2 Influence Score

We defined influence of social network as impacting right members of community at right time at previous section. To measure the influence, we have to analyze the community structure of our target, Fan Page. Owing to the movement of users will

disseminate information from Fan Pages to their friends and members of this community. When they submit their comments on the page, system will arrange comments hierarchically as Fig. 1 shown and allow people to publish comments below it, as well as function like.

We define these followed comments and like as "Re-Comment" and "Re-Like", which means they are the response of comment. To evaluate these activities in online social network and analyze them to provide a dependable result, we use Influence Score formula (IS) in (1) from Prime Time Detection method to calculate the influence. Influence Score is concerned with the relativity between three major indicators, "Like", "Comment" and "Share".

"Like" and "Comment", was found that comment, even it is short, is more gratifying to receiver than a one-click only communication experienced as like does [9]. The relativity between "Share" and "Comment" in Facebook are functional similar to "reply" and "retweet" in Twitter. Reply and comment both are easily engage in conversations in a smaller group, have lesser helpful to information dispersing. On the contrary, share and retweet have various incentives to prompt users to do so, such as request from originator or the interest of time-sensitive material can help these trending topics spread further [10]. Therefore, we can consider that the weight of importance between them will be Share > Comment > Like.

$$W_S > W_c > W_L$$
$$IS = Likes \times W_L + Comments \times W_C + \mathbf{Share} \times W_S \quad (1)$$

2.3 Gene Expression Programming

Gene expression programming (GEP) could be apply on various area, such as predicting house price [11], stock market [12], currency exchange [13], also classification rules [14] and data mining [15]. It is flexible and can be used to combine with other algorithm to approach the most appropriate solution of problem.

Fig. 1 Comments arrangement on Facebook.

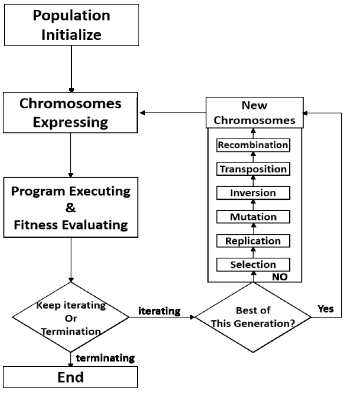


Fig. 2 The flow chart of GEP algorithm

GEP was invented by Ferreira in 1999 [16], and could be designed by computer programs or models as Fig. 2 presented.

With complicated tree structures that can learn and adapt through changing their size and composition, it incorporates both the simple, linear chromosomes of fixed length similar to the ones used in Genetic Algorithm and the ramified structures of different sizes and shapes similar to the parse trees of Genetic Programming method.

3 Methodology

3.1 Indicators

As we mentioned above, Like, Comment and Share are three major indicators of user's movement. Therefore, we take them as the feature factor of our ADF (Automatically Defined Functions) structural design in Table 1. To capture the factors and time sections from Fan Page, we develop an application to collect data from Fan Page through Facebook Graph API, FQL and web crawler. We define "Like" and "Comment" indicator as two divisions; Like and Re-Like, Comment and Re-Comment.

Table 1 Indicators used in SGEP

Indicators	**Like**	**Comment**	**Share**
Indicator variables	Like	Comment	Share
	Re-Like	Re-Comment	
ADFs	ADF_1	ADF_2	ADF_3

3.2 Social Network Gene Expression Programming (SGEP)

Combining Prime Time Detection method and gene expression programming algorithm to obtain the strategy that our model derived from training set through using slide window scheme. Each ADF is designed to reflect the problem that we intend to solve, we have three different ADFs, ADF (Like), ADF (Comment) and ADF (Share).

Fig. 3 is an instance of expression tree of our ADFs, and chromosome structure of each ADF is schemed in Fig. 4 ADF (Comment) for example, factor Coms and ReComs shall overpass the threshold given by RNC (Random Numeric Constants), selecting constants from historical data randomly. And utilize these ADFs to compose the structure of our SGEP model.

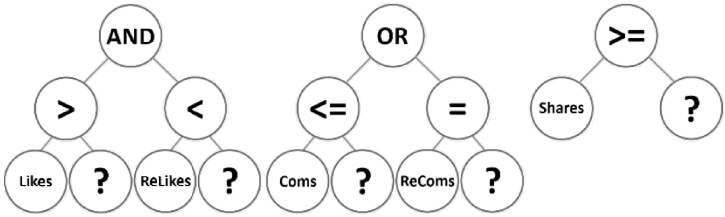


Fig. 3 Instances of ADFs Expression Tree type.

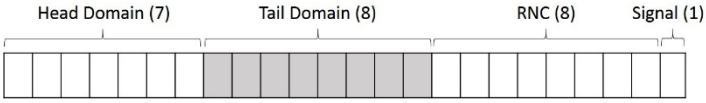


Fig. 4 Chromosome structure of ADF_2 (Comment).

GEP model will generate the initial population as the first generation, then evaluate the fitness value of each entity and will keep iterating before termination. As Fig. 3 depicted, each ADF will be filled by variable factors, and "?" will be furnished by RNC function from other time section before this detecting objective. We utilize ADF to compose the structure of our SGEP model in Fig. 5; each ADF may send a signal of confirmation (true or false) after its self-computing. Collecting them and follow Table. 2 to determine whether this time section is the prime time to post message or not. Only 2 or 3 in total will be deemed as the prime time of this Fan Page. On the other hand, 0 to 1 will be regarded as negative signal to posting strategy. After accomplished all processes of GEP and terminated, we will discover the best rule of predicting Fan Page's activities.

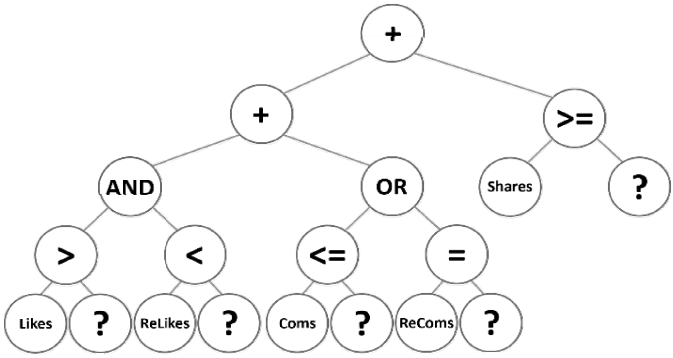


Fig. 5 Signal summing procedure of ADFs.

Table 2 Total Signal of Posting Supportive

Total Signal	0-1	2-3
Result	Negative	Positive

3.3 Strategy Design

We develop our Social Gene Expression Programming of Prime Time Detection method as training and testing phases.

In Fig. 6, at beginning, initializing three factors ADFs' structure configurations and simulate population of chromosome, then input data from training group and calculate Influence score. In order to poise the actual influence of each indicator, we give them weights in three distinct distributions with incremental series and geometric progression to reflect practical situation in Table. 3 and collect returned hits at comparing. Owing to Re-Like and Re-Comment are users' feedbacks belonged to factor Comment. Theoretically, factor Re-Like's influence level will higher than factor Like; and factor Re-Comment's influence level will higher than factor Comments as well. Hits in (2), a foundational arrangement for fitness function to calculate, means one chromosome has predicted correctly and accumulating at the same time which will be returned to fitness evaluating. Providing hits to fitness value computing for advanced particular choosing at this generation, make it rejoining evolution procedure with training period progressing until the last day of training. Result, after training period, we will achieve the best strategy of this training design which is our prime time detection strategy as well.

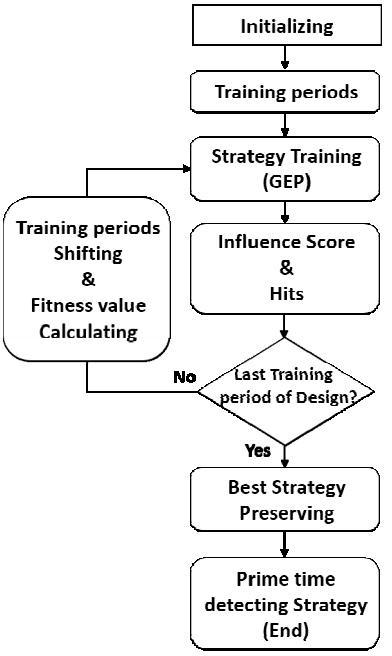


Fig. 6 Strategy Training flow chart.

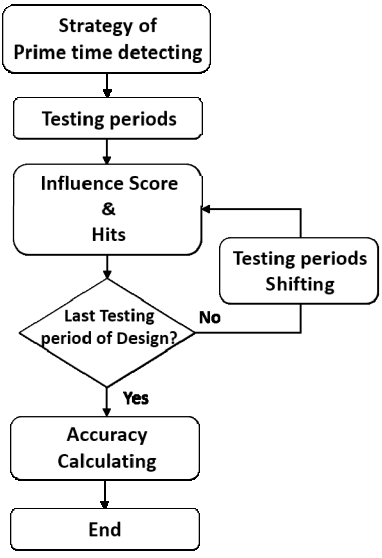


Fig. 7 Strategy Testing flow chart.

Using discovered strategy of prime time detection from training periods to testify the strategy effectivity, with progressing of testing periods, Influence score calculating, hits cumulating and fitness value computing. With this gradation, accuracy of strategy will be considered as fitness value; we can acquire each inherited chromosome's accuracy of forecasting and distinguish the most refined strategy in Fig. 7.

Table 3 Distribution of Weight

Case	Like	Re-Like	Comments	Re-Comment	Share
Case01	1/5	2/5	3/5	4/5	1
Case02	1/3		2/3		1
Case03	1_1	2	4	8	16

$$Hits \begin{cases} Signal\ Positive & \begin{cases} if\ IS_{t_x} < IS_{t_{x+1}} & , Hits + 1 \\ if\ IS_{t_x} > IS_{t_{x+1}} & , Hits = Hits \end{cases} \\ Signal\ Neegative & \begin{cases} if\ IS_{t_x} > IS_{t_{x+1}} & , Hits + 1 \\ if\ IS_{t_x} < IS_{t_{x+1}} & , Hits = Hits \end{cases} \end{cases}$$

$$Fitness = Hits / Length_{p_x} \begin{cases} p_1 : Training\ Period \\ p_2 : Testing\ Period \end{cases} \quad (2)$$

Model Comparing. To corroborate our model's practicality and improvement of social network evolution predicting, we decide to employ simple linear regression as trend forecast method (TFM) and imitating normal user behavior through random simulating (RS) decision. Our Social network gene expression programming (SGEP) model will be executed, then taking result to compare with TFM and RS respectively at the end of research

4 Experiment Result

4.1 Sliding Windows

We collected social activities from Facebook's target gaming Fan Page, embracing more than ten million active users and regular posting interval. Starting from 01/02/2013 to 12/31/2014, we develop three different type of strategy training experiments, having varied length of training periods but same testing periods of one week. Experiment A has one month training period, Experiment B for three months and Experiment C for six months. Training period will be shifted like sliding window till the day before last available testing period.

4.2 Weight of Indicators and Parameters Test

We have schemed three different weighing distributions at previous chapter. To quantify effect of these indicators, each of them had processed five times and ranked through their result. Regular parameters based on Ferreira's (2001) research was proposed to solve different type of problems, we utilize these settings (like elitist selection, inversion rate, and elitist selection rate usually been set as 3% and 5%, inversion rate was 10% and 30%.) to train our model and abstract the most proper modulus.

Mutation rate usually been arranged between 2.5%, 4.4%, 10% and crossover rate is 10% and 30% to solve most part of problems, for that matter, we execute these settings and examine how mutation and crossover will impact the result. Therefore, when inversion, elitist, mutation and crossover rate are regulated as 30%, 20%, 2.5% and 30%, it may have the better consequence at mostly conditions. Accordingly, we will proceed our experiment through these discoveries.

4.3 Experiment Result

Experiment A (one month training), B (three months training), C (six months training) have 65.1%, 66.3%, 67.3% in averaged accuracy result respectively at Table 4. We could acknowledged that Experiment C is the most preferable experimental setting from other experiments, which means training period length does influent the accuracy of prediction in our model. We have noticed that there were lesser users and simpler message posting at 2013, users' behavior were not that complicated as 2014 did. Fan Page managers improved their competence of spreading information through multimedia, cross-industrial topic manipulating, charity cooperative events, etc. These improvements is possible to be the reason to make 2014 become more difficult to predict, and decrease the accuracy of our model slightly.

Table 4 Summary of SGEP

	Experiment A	Experiment B	Experiment C
2013	66.5%	67.7%	70.4
2014	63.6%	64.9%	64.1%
Average Accuracy	65.1%	66.3%	67.3%

4.4 Comparison of Diverse Models

To verify our method's amelioration, we deployed TFM and RS to comparing our SGEP method. TFM has incremental result from 44%, 52% to 57%, we can learn that with linear regression model has adjusted each data point deviation and returned, it has stable growth and positive correlation with longer training period at

Experiment C. Random Simulating, on contrast, intend to imitate user's personal decision without any predicting factors involved. It shows low-estimation of accuracy owing to random decision making simulation, which may be considered that user's subjective judgement won't have any improvement to future social movement forecasting. In this case, Experiment A has 46.6%, B has 46.9%, but C has only 40.9%; unlike other models we applied before, training period length is not related with the result obviously.

Fig. 8 exhibits three methods' performance above, comparison of SGEP, TFM and RS. Unlike SGEP and TFM have conspicuous progressing with longer training, RS represents anomalistic instead. It shows the importance of historical data accumulating to SGEP and TFM, even some conditions may influence model's establishment. Our SGEP has the best performance among them, which we can deduce that SGEP is a better solution to acquire pattern of online social network and has meliorated the effectiveness of predicting practicability.

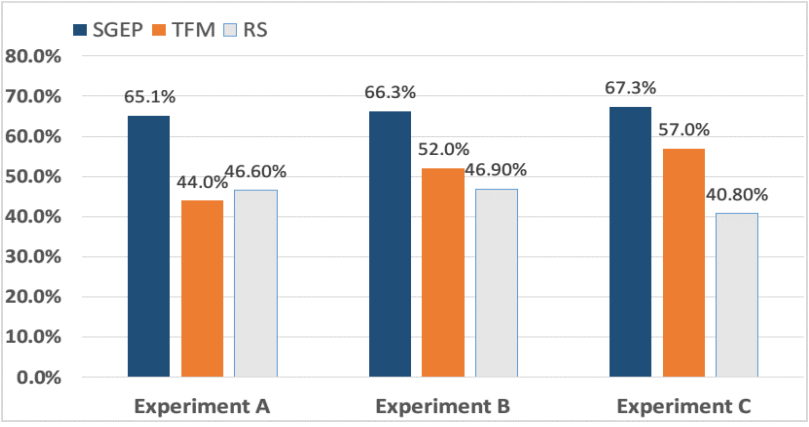


Fig. 8 Trend of variation of Mutation rate & Crossover rate.

5 Conclusion

In this research, we purpose a method for predicting the social network movements in Facebook, inaugurating with current circumstance of social network, the reason why we choose Fan Page as target, analyze user's activities from the sentimental reaction perspective and how online social network impacted marketing strategy. Taking several unique function at Fan Page as our detecting object, define indicators to evaluate user movement through Influence Score equation in Prime Time Detection method and combine with Gene Expression Programming algorithm to derive new Social network Gene Expression Programming algorithm. SGEP can develop posting strategy to Influent central members of community at opportune timing, dynamically. It has substantiated through Trend Forecast Method and Random Simulating of user behavior, outstand from them and achieve 65.1% to 67.3% of predicting accuracy.

In the future, SGEP will evolve more refined strategies to increase its predicting accurateness and accumulate more indicator information to advance the RNC threshold selecting, giving this feature selection more possibility to adapt higher affiliated with future environment. With Mining and collecting patterns at various time sections, we can diversify our interval at historical data among modeling period and experiment period. Furthermore, we may combine with sentimental and word-of-mouth analysis from text mining to advance the accuracy of our SGEP. In addition, mining data at other online social network to modify indicators of SGEP and deploy it for prime time detecting.

References

1. Hof, R.: Facebook Declares New Era for Advertising, Bloomberg Businessweek. November 2007 http://www.businessweek.com/the_thread/techbeat/archives/2007/11/facebook_declar.html(accessed February 14, 2015, from the World Wide Web)
2. Berthon, P.R., Pitt, L.F., Plangger, K., Shapiro, D.: Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. Business Horizons **55**(3), 261–271 (2012)
3. Van Belleghem, S., Eenhuizen, M., Veris, E., Social media around the world (2011). http://www.slideshare.net/stevenvanbelleghem/social-mediaaround-the-world-2011/download?lead=394fd930572c9b62fb082021af5a6d0922046ec4 (accessed November 2014, from the World Wide Web)
4. Gruen, T.W., Osmonbekov, T., Czaplewski, A.J.: eWOM: the impact of customer-to-customer online know-how exchange on customer value and loyalty. Journal of Business Research **59**(4), 449–456 (2006)
5. Trusov, M., Bucklin, R.E., Pauwels, K.: Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. Journal of Marketing **73**(5), 90–102 (2009)
6. Asur, S., Huberman, B.: Predicting the future with social media. In: 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE, vol. 1 (2010)
7. Lee, W., Xiong, L., Hu, C.: The effect of Facebook users' arousal and valence on intention to go to the festival: applying an extension of the technology acceptance model. International Journal of Hospitality Management **31**(3), 819–827 (2012)
8. Zaglia, M.E.: Brand communities embedded in social networks. Journal of Business Research **66**(2), 216–223 (2013)
9. Wang, Y.C., Burke, M., Kraut, R.E.: Gender, topic, and audience response: an analysis of user-generated content on Facebook. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 31–34. ACM (2013)
10. Boyd, D., Golder, S., Lotan, G.: Tweet, tweet, retweet: conversational aspects of retweeting on twitter. In: 2010 43rd Hawaii International Conference on System Sciences (HICSS), pp. 1–10. IEEE (2010)
11. Shekarian, E., Fallahpour, A.: Predicting house price via gene expression programming. International Journal of Housing Markets and Analysis **6**(3), 250–268 (2013)

12. Băutu, E., Bautu, A., Luchian, H.: Evolving gene expression programming classifiers for ensemble prediction of movements on the stock market. In: 2010 International Conference on Complex, Intelligent and Software Intensive Systems (CISIS). IEEE (2010)
13. Sermpinis, G., Laws, J., Karathanasopoulos, A., Dunis, C.L.: Forecasting and trading the EUR/USD exchange rate with gene expression and psi sigma neural networks. Expert Systems with Applications **39**(10), 8865–8877 (2012)
14. Karakasis, V.K., Stafylopatis, A.: Efficient evolution of accurate classification rules using a combination of gene expression programming and clonal selection. IEEE Transactions Evolutionary Computation **12**(6), 662–678 (2008)
15. Karakasis, V.K., Stafylopatis, A.: Data mining based on gene expression programming and clonal selection. In: IEEE Congress on Evolutionary Computation, CEC 2006, pp. 514–521. IEEE (2006)
16. Ferreira, C.: Gene expression programming: mathematical modeling by an artificial intelligence, vol. 21. Springer (2006)

Transport Logistic Application: Train's Adherence Evolution and Prediction Based on Decision Tree and Markov Models

Steve Ataky T. Mpinda, Marilde T.P. Santos and Marcela X. Ribeiro

Abstract In this paper it is presented a method for modelling adherence and prediction of train evolution based on classification tree and Markov Models. In day-to-day of railway operational management of traffic it is recurrent to face a deviation from the theoretic planned schedule to a certain space-time. When such deviation happens, what we also call non-adherence, there is a need for a quick decision-making to avoid, on the one hand, that primary initial non-adherence leads to a whole cascade of secondary non-adherence of other trains over the planned schedule, on the other hand that event such as a train reaches another may occurs, what is a considerable issue. We present an effective mixed approach compounded of a stochastic process (Markov Models) and classification tree for estimating and predicting the reliability of the adherence of planned and realized traffic schedule of not only rail freight transport, but also public transport, considering some parameters that may cause deviation such as weather, environmental conditions, and so on. The experiments on real data for trains of Brazilian regions show that our model is fairly realistic and deliver good results in short processing time. Moreover, the model may run for an on-line scenario where updated data are massively collected from monitoring sensors. Withal, the prediction's accuracy along with the evolution of probability distributions regarding all events over time have been evaluated. Therefrom the approach reveals to be good for predicting for train traffic evolution based on historic data and monitoring collected data as well as weather and environmental ones. As a result we obtain the increasing reliability of about 74% of prediction.

1 Introduction and Motivation

The railway traffic density tends to increase up to a point to cause the infrastructure saturation in many areas. This densification, above all, became possible, in

S.A.T. Mpinda(✉) · M.T.P. Santos · M.X. Ribeiro
Department of Computer Science, Federal University of São Carlos (UFSCar),
São Carlos, SP, Brazil
e-mail: {steve.mpinda,marilde,marcela}@dc.ufscar.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_72

part, through the scheduling optimization tools. However, the referred densification renders the incidents and disturbance most frequent, and its consequences more significant or at least much more difficult to manage effectively. The planning and real-time traffic management therefore need more and more tools to support appropriate decision.

Schematically, the problem can be summarized as follows: due to the occurrence of one or more events (accidents, weather problems, etc.), theoretical schedules are no longer achievable and a decision must be made, since on the same rail (route) there are several planned circulations, and the incident may somehow influence the realization of the planning of other circulations and eventually lead to an increased operating. The first step to confront this disturbance is basically to identify what were the causes or generator events of the incident (statistical treatment of incident categories, involving the place, date, etc.), so based on them analyze the direct and indirect consequences and then provide an improved re-planning (scheduling) estimation of train circulations to maintain the adherence [1] between the initial planned schedule and the new one, taking into account the operating constraints to be absolutely observed.

Furthermore, the train evolution accurate prediction is an important requirement for proactive and anticipative real-time control of railway traffic and transport [7]. Traffic controllers need to predict the evolution of trains within, or heading towards, their area as well as monitor them in order to manage the feasibility of timetable realisation. Wherefore, valid train evolution estimations are important for preventing or reducing non-adherence. hitherto, real-time traffic based models and transport control have mostly focused on overcoming the great combinatorial complexity of train rescheduling [4, 11], delay management [6] and rolling-stock and crew rescheduling [12]. The emerged methodologies have the capacity to tackle complex instances progressively, notwithstanding they commonly expect flawless deterministic knowledge of the input traffic state and consequent activity advancement.

As of late, the instability of train event times has been perceived as one of the significant hindrances for registering possible and implementable solutions for rescheduling issues in railroad movement [5]. The unpredictability of an event is typically represented by the likelihood dispersion of its realization. Nonetheless, a large number of the current methodologies assume fixed likelihood dispersion for the estimation of process times and do not consider the impact that real-time information on train positions and non-adherence may have on the relating dispersions, let alone the influence that some factors such as the weather and environment conditions may have.

2 Related Work

Müller-Hannemann and Schnee [9] have worked on efficient deterministic propagation of primary and secondary delays, wherefrom the authors revealed that even

[1] The term adherence is used to mean that the train circulation (in) realization is "being faithful" to the theoretic planned schedule and non-adherence otherwise.

enormous delay data streams can be spread straightaway, making this methodology attainable for real-time multi-criteria timetable information. Meester and Muns [10] and Carey and Kwieciński [2, 3] have studied in-depth stochastic models for the propagation of delays where they propose to utilize approximations of delay dispersions to diminish the computational exertion and study the error propagation for such approximations. A realistic stochastic model for delay propagation and calculation of arrival and departure time distributions in public transport was proposed by Berger et al. [1]. Mohammad [8], however, enhanced the formulas of [1] who concentrated on trains, not on entire connections, and did not investigate reliability, to calculate probability distributions for connections consisting of several trains and transfers between them. A similar approach has been taken by Kecman et al. [7] who present a method for modelling uncertainty of train delays based on a Markov stochastic process, where the dynamics of a train delay over time and space is presented as a stochastic process that describes the evolution of the time-dependent random variable. Compared to our work, his experiments are only based on traffic realisation data from the part of the passengers and high-speed corridor. We enhanced their modelling increasing rail freight transport as well as possible parameters which may cause disturbance such as weather and environmental conditions.

3 Proposed Architecture

In this section, a generic architecture is proposed which allows, besides the data with respect to the initial scheduled traffic and the realized (only in the end of the circulation or train run), obtaining the intermediate data, namely at an instant t, the process also called "monitoring", in order to verify whether everything is as planned, this is, the train run is adherent to the planned schedule. Also, it embodied the possibility to predict the possible sequential planning of traffic based on the historical data and monitoring information such as weather and environmental conditions.

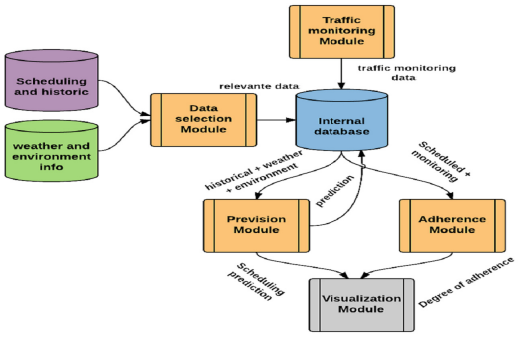


Fig. 1 Proposed General Architecture

The Data selection Module select relevant stored data from "scheduling and historical" and "weather and environment" databases for the process of forecasting and adherence calculation. Train scheduling scenarios involve a variety of data used for the prediction and adherence calculation process such as itinerary, departure time, average speed, direction, type of train, priority. Besides, the weather and environment conditions are collected, as their variation may cause disturbance and uncertainty. Traffic Monitoring Module, in turn, consists of collecting information (preferably in real-time) related to the trains circulation over a segment, as well as its current state, which implies the obtention of arrival and departure times of each vehicle in each segment, as well as its average speed. The Internal database is used as a storage system of the train circulation scenario from the origin to the destination, at the segments delimited and precise. therefore, it consists of a collection of data and / or information from the "selection modules" (data relevant to current and historical train schedules), "Monitoring" (data for the scenario that is effectively occurring).

The Adherence Module, in turn, initially, carries out adherence verification, i.e. verify whether there is a bi-univocal correspondence between the scheduled planning and current situation (monitoring). This process compares each scheduled data with its correspondent (realized or in realization). For example, knowing the current time and position of the train, verify that it converges to what was scheduled taking into account the pre-defined degree of freedom. Obviously, the non-adherence makes it clear that an event occurred, which can be foreseen or unforeseen, whose impact negatively influenced in making the initial planning possible. In this purpose, it is essential to identify the incident (what happened?), categorize its impact (what impact?) And, finally, if necessary, generate data for re-planning. On the other hand, Prediction Module is responsible for registering the unexpected event occurrences, as well as environmental factors (climatic and non-climatic) along the route of the trains, and, based on them, predict future behavior of the traffic. Furthermore, the prediction should allow to have control of what may be realized with the highest degree of probability.

4 Description of Sequential Data Analysis

This section presents the methodology of modelling train circulation as sequential data analysis. Each circulation in each segment consist of a set of heterogeneous sequential data according to the Figure 2

Let $S = \{S_1, S_2, \ldots, S_n\}$ be a set of n sequences which may be of different lengths, where each sequence $S_i = \left\{e_{i,1}, e_{i,2} \ldots, e_{i,T_i}\right\}$ is given by a set of T_i states $e_{i,j}$ observed successively by a train i and j from 1 to T_i informs the observation time of the event e_{ij}. Each state $e_{i,j}$ is described over a set of p heterogeneous variables $Y = \left\{Y_1, Y_2, \ldots, Y_p\right\}$. The classification process considered in this work is to structure the sequences contained in $S = \{S_1, S_2, \ldots, S_n\}$ in terms of their similarities, under as a set of homogeneous and significant classes. In this case where there are complex sequences (Figure 2), where, for example, in the segment

AB the estimated time is 137min according to the following conditions: average temperature, no rain, moderate wind, sunny weather, good rail condition, no fallen leaves on the rail, no water in the rail, and the circulation occurs by the morning. In this situation, it is necessary to structure these data as well as homogenize them in order to exploit them for decision making, which is the goal of this work. Thus, it is considered, that each state $e_{i,j}$ is no longer described by a set of p heterogeneous variables, but by a group $g_{i,j}$ obtained by classification of p descriptive variables. The sequence constructed becomes, therefore, $S_i = \{g_{i,1}, g_{i,2}, \ldots, g_{i,T_i}\}$. In this work such homogenization is called *Homogeneous Group of Adherence* (GHA).

It is noteworthy that in this case it is known in advance the number of classes that characterize trajectories studied, to wit : on time (N), delayed (T) and in advance (D). The approach is applied to sequences of any length whose states are described by heterogeneous variables and use the modified Edit Distance called *DEFlex* to better reflect the presence of common states between the sequences of different lengths.

From another perspective, to better explain the dimensions related to states-circulation, it is necessary to cross the variable *adherence group* (GHA) with the variable *average speed* (VM) in the description of the states of the sequences performed. In practice, the GHA represent the operational aspect, while the VM provides the best idea on the train state. Table1 shows the example of some circulation trajectories.

Table 1 Trajectory of circulation with crossed informations.

Trajetory (S_i)	State 1	State 2	...	State n
S_1	GHA-VM-1	GHA-VM-2	...	GHA-VM-N
S_2	GHA-VM-1	GHA-VM-2	...	GHA-VM-N
S_3	GHA-VM-1	GHA-VM-2	...	GHA-VM-N
...	...	...	...	...
S_n	GHA-VM-1	GHA-VM-2	...	GHA-VM-N

Where: **GHA**: Homogeneous Group of Adherence can take the values (N: *on time*; T:*delayed*; D:*in advance*), of symbolic data type. **VM**: average speed, of classical data type. Given a qualitative nature of data obtained from the sates $g_{i,j}$ of sequences S_i to analyze, the Edit Distance is suitable to evaluate the similarity between the temporal sequences. In practice, the problem of evaluating the edit distance between two time sequences (considered as two strings) is a widespread problem of evaluating the length of the longest common subsequence (LCS) between these two sequences. However, this distance has certain limitations to be applied in the context studied in this work. The edit distance is as shown in the equation 1 and was modified as in the equation 2.

$$d_E(S_i, S_j) = |S_i| + \left|S_j\right| - 2 * LCS(S_i, S_j) \tag{1}$$

$$d_{Flex}(S_i, S_j) = 1 - \left[\left\{X - 2 * LCS(S_i, S_j)\right\}\right] * X^{-1} \tag{2}$$

$$d_{Flex}(S_i, S_j) = 1 - \frac{\{X - 2 * LCS(S_i, S_j)\}}{X}$$

where:

$$\begin{cases} X = |S_i| + |S_j|, & \text{considering that } |S_i| \neq |S_j| \\ X = 2 * min(|S_i, S_j|), & \text{considering that } |S_i| = |S_j| \end{cases}$$

where $min(|S_i, S_j|)$ is the shortest lengths between S_i e S_j knowing that $|S_i, S_j| \leq |S_i| + |S_j|$. This is useful for calculating the distance between two sequences, taking into account the equidistant trajectories.

5 Case Study

The architecture presented in the section 3 is used to model adherence evolution of trains running over rail freight lines in Brazilian regions. These lines usually are in non-urban areas where environmental conditions require special care. Moreover, in such regions, during the periods of rainfall and wind flooding may occur and obstacles such as (mud, fallen tree, foliage or fallen leaves, etc.) may be found.

There are in total 1235 collected data, this is, 1235 train circulations with respected to almost 2 years.

In order to discover the sequences patterns, trains behavior in different circulations over the same route, categorize the events (incidents) based on their impact, a data sheet built representing all scenarios of the sample train trajectories obtained in the database. In practice, we found some factors or parameters that have a key role in determining the trains behavior during the circulation, leading to degree of adherence (on-time, delayed, in advance) arriving to the end of the trajectory. These factors are: temperature, rain, wind, weather, rail condition, foliage and water on the rail and, but not necessarily, the time of the day, and each of the factors can have different values. For example, the temperature may have values (symbolic) as very hot, hot, medium, with numeric ranges defined for each value. An instance, combination of the attributes, in turn, determines a possible degree of adherence of a trajectory em comparison to the planned schedule. To make sure of that, a data mining process was performed, the built dataset was trained in order to discover knowledge behind this massive range of data. As a result of performed in training, we obtained accuracy of 91% after 100 training, what means that the generated classification model is considered efficient.

5.1 Railway Configuration and Train Planning Scenario

Let the Figure 2, represent the macroscopic configuration of a railway, railway configuration with estimated distance of segment and train planned schedule with estimated duration related to the distance of the segment.

Let suppose that a train realized different circulation over the same network in different period of time. The Figure 2 exemplify this scenario where the first object (blue vehicle) is the planned schedule and the followed three ones (orange vehicles) are some of the effectively realized circulations. It is evident to observe the variation in term of duration time spent over each segment comparing to its correspondent in other circulations as well as the theoretic planned schedule.

Such variations bring us to find out the cause of such deviation between planned schedule and effectively realized. Moreover, it is important to understand the meaning of aforementioned deviation, in term of adherence and non-adherence, what is one of the objective of this work. To do so, based on the Table 2 one can determine whether the spent time of the realized circulation is on-time, delayed or in advance regarding the planned one. Thereby, there are have three adherence classes which are on-time, delayed and in advance that will be henceforth represented by the letters N, T and D respectively. Hence, one says that the realized circulation is adherent if and only if its spent time over a determined segment taking into account the estimated time (planned time) belongs to the on-time degree of freedom, and non-adherent otherwise.

Table 2 Degree of freedom Table

Segment	estimated time	on-time	delayed	in advance
AB	150	$120 < t \leq 165$	$t > 165$	$t < 120$
BC	240	$192 < t \leq 264$	$t > 264$	$t < 192$
CD	150	$120 < t \leq 165$	$t > 165$	$t < 120$
DE	214.2	$171.36 < t \leq 235.62$	$t > 235.62$	$t < 171.36$
EF	360	$288 < t \leq 396$	$t > 396$	$t < 288$

From this point, the main purpose of the present paper is to find out the possible causes of non-adherence and examine their effect over the scenario. Likewise, once detected (in advance) such causes and effects that prediction horizon and incoming information about a running train may have on the predictability of subsequent evolution of that train, it is possible to offer a more clear understanding of how does the probability distribution of adherence of an event change over time when detected causes that lead to non-adherence occur.

5.2 Data Operations and Adherence Calculation

Given the availability of the dataset, such as planned itineraries of trains, data regarding the evolution of the trajectory realized until an instant t (monitoring module), both indirectly from internal data repository, it becomes hence possible to carry out the procedures for the adherence calculation, that is, verify whether the planned scenario corresponds to what is happening or not. Moreover, one can make the prediction

based on the information contained in the dataset, as well as the discovered patterns, and current environmental and weather information.

Suppose the scenario of Figure 2 where we have the planned schedule with estimated values and three different circulations carried out by train T1 with heterogeneous information on weather and environmental relevant factors.

The Figure 2 and 3 represent the trajectory planning to be traveled by train T1 and different circulations carried out with information on weather and environmental relevant factors and Planning and trajectories traveled by train T1 and different circulations realized with information on the degree of adherence, respectively.

Taking into account the homogenization presented in the section 4 and considering the Table 2 of degree of freedom of the section 5.1 to determine the adherence, such scenario can be represented according to the Figure 3

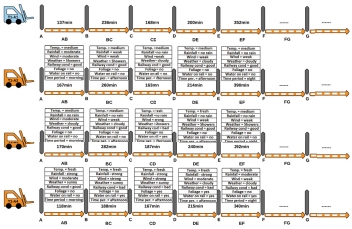


Fig. 2 Trajectory planning to be traveled by train T1

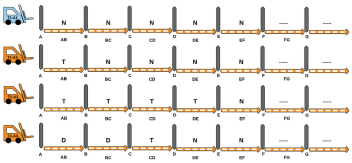


Fig. 3 Planning and trajectories traveled by train T1

Thus, one can, from Figure 1, forms strings gathering degrees of adherence in respect of each circulation. In this case, we will have four strings as follows:

1. *Planned schedule* **T1-A1**: (NNNNN).
2. *First circulation* **T1-A2**: (TNNNN)
3. *Second circulation* **T1-A3**: (TTTTN)
4. *Third circulation* **T1-A4**: (DDTNN)

To discover the adherence of each circulation performed with respect to the planned schedule, it is necessary to verify how much similar they are. To do so, given these strings, it was used proposed distance as follow:

– The adherence degree between the first circulation realized **T1-A2** em comparison to the planned **T1-A1** is:

$$d_{Flex}(T1 - A2, T1 - A1) = 1 - \frac{|5| + |5| - 2 * 4}{|5| + |5|} = 0,8$$

This is, T1-A2 is 0,8 or 80% adherent to T1-A1.

– The adherence degree between the second circulation realized **T1-A3** in comparison to the planned **T1-A1** is:

$$d_{Flex}(T1 - A3, T1 - A1) = 1 - \frac{|5| + |5| - 2 * 1}{|5| + |5|} = 0,2$$

This is, T1-A3 is 0,2 or 20% adherent to T1-A1. And so on.

On the other hand, suppose that we have the following circulation: let T1-A5 = (DDTTDD) be a realized circulation whose intersection with T1-A0 (NNNNNN), initial planning is null. Applying the original editing distance we have:

$$d_E(T1 - A5, T1 - A0) = |6| + |6| - 2 * 0 = 12$$

We may note that the value 12 does not make clear its meaning therefore inconclusive in the context of this work. Notwithstanding, applying the modified and proposed distance we have:

$$d_{Deflex}(T1 - A5, T1 - A0) = 1 - \frac{|6| + |6| - 2 * 0}{|6| + |6|} = 1 - 1 = 0$$

This clearly explains that the two sequences are disjoint, or completely dissimilar, since the adhesion level range is from [0, 1], where 0 refers to the total non-adhesion and 1 total adherence.

In addition to the previous scenario, there is another case where one wants to verify a circulation adherence between its initial and ongoing planning. Let (T, N, N, ?) be the ongoing planning.

In that case that we have two different lengths, using the first variant of the modified distance we get:

$$d_{Flex}(T1 - A3, T1 - A1) = 1 - \frac{|3| + |5| - 2 * 2}{|3| + |5|} = 1 - 0.75 = 0.25$$

But if we need to verify and calculate the adherence of the ongoing circulation with its corresponding planned, i.e., of equal length, which leads to use the second variant, which condition the dissimilarity and adherence verification between two sequences with equal lengths.

Applying this for the monitoring case we obtain:

$$d_{Flex}(T1 - A3, T1 - A1) = 1 - \frac{(2 * 3) - 2 * 2}{2 * 3} = 1 - 0.33 = 0,66$$

This is, the ongoing circulation is 66 % adherent to the corresponding planned. Therefore, we can make certain decisions if necessary because it has been perceived the actual situation.

Prediction using probabilistic and Markov Models

For the prediction, we need first to understand what is the probability and quantify it. Secondly, it is necessary to find the relation between the probability and ignorance (lack of information) and finally distinguish causality and correlation. The Table 3

illustrates the correlation with respect to the wind and rainfall regarding the collected dataset. This is done for all other considered parameters (temperature, foliage, etc.)

Table 3 Probabilities od transition based on parameter rain

Wind				
Adherence degree	Violent	Strong	Moderate	weak
On time (N)	0.0%	7%	50%	76%
Delayed (T)	100%	93%	50%	10%
In advance (D)	0.0%	0.0%	0.0%	14%
Rain				
Adherence	Strong	Moderate	Weak	no rain
NN	0.0%	0.0%	66.6%	78%
NT	0.0%	0.0%	33.3%	16%
ND	0.0%	0.0%	0.0%	0.7%
—	—	—	–	—
TT	99%	99%	20%	38%
TN	0.0%	0.0%	60%	53%
TD	0.0%	0.0%	0.0%	0.0%
—	—	—	–	—
DD	0.0%	0.0%	0.0%	50%
DN	0.0%	0.0%	0.0%	16%
DT	0.0%	0.0%	0.0%	33%

In order to use Markov Models, it was necessary to initialize the probability matrix, what is done from the classification obtained by classification tree which distributed the individuals in the case circulations in due classes. These probabilities are calculated based on historical data. To illustrate, we have chosen the parameters rain according to the table 3, where the rain value is weak and which leads to:

$$A = \left\{a_{ij} = P(s_j \mid s_i)\right\} = \begin{bmatrix} 0.78 & 0.53 & 0.16 \\ 0.16 & 0.38 & 0.33 \\ 0.07 & 0.0 & 0.5 \end{bmatrix}$$

$$A = \left\{a_{ij} = P\left[q_i = S_j \mid q_{t-1} = S_i\right] = P(o_j \mid o_i)\right\}, \forall i \in [1, N] \text{ e } j \in [1, N].$$

And initial calculated probabilities $\Pi = \{\pi = P(s_i)\} = \begin{bmatrix} 0.69 \\ 0.21 \\ 0.1 \end{bmatrix}$.

Thus, we can predict calculating the probabilities (when it is raining lightly) as following:

1. Probability of a sequence to be realized

- the probability of obtaining the sequence P(On time, On time, On time, On time) = P(N, N, N, N) or P(NNNN) = $0.69 * 0.78 * 0.78 * 0.78 = 0.32744$
- the probability of obtaining the sequence P(In advance, On time, On time, Delayed) = P(D, N, N , T) or P(DNNT) = $0.1 * 0.16 * 0.78 * 0.16 = 0.0199$

2. Prediction of a future state The event system consist of: $P(X(t) = o_i) = \sum_{j=1}^{m} \left[P(X(t) = o_i \mid X(t-1) = o_j) * P(X(t-1) = o_j\right]$. On the assumption of stationarity, that is, a change in a state independent of time, this system becomes: $P(X(t)) = o_i \mid X(t-1) = o_j) = P(o_i \mid o_j) = a_{ij}$. That said, we can calculate, for example, the probability of having "On time" adherence in the instant $t+3$: $P(X(3) = N) = 0.78*(0.78*0.69+0.16*0.21+0.07*0.1)+0.16*(0.53*0.69+0.38*0.21+0*0.1)+0.07*(0.16*0.69+0.33*0.21+0.5*0.1) = 0.5380$.
3. Prediction of a future state from a known state We use the formula: $P(X(t_3) = o_k \mid X(t_1) = o_k) = \sum_{j=1}^{m} P(X(t_3) = o_k \mid X(t_2) = o_j) P(X(t_2) = o_j \mid X(t_1) = o_1)$. One can calculate the probability of having adherence "On time" at instant $t+4$, knowing that the current degree of adherence is "Delayed" at time t: Analytically we get: $P(X(t_4) = \text{“}N'' \mid X_1 = \text{“}T'') = 0.78*0.78*0.53+0.53*0.38*0.38+0.16*0.5*0 = 0.40$. As at instant $t+3$: $P(X(t_3) = \text{“}N'' \mid X_1 = \text{“}T'') = 0.78*0.53+0.53*0.38+0.16*0 = 0.61$.

6 Validation

To evaluate our approach, we used the Prediction performance index (PP) whose aim is to examine the rate of good prediction of the evolution of sequences $S_I = (g_{i,1}, g_{i,2}, \ldots, g_{i,T_i})$, where $(g_{i,j} = (GHA_{i,j}, VM_{i,j}))$ is a set of sequential data Y. This evaluation process consist of:

- delete the last state g_{i,T_i} of the sequence S_i;
- classify a new sequence truncated sequence (S_{tr}) into k existent classes;
- predict the state q more likely to occur and appear in the end this truncated sequence. This state is thereby compared with the actual state g_{i,T_i} that was deleted from the sequence.

We also used the intraclass Homogeneity Index (HI) which validate the automatic classification problem of sequential data. In this work, the intraclass homogeneity is regarded as a probabilistic index reflecting the stability, cohesion and ease of interpretation of the classes obtained by an automated classification process. Where the higher the intraclass homogeneity value, the greater the possibility of partition classes are compact and easy to interpret. Thus, it was possible to confirm the effectiveness of decision-making models of the proposed classification and prediction approach. Table 4 shows the results obtained on samples of real and simulated data trajectories in term of performance **prediction** and **intraclass homogeneity**. This result clearly indicates that the proposed approach showed good performance. Where TRD, TSD, TsRD and TsSD stand for Training Real Data, Training Simulated Data, Test real Data and Test Simulated Data respectively.

Table 4 The first line indicates the prevision performance based on 217 circulations and the second the Prevision performance based on 623 circulations

HI	PP (TRD)	PP (TSD)	PP (TsRD)	PP (TsSD)
87,0%	72,6%	75,2%	69,8 %	73,1%
92,0%	74,2%	79,7%	73,8 %	75,3%

7 Conclusion

The evaluation of the experiments show the usability of the proposed approach in this paper as a decision-making support in the railway context of the type studied (rail freight transport), the ability to respond to problems of classification and prediction of trajectories or behavior of vehicles. In addition to the above, the operating sector can, based on this approach, predict future traffic behavior, which will allow good organization at operational level of traffic, management of materials and routing or route planning. Finally, the approach proposed is basically data driven and hence directly applicable to other type of networks, operated by trains and considering the parameters capable of causing disturbance.

References

1. Berger, A., Gebhardt, A., Müller-Hannemann, M., Ostrowski, M.: Stochastic delay prediction in large train networks. In: Caprara, A., Kontogiannis, S. (eds.) 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems. OpenAccess Series in Informatics (OASIcs), vol. 20, pp. 100–111. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2011). http://drops.dagstuhl.de/opus/volltexte/2011/3270
2. Carey, M., Kwieciński, A.: Stochastic approximation to the effects of headways on knock-on delays of trains. Transportation Research Part B: Methodological **28**(4), 251–267 (1994). http://www.sciencedirect.com/science/article/pii/0191261594900019
3. Carey, M., Kwieciński, A.: Properties of expected costs and performance measures in stochastic models of scheduled transport. European Journal of Operational Research **83**(1), 182–199 (1995). http://www.sciencedirect.com/science/article/pii/0377221793E0248V
4. Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M.: Dispatching and coordination in multi-area railway traffic management. Computers & Operations Research **44**, 146–160 (2014). http://www.sciencedirect.com/science/article/pii/S0305054813003249
5. Corman, F., Meng, L.: A review of online dynamic models and algorithms for railway traffic control. In: 2013 IEEE International Conference on Intelligent Rail Transportation (ICIRT), pp. 128–133, August 2013
6. Dollevoet, T., Schmidt, M., Schöbel, A.: Delay management including capacities of stations. In: Caprara, A., Kontogiannis, S. (eds.) 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems. OpenAccess Series in Informatics (OASIcs), vol. 20, pp. 88–99. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2011). http://drops.dagstuhl.de/opus/volltexte/2011/3269
7. Kecman, P., Corman, F., Meng, L.: Train delay evolution as a stochastic process. In: Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis: RailTokyo2015 (2015)

8. Keyhani, M.H., Schnee, M., Weihe, K., Zorn, H.P.: Reliability and delay distributions of train connections. In: Delling, D., Liberti, L. (eds.) 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems. OpenAccess Series in Informatics (OASIcs), vol. 25, pp. 35–46. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2012). http://drops.dagstuhl.de/opus/volltexte/2012/3701
9. Lennart, F., Müller-Hannemann, M., Schnee, M.: Efficient on-trip timetable information in the presence of delays. In: ATMOS 2008 - 8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems, Karlsruhe, Germany, September 18, 2008 (2008). http://drops.dagstuhl.de/opus/volltexte/2008/1584
10. Meester, L.E., Muns, S.: Stochastic delay propagation in railway networks and phase-type distributions. Transportation Research Part B: Methodological **41**(2), 218–230 (2007). http://www.sciencedirect.com/science/article/pii/S0191261506000221, advanced Modelling of Train Operations in Stations and Networks
11. Meng, L., Zhou, X.: Simultaneous train rerouting and rescheduling on an n-track network: A model reformulation with network-based cumulative flow variables. Transportation Research Part B: Methodological **67**, 208–234 (2014). http://www.sciencedirect.com/science/article/pii/S0191261514000782
12. Nielsen, L.K., Kroon, L., Maróti, G.: A rolling horizon approach for disruption management of railway rolling stock. European Journal of Operational Research **220**(2), 496–509 (2012). http://www.sciencedirect.com/science/article/pii/S0377221712000732

A Method for Match Key Blocking in Probabilistic Matching
(Research-in-Progress)

Pei Wang, Daniel Pullen, John R. Talburt and Cheng Chen

Abstract The pair-wise nature of Entity Resolution makes it impractical to perform on large datasets without the use of blocking. Many blocking techniques have been researched and applied to effectively reduce pair-wise comparisons in Boolean rule based systems while also providing 100 % match recall. However, these approaches do not always work when applied to probabilistic matching. This paper discusses an approach to blocking for probabilistic scoring rules through the use of match key indexing.

Keywords Entity Resolution · Scoring rule · Match key indexing · Probabilistic match · Boolean rule

1 Background

1.1 Entity Resolution

Entity Resolution (ER) is the process of determining whether two references to real world objects in an information system are referring to the same object or to different objects [4]. The references are made up of attributes and the values of the attributes describe the real world entity to which they refer. The ER processes discussed in this paper use Boolean match rules to make their decisions. Boolean match rules do not produce a score or weight when comparing a pair of references,

P. Wang(✉) · D. Pullen · J.R. Talburt
University of Arkansas at Little Rock, Little Rock, AR, USA
e-mail: {pxwang,dlpullen,jrtalburt}@ualr.edu

C. Chen
Black Oak Analytics, Little Rock, AR, USA
e-mail: cchen@blackoakgroup.com

S. Latifi (ed.), *Information Technology New Generations,*
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_73

only a True/False decision. If two references satisfy a Boolean match rule, i.e. the rule is “true”, the references are linked together. After the application of transitive closure, all of the references that can be linked together form an entity identity structure (EIS) [8]. An EIS is often labeled as a cluster in ER literature. This terminology is too general and applied to many topics outside of ER. EIS, a term from Entity Identity Information Management (EIIM), is a specific term that only applies to ER and satisfies the needs of EIIM.

The ER processes discussed and undertaken throughout this paper are based on the Fellegi-Sunter model [9]. Both Boolean match rules and scoring rules will be discussed in this research [5]. Boolean match rules and scoring rules are used to determine the outcome as "link" pairs, "non-link" pairs, or “possible-link” pairs. The latter of the three is not available in Boolean match rules. The scoring rule is a probabilistic method for performing ER. If the values meet or exceed the prescribed level of similarity, the match for that attribute is assigned a numeric score called an “agreement weight”. If the values for that attribute fail to meet the similarity condition, they are assigned a different numeric score called a “disagreement weight”. The overall score f or the match between two records is simply the total of all the agreement/disagreement scores as determined by the similarity function [5] across all attributes being compared. In evaluating the outcome of an ER process, the results of matches between all pairs of references can be categorized into four outcomes: true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). TPs are correctly labeled "link" pairs. TNs are correctly labeled “non-link” pairs. In contrast to these correct results, there are two types of incorrect linking results. FPs are pairs of records identified as matches or “link” pairs by the ER process but actually refer to two different real world entities. FNs are pairs of records identified as non-matches or “non-link” pairs by an ER process but actually refer to the same real world entity [5]. The goal of an ER process is to produce the lowest number of FPs and FNs.

2 Problem Statement

Blocking is often used to speed up the processing of large data sets and is a mandatory step for effectively working with large data sets. For example, without blocking, one data set which contains 3,232,292 records will require on the order of $2^{3232392}$ pair-wise comparisons and require an unreasonable amount of time to process. If a blocking technique is applied, the total number of pair-wise comparison can be dramatically reduced. Blocking breaks the data up into subsets called “blocks.” Comparisons are only performed on pairs of records within each block [7]. Many different types of blocking approaches have been developed. But, this research discusses one in particular. Match key indexing, also known as standard blocking, is an inverted index. For example, if a user defines last name as match key index, the value of last name is used to separate references into blocks. This means only those records which share the same last name will be blocked together. This approach, one of the more common methods of blocking in ER, was used

for the approach outlined in this paper. Match key indexing breaks the data into blocks based on the value of a set of attributes.

When applying any blocking approach, it is important to consider alignment with the match rule regardless of the type of rule, e.g. scoring rule or Boolean rules. Without proper alignment, the final result will contain many FN outcomes that may be difficult to track down or explain by looking solely at the match logic. A blocking technique is considered to be in complete alignment with the match rule if any two records matched by the match rule are always in the same block. Figure 1 visually demonstrates both an aligned and misaligned blocking technique.

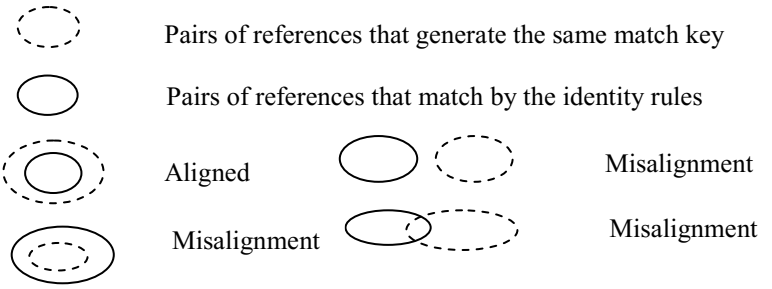


Fig. 1

Match key indexing is Boolean in nature. This means that sets of attributes can be selected that directly match the sets of attributes contained in the rules. Examples of this have been discussed in detail with examples by other researchers [11]. However, it is more difficult to create match keys for the scoring rule that are correctly aligned than for Boolean rules.

3 Match Key Index Approach for Probabilistic Matching

3.1 The OYSTER ER System

The ER processes in this paper were performed with OYSTER (Open sYStem for Entity Resolution). OYSTER is an open source ER system developed by the Center for Advance Research in Entity Resolution and Information Quality (ERIQ) at the University of Arkansas at Little Rock (UALR). OYSTER was specifically designed to support EIIM [8]. Although OYSTER can be run in several different configurations to support the various phases of the entity identity information life cycle, only the identity capture configuration was used for the results given in this paper [10].

3.2 Data and Rule Sets

The data used in this research consists of low quality synthetic data comprised of 2,866 records and 200 EIS containing 15 attributes in total. This research only utilizes nine of the total attribute set. The total true EIS and Records of data set are listed below in Table 1.

Table 1 Data Sets

	Set A
Total True EIS	200
Total Records	2,866

The data profiling of the 10 main attributes are shown in Table 2. The uniqueness percentage in Table 2 represents the number of values that have no duplicates and the distinct percentage represents the number of non-null values that are different from each other. Most of the attributes shown in Table 2 have null values. This is especially true of the DOB (date of birth) field with 75% values in this field missing. In this data set, the unique and distinct percentages of each attribute are not very high. Most of them are under 50% which shows the quality of this data set is not very good. Likely, this is caused by variations in the data formats and contents. Furthermore, some of the main attributes have several different value patterns, such as DOB and phone number. The value patterns of DOB and phone number are shown in Table 3. In addition to these identified quality problems, approximately 300 rows suffer from delimiter alignment issues. Together, these data quality problems demonstrate this data set has a relatively low degree of data quality.

Table 2 Data Profile

Attrib-utes	First-Name	Middle-Name	Last-Name	Address	City	State	ZIP	SSN	DOB	Phone
Null	4%	28%	0%	28%	29%	28%	28%	52%	75%	32%
Unique	10%	0%	10%	53%	20%	1%	8%	17%	8%	37%
Distinct	28%	2%	28%	61%	34%	1%	21%	27%	14%	50%

Table 3 Value Patterns for Data of Birth and Phone

DOB	Percentage	PhoneNumber	Percentage
NULL	75.47%	NULL	31.79%
DDDDDDDD	7.43%	DDD.DDD-DDDD	11.30%
DDDD	7.08%	(DDD)DDD.DDDD	11.17%
D/DD/DDDD	2.72%	DDD-DDD.DDDD	11.13%
DDD-DD-DDDD	2.48%	(DDD)DDD-DDDD	10.99%
DD-LLL-DD	1.64%	DDD-DDD-DDDD	10.96%
D/D/DDDD	1.05%	DDD.DDD.DDDD	9.53%
D-LLL-DD	0.73%	DDDD	1.33%
DD/DD/DDDD	0.63%	DDDDDDDD	0.63%
DD/D/DDDD	0.38%	D/DD/DDDD	0.35%
DDDDDDDDD	0.21%	DD-LLL-DD	0.24%
LL DDDDD"	0.07%	DD/DD/DDDD	0.21%
DDDDDDD	0.03%	D/D/DDDD	0.17%
LL DDDD"	0.03%	D-LLL-DD	0.14%
LLLLL. DDDDD"	0.03%	DD/D/DDDD	0.07%

3.3 Scoring Rule

A scoring rule is defined by a set of attributes each with a corresponding agreement and disagreement weights. The agreement weight refers to the weight applied when an agreement occurs between corresponding attribute values from two references being compared, as specified by the similarity function. The disagreement weight refers to the weight applied when a disagreement occurs between corresponding attribute values from two references being compared, as specified by the similarity function.

3.4 Weight Calculation

The algorithm used to calculate the attribute agreement weights and disagreement weights is the Fellegi-Sunter probabilistic model for estimated weights [2] under the assumption of conditional independence of the identity attributes [6]. This research does not apply frequency based weighting. A previous Boolean result is used to approximate the truth to determine equivalence during the weight calculation. In order to calculate the probability that the equivalent records agree on attribute (m_i) and the probability of the non-equivalent records agree on attribute (u_i), the equivalent pairs (E_i) and non-equivalent pairs that agreed on attribute i ($!E_i$) must first be counted. Secondly, the numbers of equivalent pairs (E_h) and non-equivalent pairs of the whole data set ($!E_h$) need to be counted. The calculation of the probability of equivalent records and non-equivalent records are shown below in equations (1) and (2).

$$m_i = E_i / E_h \tag{1}$$

$$u_i = !E_i / !E_h \tag{2}$$

The agreement weight and disagreement weight of individual attributes are calculated as:

$$w_i = \begin{cases} \log_2\left(\frac{m_i}{u_i}\right) & \text{if agreement in attribute i} \\ \log_2\left(\frac{(1-m_i)}{(1-u_i)}\right) & \text{otherwise} \end{cases} \tag{3}$$

In this research, a Java program was written to automatically calculate the agreement and disagreement weights from the synthetic data. The attributes weights are shown in Table 4.

Table 4 Scoring Rule Weight

	First-Name	Middle-Name	Last-Name	DOB	Phone	Address	City	State	Zip
Rule	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact
Agree Weight	76	29	77	0	13	10	15	6	16
Disagree Weight	-17	-34	-18	0	-3	-1	-3	-3	-3

4 Experimental Results

4.1 Scoring Rule Match Key Index

When using the scoring rule, the combinations of attributes are called patterns. The goal is to produce an index configuration that covers all the patterns above an established threshold. In order to achieve this, the patterns with weights of all possible combinations should been generated first. An attribute set of 9 generates 512 patterns. Because this is a binary pattern, 0 for disagreement and 1 for agreement, the total number of patterns is calculated as 2^9. The example of the combination patterns are listed in Table 5 and the distribution of the patterns are listed in Figure 2. The sequences of attributes inside of the pattern are followed by the cumulative weights calculated from the values listed in Table 4. In a pattern, one (1) stands for two references that match on the attribute and will be assigned the agreement weight. Zero (0) means that two references do not match on this attributes and will be assigned the disagreement weight.

Table 5 Example of Patterns

Pattern	AgreeWeight	DisAgreeWeight	TotalWeight
111011111	242	0	242
111111111	242	0	242
111011101	236	-3	233
111111101	236	-3	233
111010111	232	-1	231
111110111	232	-1	231

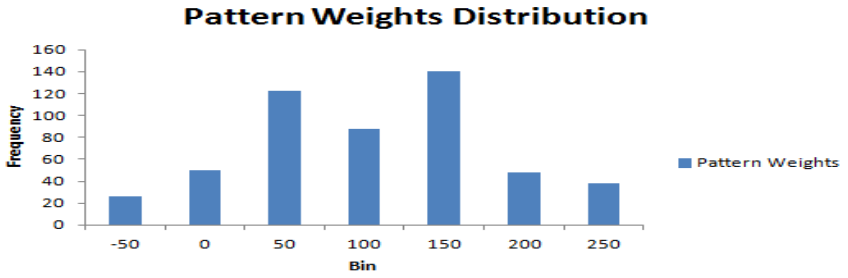


Fig. 2

A match threshold needs be identified after generating all the patterns. Based on this threshold, all of the patterns above will be considered a match. This research chose 25 as the threshold. There are 396 patterns above this threshold. The attribute pattern is one type of binary pattern. It has the feature that one pattern with fewer attribute agreement scan cover several patterns that have more attributes agreements. This is much like aligning Boolean rules. For example, the pattern 111000000 can cover all patterns that agree on first 3 attributes regardless of the agreement or disagreements on the remaining attributes. The patterns which have fewer attribute agreements are referred to as parent pattern. The patterns which have more attribute agreements are called children patterns.

To find appropriate parent patterns from all the patterns above threshold 25 is the main step in selecting indices for probabilistic matching. Two factors used to identify the appropriate pattern include: (1) The number of attributes in the pattern (2) The number of records for each candidate group.

As mentioned before, the less attributes agree in parent pattern, the more children patterns can be covered. At the same time, the size of each candidate group needs to be taken into consideration. The group size will heavily affect the number of pair-wise comparisons and performance of OYSTER. In order to keep a balance of these two factors, the first generation of the parent patterns start with pair patterns, which have two attributes in agreement in the pattern. To find out all of the pair patterns above the threshold, remove the pattern itself and all of the children patterns of this parent pattern from the pattern pool. After the all the pair

patterns have been assessed, the triple patterns are the next generation of the parent pattern. The iterative approach of the parent pattern generation continues and keeps identifying the parent patterns until the all of the patterns that need to be covered are removed from the pool. For example, if there are four attributes, pair pattern 1100 is one of the parent patterns for 1101, 1110 and 1111. A tree example is shown in Figure 3.

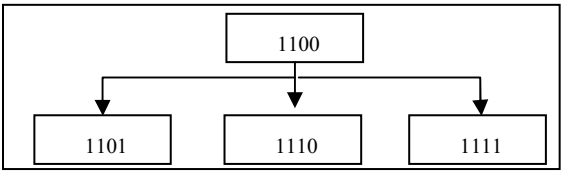


Fig. 3

A second Java program was written to automatically identify the maximum candidate groups for match keys of each index pattern and its coverage of the match patterns.

There are 17 patterns out of 396 patterns that satisfy the criteria. All of them were used in one scoring run. The attributes and indices are listed in Table 6.

Table 6 Scoring Rule Index

	First-Name	Middle-Name	Last-Name	DOB	Phone	Address	City	State	Zip
Index 1	Exact		Exact						
Index 2		Exact	Exact						
Index 3	Exact	Exact							
Index 4			Exact						Exact
Index 5			Exact				Exact		
Index 6	Exact								Exact
Index 7			Exact		Exact				
Index 8	Exact						Exact		
Index 9	Exact				Exact				
Index 10			Exact			Exact		Exact	
Index 11	Exact					Exact		Exact	
Index 12		Exact			Exact		Exact		Exact
Index 13		Exact				Exact	Exact		Exact
Index 14		Exact					Exact	Exact	Exact
Index 15		Exact			Exact	Exact			Exact
Index 16		Exact			Exact	Exact	Exact		
Index 17		Exact			Exact			Exact	Exact

By comparing the results of the scoring rule with threshold 25 using indices listed in Table 6 against the truth, the TWi was calculated as 0.68731. By comparing results of the scoring rule with and without indices using the same threshold, the TWi was calculated as 1.0. This means the indices are in perfect alignment with the scoring rule.

4.2 Using Scoring Rule to Define Boolean Rule Sets

It is hard to define the appropriate Boolean rule without knowing the data well. For the scoring rule, the user does not need to choose the combination of attributes when designing the rule. Instead, the scoring rule can automatically handle all of the combinations of attributes by comparing them against the match threshold during run time. However, it does require some type of equivalence link for attribute weight calculation. With the scoring rule, the user can use results from simple Boolean rules to calculate the weights for the scoring rule. The weights are good enough to produce effective and accurate scoring rule results [3]. The approach for defining the Boolean rule using scoring rules is similar to the approach to finding the indices of a scoring rule. This Boolean approximation of a scoring rule has the same goal to produce the same results as the scoring rule without index.

Table 7 Boolean Rule

	First-Name	Middle-Name	Last-Name	DOB	Phone	Address	City	State	Zip
Rule 1	Exact		Exact						
Rule 2		Exact	Exact						
Rule 3	Exact	Exact							
Rule 4			Exact						Exact
Rule 5			Exact				Exact		
Rule 6	Exact								Exact
Rule 7			Exact		Exact				
Rule 8	Exact						Exact		
Rule 9	Exact				Exact				
Rule 10			Exact			Exact		Exact	
Rule 11	Exact					Exact		Exact	
Rule 12		Exact			Exact		Exact		Exact
Rule 13		Exact				Exact	Exact		Exact
Rule 14		Exact					Exact	Exact	Exact
Rule 15		Exact			Exact	Exact			Exact
Rule 16		Exact			Exact	Exact	Exact		

The indices generated to align with the scoring rule are the minimum set of parent patterns covering all of the children patterns above the identified threshold. These parent patterns can not only be treated as indices but also as Boolean rules to get the same results of an identified threshold for scoring rule. There are two reasons why it should produce the same results.

1) If the parents pattern is above the threshold, it's children pattern should also above the threshold
2) If the records match by a children pattern, it should also be matched together by the parent pattern.

Additionally, the result of parent patterns is the transitive closure (TC) of all the children pattern results. For the transitive closure, it means to group any records together that contain a link to records in another group [7]. The rules are listed in Table 7. By comparing the results of using the 16 Boolean rules and the scoring rule with threshold 25, the TWi was calculated as 1.0. This means the two produce identical results.

5 Conclusion

The calculation of the TWi from scoring rule, with and without indices, demonstrates that both results are the same and shows that the indices can be aligned with the scoring rule for a given match threshold. In order to align the index configuration to the probabilistic match rule, the goal is to cover all the match patterns that generate scores above the threshold. This method finds the parent patterns above the threshold to cover the children pattern. A total of 396 patterns are above the threshold, but only 17 of them are parent patterns, which cover the remaining 379 patterns. Furthermore, the parent patterns used as indices can also be used as Boolean rules to produce the same results as the scoring rule for a given threshold without using value frequency-based matching. The comparison of the 17 Boolean rules with the scoring rule using a threshold 25 demonstrates that the two approaches produce the same results.

6 Future Work

Frequency-based matching that uses separate weights specific to a particular attribute value should be explored further. Some names may have a much stronger probability of indicating a TP match. For example, “Pei” is an uncommon name in the United States of America and may indicate a stronger probability of a TP outcome in data from the United States of America. Similarly, “Daniel” is an uncommon name in China and may indicate a stronger probability of a TP outcome in data from China.

Furthermore, the size of blocks from multiple match-key index generators after transitive-closure can be difficult to predict. Though a particular blocking

configuration may be in alignment with the match logic, it may still be unfeasible to perform all of the pair-wise comparisons within a very large block. Heuristic logic should be explored to see if it is possible to dynamically adjust the configuration of a match key blocking technique in order to reduce the total pair-wise comparisons without sacrificing the accuracy of the results.

References

1. Pullen, D., Wang, P., Wu, N., Talburt, J.R.: Mitigating data quality impairment on entity resolution errors in student enrollment data. In: Proceedings: Information and Knowledge Engineering Conference, 2013, pp. 96–100 (2013)
2. Winkler, W.: Using the EM Algorithm for Weight Computation in the Felligi-Sunter Model of Record Linkage. Report No. RR2000/05, Statistical Research Division, Methodology and Standards Directorate, U.S. Bureau of the Census, Washington, DC (2000)
3. Wang, P., Pullen, D., Wu, N., Talburt, J.R.: Iterative approach to weight calculation in probabilistic entity resolution. In: Proceedings: International Conference on Information Quality (ICIQ-19), Xi'an, China, August 1–3, 2014, pp. 245–258 (2014)
4. Talburt, J.R.: Entity Resolution and Information Quality. Morgan Kaufmann/Elsevier, San Francisco (2011)
5. Christen, P.: Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Springer, Berlin (2013)
6. Ihara, S.: Entropy. Information Theory for Continuous Systems, pp. 1–2. World Scientific, Singapore (1993)
7. Talburt, J.R., Zhou, Y.: Entity Information Life Cycle for Big Data: Master Data Management and Information Integration. Morgan Kaufmann (2015)
8. Zhou, Y., Talburt, J.R.: Entity Identity Information Management (EIIM). In: Proceedings: International Conference on Information Quality (ICIQ-11), Adelaide, Australia, November 18–20, 2011, pp. 327–341 (2011)
9. Fellegi, I., Sunter, A.: A Theory for Record Linkage. Journal of the American Statistical Association **64**(328), 1183–1210 (1969)
10. Zhou, Y., Talburt, J.R., Kobayashi, F., Nelson, E.D.: Implementing Boolean matching rules in an entity resolution system using XML scripts. In: Proceedings: Information and Knowledge Engineering Conference, 2012, pp. 332–337 (2012)
11. Zhou, Y., Nelson, E., Talburt, J.R.: User-defined inverted index in boolean, rule-based entity resolution systems. In: Proceedings: International Conference on Information Technology: New Generations, Las Vegas, Nevada, April, 2013, pp. 608–612 (2013)

Accuracy Assessment on Prediction Models for Fetal Weight Based on Maternal Fundal Height

Applications in Indonesia

Dewi Anggraini, Mali Abdollahian and Kaye Marion

Abstract Delivery weight is a significant indicator of pregnancy outcomes. Both low birth weight and macrosomia ranges have negative impact on neonatal health yet low birth weight is the leading cause of neonatal mortality in most developing countries. Since fetal weight cannot be directly measured, the prediction has become increasingly important in routine antenatal care. Early birth weight estimation during expectancy assists medical practitioners to make an informed decision on whether intervention is required prior to delivery. Several prediction models for fetal weight have been developed in Indonesia based on clinical assessment of fundal height as an alternative to ultrasound. However, most prediction models only would provide weight estimates very close to delivery or at more than 35 weeks of gestational age. Some researchers carried out comparison study among these prediction models. However, there has been little discussion on their forecast accuracy measures. This paper aims to evaluate and compare the accuracy of existing prediction–based fundal height models for estimating neonatal delivery weight using fundal height measurements between 20 and 35 weeks of gestational age.

Keywords Forecast accuracy · Fetal weight · Fundal height · The Johnson-Toshach's model · The Risanto's model

1 Introduction

Birth weight is one of the most important indicators for safe pregnancy outcomes. A normal delivery weight is one of the most expected pregnancy outcomes [1]. Deviation from normal delivery weight (2500 – 3999 g), such as low birth weight (< 2500 g) and macrosomia (> 4000 g) leads to some negative consequences on

D. Anggraini(✉) · M. Abdollahian · K. Marion
School of Mathematical and Geospatial Sciences,
Royal Melbourne Institute of Technology (RMIT) University, Melbourne, Australia
e-mail: {dewi.anggraini,mali.abdollahian,kaye.marion}@rmit.edu.au

S. Latifi (ed.), *Information Technology New Generations,*
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_74

neonatal health [1-3]. Since fetal weight cannot be measured directly during pregnancy, its estimation from fetal or maternal anatomical characteristics [4] has increasingly become one of the most important measurements in maternal and neonatal health care centers, particularly in routine antenatal care. A reliable delivery weight prediction between 20 and 35 weeks would be critical for intervention for safe delivery [2].

Simple to advanced clinical approaches together with scientific modeling techniques have been developed to predict fetal weight prior to delivery. In most developed countries, the use of radio frequency volume reduction (ultrasound) [5-8] is common to monitor fetal intergrowth. Unfortunately, this is not the case in most developing countries where the provision of health resources and trainings is limited for ultrasound. Therefore, abdominal palpation method [1, 9] and fundal height measurement [1, 3, 9-15] remain the most applicable clinical alternatives to assess fetal intra uterine growth. The latter method is more objective than the former approach [16].

In recent years, there has been an increasing interest in the accuracy of fetal weight prediction models. It is essential to record the fetal weight on a continuing basis. This will assist clinicians to plan appropriate care for the safety of both mother and fetus during the process of pregnancy and delivery [3]. A considerable amount of literature has been published on the comparison between predicted delivery weights based on clinical and ultrasonic measurements [4, 16-19]. Most of these studies conclude that clinical method was as accurate as ultrasonic method in estimating fetal weight, particularly for the normal newborns but not for the low birth weight and macrosomia neonates. Therefore, clinical approaches can be used as alternatives to ultrasound for estimating fetal weight in low income countries.

In Indonesia, the Johnson-Toshach's model is well-recognized and deployed as a standard clinical model to estimate fetal weight. This model was first developed in 1954 by Johnson and Toshach [10] and later on was simplified in 1957 [20]. The development criteria were based on the measurement of fundal height and the knowledge of fetal engagement status on ischial spine among the US data. The measurements were taken prior to delivery or at more than 35 weeks of gestational age.

The other commonly used prediction model for delivery weight in Indonesia is the Risanto's model. This model was first established in 1995 by Siswosudarmo [21] cited in [15]. It was redeveloped in 2014 [22]. The development criteria were similar to the former model minus the fetal station information. The original model was developed using a sample of size 560 pregnant women in the Province of Yogyakarta in Indonesia. It was then improved using a sample of size 655 from the same province.

Several attempts have been made to compare these models [15, 23]. However, the majority of these studies only included term pregnancies (> 35 weeks of gestational age) and little attention has been paid to the efficacy assessment of these models and how well they predict neonatal weight during antenatal care or between 20 and 35 weeks of gestational age. This paper for the first time will focus on comparing the accuracy of the above models in predicting fetal weight between 20 and 35 weeks of gestational age in Indonesia.

2 Method

This study has been carried out in two steps: a systematic literature review and a case study using a secondary data set.

2.1 A Systematic Literature Review

A systematic review was conducted to compare the Johnson-Toshach's and the Risanto's prediction models. The models were selected because they are commonly used in Indonesia and are developed based on fundal height measurement taken prior to delivery. Note that the Johnson-Toshach's model is developed using the US data while the Risanto's model is based on Indonesian data.

2.2 A Case Study and Data Analysis

In order to compare the accuracy of the models, a secondary data set of pregnant women from a midwife clinic in South Kalimantan, Banjarmasin, Indonesia is used. The data contains the fundal height measurements, the information of fetal head station during expectancy or between 20 and 35 weeks of gestation age, and the respective actual birth weights. We will use multi assessment criteria to assess the accuracy. The accuracy of three forecasting models (the Johnson-Toshach's and the two Risanto's models) are listed in Table 2-4. Data management and analysis were performed using Microsoft Excel and Minitab 17.

3 Existing Prediction Models of Fetal Weight

The Johnson-Toshach's and the Risanto's formulas are the two most commonly used models deployed in Indonesia to predict the fetal weight.

3.1 The Johnson-Toshach's Formula

The basic assumption of the Johnson-Toshach's model was that the height of the fundus should be 35 cm when the fetal head is at station zero or engaged for normal fetal weight (7 pounds and 8 ounces or approximately 3400 grams) [10].

Their proposed model is:

$$\text{EFW (pounds)} = [7 \text{ pounds} + 8 \text{ ounces}] + [(M + S - 35 \text{ cm}) (5 \text{ ounces})] \quad (1)$$

Where EFW is the estimated fetal weight in pounds, M is the height of fundus measured in centimeters and S is the fetal station and its value is defined by S = -1 for "minus" station, S = 0 for "zero" station, and S = 1 for "plus" station.

Johnson in 1957 simplified the model and changed the unit into grams [20]. The simplified model is:

$$\text{EFW (grams)} = (\text{FH} - n) \times 155 \quad (2)$$

Where EFW is the estimated fetal weight in grams, FH is the fundal height measured in centimeters. The value of n = 13 if vertex is above ischial spine or fetal head is unengaged, n = 12 if vertex is at ischial spine or fetal head is engaged, and n = 11 if vertex is below ischial spine or fetal head is engaged [24]. Equation (2) also takes into account maternal weight adjustment, i.e., if a pregnant woman weighs more than 91 kg, 1 cm is subtracted from the fundal height [4].

Original results based on (1) showed that the average prediction error was 12.45 ounces [10, 20]. The study also showed that the fundal height of 26 cm can potentially indicate a premature birth weight (this was true for 80% of the premature newborns). The fundal height of 30 cm is necessary to ensure that the baby would have normal weight.

3.2 The Risanto's Models

The Risanto's model (3) proposed by Siswosudarmo in 1995 [21]. A modified version of (3) was proposed in 2014 (4) [22]. Both models were developed using Indonesian data where measurements were taken prior to delivery or between 37 and 42 weeks of gestational age. The models are:

$$\text{EFW (grams)} = 126.7\ \text{FH} - 931.5 \quad (3)$$

$$\text{EFW (grams)} = 125\ \text{FH} - 880 \quad (4)$$

Deploying both models, the mean difference between (3) and the actual birth weight was 100.8 ± 86.1 grams. The results (4) also showed that correlation between the measurement of fundal height and neonatal weight is r = 0.93 and 86% of variation in the delivery weight is explained by the maternal fundal height ($R^2 = 0.86$).

4 Accuracy Assessment

The accuracy of prediction is fundamental to ensure that the forecast correctly reflects the future [26-28]. In this paper, we have used the following error measurements to assess the efficacy of the different models; mean error (ME), median error (MEDE), mean percentage error (MPE), median percentage error (MEDPE), mean absolute error (MAE), median absolute error (MEDAE), mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), median absolute percentage error (MEDAPE), mean square percentage error (MSPE), and root mean square percentage error (RMSPE) [25, 26].

In the scale-dependent category, bias can be measured by ME and MEDE, while MPE and MEDPE are used in the non-scale category. Both ME and MPE represent the average of prediction errors but in different terms. ME is expressed

in the units of observations and written as $ME = \frac{\sum(A_i - P_i)}{n}$, whereas MPE is stated in the percentage and written as $MPE = \frac{\sum\left(\frac{A_i - P_i}{A_i}\right) \times 100}{n}$, where A_i is the actual birth weights, P_i is the predicted values and n is the number of observations.

These two measures can determine bias by identifying the number of underestimated and overestimated forecasts but they are very sensitive to extreme errors. Therefore, to anticipate false interpretation due to error distortion, it is recommended [25] to use the $MEDE = Median\,(A_i - P_i)$ or $MEDPE = Median\left(\frac{A_i - P_i}{A_i} \times 100\right)$.

On the other hand, the precision, in the scale-dependent category, can be assessed by MAE, MEDAE, MSE and RMSE, while in the non-scale category it can be explained by MAPE, MSPE, RMSPE, and MEDAPE. The MAPE measure is commonly used in cross-sectional study to compare different prediction models over a period of forecasts interval even though it tends to have larger values than MEDAPE [26].

5 Comparison between Fetal Weight Prediction Models Based on Fundal Height Measurement

Sequence plot is used to visually compare the delivery weight predictions based on the Johnson-Toshach's and the Risanto's models with the actual recorded birth weights (see Fig. 1). The prediction models used 39 recorded data. The fundal height was measured between 20 and 35 weeks of gestational age.

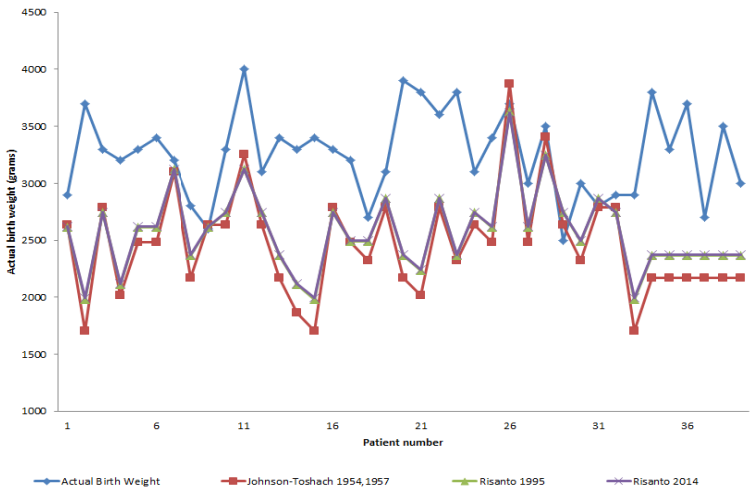


Fig. 1 Sequence plot of fetal weight estimated using the Johnson-Toshach's and the Risanto's models where fundal height is measured between 20 and 35 weeks of gestational age

Fig. 1 shows that both models underestimate the delivery weights and there is no significant difference between their forecasting abilities. The authors have also carried out hypothesis testing to assess if there is a significance difference between the actual recorded data and forecasted values based on the three models.

Table 1 presents the corresponding p-values and the 95% confidence intervals for the differences between the actual and estimated delivery weights. The results show that there is a statistically significant difference between the actual delivery weights and the forecasts based on the existing models, when forecasts are based on fundal height measured between 20 and 35 weeks.

Table 1 Mean difference between the Johnson-Toshach's and the Risanto's models and actual birth weight

Models	Estimated mean difference	P-value	Pooled standard deviation	95% confidence interval
(2)	794.9	≤ 0.0005	422.04	(604.5, 985.2)
(3)	691.6	≤ 0.0005	366.02	(526.5, 856.7)
(4)	687.1	≤ 0.0005	363.75	(523.0, 851.1)

The performance of the three models has been presented in Table 2. We can conclude based on the p-values (> 0.05) that there is no significant difference between the Johnson-Toshach's and the two Risanto's models. We can also conclude that the two Risanto's models are not significantly different.

Table 2 Mean difference between the Johnson-Toshach's and the Risanto's formulas

Models	Estimated mean difference	P-value	Pooled standard deviation	95% confidence interval
(2) – (3)	-103.3	0.269	409.91	(-288.1, 81.6)
(2) – (4)	-107.8	0.247	407.89	(-291.8, 76.1)
(3) – (4)	-4.6	0.954	349.60	(-162.2, 153.1)

To assess forecast accuracy between the Johnson-Toshach's and the Risanto's models, we have measured the bias and precision using a sample of size 39. The results are given in Table 3 and Table 4.

Table 3 Measures of bias for 39 observations during expectancy at gestational age between 20 and 35 weeks

Models	Measures of bias (direction problem)					
	Scale dependent		Non-scale dependent		Direction	Bias (%)
	ME	MEDE	MPE	MEDPE		
(2)	794.87	720	23.41	22.50	Underestimate	92
(3)	691.61	637	20.18	20.29	Underestimate	92
(4)	687.05	630	20.04	20.28	Underestimate	92

It is apparent from this table that all the proposed models have an equal pattern of bias. More than 90% of the forecasts underestimate the birth weight. Both Risanto's models perform slightly better than the Johnson-Toshach's formula (smaller ME, MEDE, MPE and MEDPE).

Table 4 Measures of precision for 39 observation during expectancy at gestational age between 20 and 35 weeks

Models	Measures of precision (magnitude problem)							
	Scale dependent				Non-scale dependent			
	MAE	MED AE	MSE	RMSE	MAPE	MSPE	RMS PE	MED APE
(2)	813	720	965123	982	24	801	28	20
(3)	708	637	731686	855	21	597	24	17
(4)	704	630	722294	850	21	589	24	17

Table 4 confirms that the Johnson-Toshach's model tended to produce more error than the other two models. Table 5 presents the percentage of estimated weight that fall outside the 2 standard deviation from mean of the actual data.

Table 5 Percentage error of individual estimate at the 2 standard deviation of the mean actual birth weights

Models	Complete samples	$\bar{x}$ (grams)	s (grams)	$\bar{x} \pm 2s$ (grams)	Number of estimates outside the range	
					n	%
(2)	39	3267.7	378.1	(2511.5 – 4023.9)	23	58.97
(3)	39				19	48.72
(4)	39				19	48.72

The results indicate that 59% of the estimated delivery weights using the Johnson-Toshach's model fall below the 2 standard deviation of the mean actual birth weights in comparison with 49% based on the two Risanto's models. The 2 standard deviation range is recommended [27] for standard practice or "alarm signals" in health care.

6 Discussion

A strong relationship between maternal fundal height (measured after 35 weeks of gestational age) and birth weight has been reported in the literatures [11-13, 22, 28-29]. The analysis of the recorded data in this paper also confirms this relationship with the correlation coefficient of r = 0.93 when using fundal height measurement recorded after 35 weeks. In Indonesia, the Johnson-Toshach's model (developed based on the US data) has been used nationally to estimate delivery weight versus the less commonly used the Risanto's models (developed based on Indonesian data). Currently, these models are used after 35 weeks of gestation age to predict delivery weight.

In this paper, we deployed the models to predict the delivery weight based on fundal height measurements recorded between 20 and 35 weeks of gestation age. The comparison demonstrates that the Risanto's models perform slightly better than the Johnson-Toshach's model. The results also indicate that using early

measurements of fundal height, 49% of the Risanto's delivery weight predictions fall below the 2 standard deviation of the actual recorded delivery weights versus 59% for the Johnson-Toshach's model.

While both models are based on fundal height, the Johnson-Toshach's model also requires the knowledge of fetal head position (whether it is below, at, or above ischial spine). A subjectivity issue may emerge from the decision on fetal head engagement unless there is a standard protocol to determine fetal station with less error, yet the results show that inclusion of this measurement has not reduced the prediction error in delivery weight estimation [13]. Therefore, one should question its clinical usefulness.

7 Conclusion

In recent years, there has been increasing interest in predicting birth weight based on maternal fundal height prior to delivery or at full term pregnancy. However, early delivery weight estimation can assist the medical practitioners to decide whether early clinical intervention for safe delivery is required. It can also assist the local government to make informed medical resource allocation decisions particularly in rural area.

The most commonly used prediction models in Indonesia are the Johnson-Toshach's and the Risanto's models. The current predictions are based on fundal height measured after 35 weeks of gestation age and predict the weight very close to delivery. This paper, for the first time is using these models to predict the delivery weight based on fundal height measured between 20 and 35 weeks of gestation age.

The paper has deployed the three commonly used models to predict the delivery weight using real data form a rural clinic in Indonesia. The predicted delivery weight for individual mother is then compared with the actual delivery weight to assess the efficacy of these models. The accuracy measures together with the percentage of estimates outside the 2 standard deviation range are provided. The results show that the Risanto's models slightly outperform the Johnson-Toshach's model and yet are simpler to deploy as they only require fundal height measurements in contrast to the Johnson-Toshach's model that requires the knowledge of fetal head position as well.

One of the significant outcomes that is hoped to emerge from this research is that an early delivery weight prediction will assist the medical practitioners to plan appropriate clinical interventions well in advance if it is required. The Risanto's models may predict better prior to delivery than the Johnson-Toshach's formula but both are incapable to accurately predict fetal weight based on the fundal height measurements recorded between 20 and 35 weeks. One way to improve the accuracy is to investigate the impact of incorporating other characteristics of mother or fetal in the models. The other alternative would be to examine whether one can define a suitable correction factor in both formula. Both of these alternatives are part of future research in this study.

Author Contributions. D.A., M.A., and K.M. contributed to the study concept, design, analysis and interpretation. D.A. collected and summarized information on relevant research articles and policies in Indonesia and drafted this article.

Acknowledgment This paper presents part of doctoral degree research that is funded by the Australia Awards Scholarships (AAS). The authors would like to thank Supri Nuryani, SSiT, MKes, as a senior midwife and her dedicated student, Rahma An Nisa, Amd. Keb., for their effort in collecting the data from the clinic in Banjarmasin, South Kalimantan, Indonesia, and providing them to us for our research.

References

1. Farid and Sukarya, W.: Birth weight estimation based on the modified Niswanser formula—Taksasi berat badan anak berdasarkan modifikasi rumus Niswander. Majalah Obstetri dan Ginekologi Indonesia **23**, 188–193 (1999)
2. Belete, W., Gaym, A.: Clinical estimation of fetal weight in low resource settings: comparison of Johnson's formula and the palpation method. Ethiopian Medical Journal **46**, 37–46 (2008)
3. Parvin, Z., Shafiuddin, S., Uddin, M.A., Begum, F.: Symphysio Fundal Height (SFH) Measurement as a Predictor of Birth Weight. Faridpur Medical College Journal **7**, 54–58 (2013)
4. Khani, S., Ahmad-Shirvani, M., Mohseni-Bandpei, M.A., Mohammadpour-Tahmtan, R.A.: Comparison of abdominal palpation, Johnson's technique and ultrasound in the estimation of fetal weight in Northern Iran. Midwifery **27**, 99–103 (2011)
5. Hadlock, F., Harrist, R., Carpenter, R., Deter, R., Park, S.: Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. Radiology **150**, 535–540 (1984)
6. Dudley, N.: A systematic review of the ultrasound estimation of fetal weight. Ultrasound in Obstetrics & Gynecology **25**, 80–89 (2005)
7. Campbell, S., Wilkin, D.: Ultrasonic measurement of fetal abdomen circumference in the estimation of fetal weight. BJOG: An International Journal of Obstetrics & Gynaecology **82**, 689–697 (1975)
8. Spinnato, J.A., Allen, R.D., Mendenhall, H.W.: Birth weight prediction from remote ultrasound examination. Obstetrics & Gynecology **71**, 893–898 (1988)
9. Niswander, K.R., Capraro, V.J., Van Coevering, R.J.: Estimation of birth weight by quantified external uterine measurements. Obstetrics & Gynecology **36**, 294–298 (1970)
10. Johnson, R., Toshach, C.: Estimation of fetal weight using longitudinal mensuration. American Journal of Obstetrics and Gynecology **68**, 891 (1954)
11. Mongelli, M., Gardosi, J.: Estimation of fetal weight by symphysis–fundus height measurement. International Journal of Gynecology and Obstetrics **85**, 50–51 (2004)
12. Buchmann, E., Tlale, K.: A simple clinical formula for predicting fetal weight in labour at term: derivation and validation. SAMJ South African Medical Journal **99**, 457–460 (2009)
13. Bothner, B., Gulmezoglu, A., Hofmeyr, G.: Symphysis fundus height measurements during labour: a prospective, descriptive study. African Journal of Reproductive Health **4**, 48–55 (2000)

14. Gayatri, D., Afiyanti, Y.: Validation of fetal weight estimate formula to predict birth weight based on maternal fundal height—Validasi rumus taksiran berat janin (TBJ) untuk prediksi berat badan lahir berdasarkan tinggi fundus uterus ibu hamil. Jurnal Keperawatan Indonesia **10** (2006)
15. Titisari, H.I., Siswosudarmo, R.: Risanto's formulas is more accurate in determining estimated fetal weight based on maternal fundal height. Indonesian Journal of Obstetrics and Gynecology **1**, 149–151 (2013)
16. Banerjee, K., Mittal, S., Kumar, S.: Clinical vs. ultrasound evaluation of fetal weight. International Journal of Gynecology & Obstetrics **86**, 41–43 (2004)
17. Shittu, A., Kuti, O., Orji, E.: Comparison of clinical and ultrasonographic estimation of fetal weight. International Journal of Gynecology and Obstetrics **90**, 140–141 (2005)
18. Torloni, M.R., Sass, N., Sato, J.L., Renzi, A.C.P., Fukuyama, M., de Lucca, P.R.: Clinical formulas, mother's opinion and ultrasound in predicting birth weight. Sao Paulo Medical Journal **126**, 145–149 (2008)
19. Malik, N., Shahzad, S., Malik, S.M., Anwar, S.: Comparison of two different methods for estimation of fetal weight at term. RMJ 38–41 (2012)
20. Johnson, R.W.: Calculations in estimating fetal weight. American Journal of Obstetrics & Gynecology **74**, 929 (1957)
21. Siswosudarmo, H.: Detection of low birth weight neonates at term pregnancy using fundal height measurement—Deteksi bayi berat lahir rendah pada kehamilan aterm dengan pengukuran tinggi fundus. Berkala Epidemiologi Klinik & Biostatika Indonesia **1**, 78–84 (1995)
22. Siswosudarmo, R., Titisari, I.: Developing a new formula for estimating birth weight at term pregnancy. Jurnal Kesehatan Reproduksi **1** (2014)
23. Wijayanti, Y.: Accuracy difference between the Risanto's and the Johnson's formulas to estimate fetal weight based on fundal height—Perbedaan akurasi antara rumus Risanto dan rumus Johnson untuk mengestimasi berat bayi berdasarkan tinggi fundus uteri Universitas Sebelas Maret (2013)
24. Chithra, S.C., Kumari, L.K., Sangeereni, M.: Comparative study of fetal weight estimation using Hadlock's and Johnson's formula and its correlation with actual birth weight. International Journal of Scientific Study **2**, 163–170 (2014)
25. Levenbach, H.: Taming uncertainty: all you need to know about measuring forecast accuracy, but are afraid to ask (2015)
26. Swanson, D.A., Tayman, J., Bryan, T.: MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts. Journal of Population Research **28**, 225–243 (2011)
27. Carey, R.G.: Improving Healthcare with Control Charts: Basic and Advanced SPC Methods and Case Studies, 1st edn. ASQ Quality Press, Milwaukee (2003)
28. Morse, K., Williams, A., Gardosi, J.: Fetal growth screening by fundal height measurement. Best Practice & Research Clinical Obstetrics & Gynaecology **23**, 809–818 (2009)
29. Santjaka, H.I., Handayani, R.: Accuracy of fetal weight estimation based on statistics and fundal height—Studi ketepatan taksiran berat janin berdasarkan statistik dan tinggi fundus uteri. Journal Bidan Prada, **2** (2011)

Ensemble Noise Filtering for Streaming Data Using Poisson Bootstrap Model Filtering

Ashwin Satyanarayana and Rosemary Chinchilla

Abstract Ensemble filtering techniques filter noisy instances by combining the predictions of multiple base models, each of which is learned using a traditional algorithm. However, in the last decade, due to the massive increase in the amount of online streaming data, ensemble filtering methods, which largely operate in batch mode and requires multiple passes over the data, cause time and storage complexities. In this paper, we present an ensemble bootstrap model filtering technique with multiple inductive learning algorithms on several small Poisson bootstrapped samples of online data to filter noisy instances. We analyze three prior filtering techniques using Bayesian computational analysis to understand the underlying distribution of the model space. We implement our and other prior filtering approaches and show that our approach is more accurate than other prior filtering methods.

1 Introduction

Data stream mining is the process of extracting knowledge structures from continuous, rapid data records. The goal in many data stream mining applications is to form a generalization, from a set of previous labeled training data streams such that classification accuracy for previously unobserved instances is maximized. The maximum accuracy achievable depends on the quality of the data and on the chosen learning algorithms. This paper focuses on improving the quality of training data by identifying and eliminating mislabeled instances on streaming data.

Quinlan [1] demonstrated that as noise level increases, removing noise from the mislabeled training instances (class noise) increases the predictive accuracy of the

A. Satyanarayana(✉) · R. Chinchilla
Department of Computer Systems Technology, New York City College of Technology (CUNY), N-913, 300 Jay St, Brooklyn, NY 11201, USA
e-mail: asatyanarayana@citytech.cuny.edu, rosemary.chinchilla@mail.citytech.cuny.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_75

resulting classifier. This has been empirically verified by many researchers over the last decade. For example, Brodley and Friedl ([2][3]) illustrate that for class noise levels of less than 40%, removing mislabeled instances from the training set resulted in higher predictive accuracy relative to classification achieved without "cleaning" the training data.

Gamberger et al. [4] point out that the separation of noise detection and hypothesis formation has the advantage that noisy examples do not influence hypothesis construction. This is different from the approaches that handle noise in the hypothesis formation process by trying to avoid overfitting the noisy example set.

The remainder of the paper is organized as follows. Section 2 discusses three prior significant filtering techniques. In Section 3, we mention the contributions of this paper. In Section 4 we analyze the current filtering techniques with Bayesian computational analysis. In Section 5 we present our bootstrap model filtering algorithm. In Section 6 we show empirically how our technique compares with other filtering techniques, and we finally conclude in Section 7.

2 Prior Work in Ensemble Filtering

The problem of handling noise has been the focus of much attention in machine learning especially in online data stream mining. For example, pruning in decision trees is designed to reduce the chance that the tree is overfitting to noise in the training data.

Weisberg [5] suggested the use of regression analysis to identify outliers in the training data. Those cases that could not be described by the model and have the largest residual errors are outliers. Motivated by the same idea, filtering was introduced to remove mislabeled training data. In this section, we survey three significant results in the area of filtering.

2.1 Ensemble Filtering

An ensemble classifier detects noisy instances by constructing a set of classifiers (base level detectors) [2]. A majority vote filter tags an instance as mislabeled if more than half of the m classifiers classify it incorrectly. A consensus filter requires that all classifiers must fail to classify an instance as the class given by its training label. In their empirical results, it was shown that the majority filter performs better than a consensus filter.

2.2 *Classification Filtering*

Classification filtering [4] is a filtering approach in which the training set E is partitioned into n subsets, and the set of classifiers trained from the aggregation of any $n - 1$ subsets are used to classify the instance in the excluded subset. There is only one base-level classifier used in classification filtering.

2.3 *Partition Filtering*

Partition filtering proposed by Zhu et al. [6] partitions a dataset into equally sized subsets, and each instance is evaluated by more than one model, and the results are averaged to identify the noisy instances. The major difference between classification and partition filtering is that in partition filtering each instance is evaluated by $k - 1$ hypotheses, whereas in classification filtering, each instance is evaluated once by a single model built from the remaining instances in the other $k - 1$ partitions.

3 Contributions of This Paper

The novel contributions of this paper are as follows:

1. We perform Bayesian analysis to understand the underlying noise model distribution in ensemble techniques, and show that the three prior filtering techniques use crude approximations of model averaging (Section 4).
2. We present our novel ensemble filtering technique which extends the model space by Poisson bootstrap samples and learns on those samples using multiple inductive learners (Section 5).
3. We compare our approach with a single filter, a prior filtering technique (partition filtering) and online bagging, and plot the different techniques using (a) an accuracy vs noise level graph and (b) learning curves.

4 Bayesian Analysis

Ensemble techniques can be considered as performing *model averaging*. Model averaging framework is to build a set of models, obtain model specific estimates and use the rules of probability to average over all the instances. It has been shown that model averaging works by extending the model space [9] and that several classifiers produces a higher accuracy as compared to any single model [7]. In this section we use Bayesian analysis to understand the noise model distribution in ensemble techniques.

Let $\vec{x}$ represent the original training set, $\vec{y}$ the corresponding class labels and θ a model (or hypothesis) in the model space Θ. Mathematically, an unseen test instances x, is assigned to a class y that maximizes the following equation:

$$Pr(y|\vec{x}, \Theta) = \sum_{\theta \in \Theta} Pr(y|x, \theta).Pr(\theta|\vec{x}, \vec{y}) \tag{1}$$

By Bayes' theorem, and assuming the instances are drawn independently, the posterior probability of θ is given by:

$$Pr(\theta|\vec{x}, \vec{y}) = \frac{Pr(\theta)}{Pr(\vec{x}, \vec{y})} \prod_{i=1}^{n} Pr(x_i, y_i|\theta) \tag{2}$$

where $Pr(\theta)$ is the prior probability of θ, and the product of $Pr(x_i, y_i|\ \theta)$ is the likelihood. The data prior $Pr(\vec{x}, \vec{y})$ is the same for all models and can thus be ignored.

In order to compute the likelihood in equation (2), it is necessary to compute the probability of a class label y_i, given an unlabeled instance x_i, and the hypothesis θ since:

$$Pr(x_i, y_i|\theta) = Pr(x_i, \theta).Pr(y_i|x_i, \theta) \tag{3}$$

The probability $Pr(y_i|x_i, \theta)$ is called the *noise model*. We assume a uniform class model in which each instance's class is corrupted with probability ϵ, and thus:

$Pr(y_i|x_i, \theta) = 1 - \epsilon$ if θ predicts the correct class y_i, and

$Pr(y_i|x_i, \theta) = \epsilon$ if θ predicts an incorrect class

Equation (2) then becomes:

$$Pr(\theta|\vec{x}, \vec{y}) \propto Pr(\theta)(1 - \epsilon)^s \epsilon^{n-s} \tag{4}$$

where s is the number of instances correctly classified by θ, and the noise level can be estimated by the models' average error rate.

Using Equation (4) in (1) we get:

$$Pr(y|\vec{x}, \Theta) \propto \sum_{\theta \in \Theta} Pr(y|x, \theta).Pr(\theta).(1 - \epsilon)^s \epsilon^{n-s} \tag{5}$$

As $n \to \infty$ the distribution of the noise model tends to $Poisson(1)$ distribution. Therefore for each training instance (as discussed in Section 5), we choose that instance $K \sim Poisson(1)$ times independently. We now analyze the computation of the three prior filtering techniques.

4.1 Bayesian Analysis of Ensemble Filtering

In the case of a majority vote filter, the probability of classifying an unseen instance x to a true label t, can be represented by the following equation:

$$Pr(y = t|x, \Theta) = \frac{1}{|\Theta^*|} \sum_{\theta \in \Theta^*} \delta(t, \theta(x)) \tag{6}$$

where $\delta(t, \theta(x)) = 1$ if the class label predicted by the classifier for the test case x is the true label t, and 0 otherwise. Θ^* represents the most probable classifiers of the model.

If $Pr(y = t|x, \Theta) < \frac{1}{2}$, then the instance is noisy and filtered.

4.2 Bayesian Analysis of Classification Filtering

As mentioned in the previous section, classification filtering performs k-fold cross validation. K-fold cross validation is a commonly used technique which takes a set of n examples and partitions them into k sets of size $\frac{n}{k}$. For each fold, a classifier is trained on the other folds and tested on the current fold. Thus k hypotheses $\theta_1, \theta_2, \ldots, \theta_k$ are generated. This prediction is equivalent to outputting the average of k-hypotheses as shown below:

$$Pr(y = t|x, \Theta) = \frac{1}{k} \sum_{i=1}^{k} \delta(t, \theta_i) \tag{7}$$

where δ is a 0-1 loss function, which returns 1 if θ_i predicts the correct label t, else returns 0. Equation (7) captures the cross validation step of the classification filtering algorithm. If t is the true class label, and θ is the final model built, then it performs the following computation:

$$Pr(y = t|x) = Pr(y = t|\vec{x}, \theta).Pr(\theta|\vec{x}, \vec{y}) \tag{8}$$

If $Pr(y = t|x) < \frac{1}{2}$, then the instance is noisy and filtered.

4.3 Bayesian Analysis of Partition Filtering

In the case of partition filtering, there are k different hypotheses and unlike the previous case, each instance is evaluated on each of the $k - 1$ hypotheses. If more than half of the hypotheses classify the instance as noisy, then it is eliminated. That is, each instance is evaluated by an average of the other models built.

$$Pr(y = t|x) = \frac{1}{k-1} \sum_{i=1}^{k-1} \delta(t, \theta_i) \tag{9}$$

where δ is a 0-1 loss function, which returns 1 if θ_i predicts the correct label t, else returns 0

If $Pr(y = t|x) < \frac{1}{2}$, then the instance is noisy and filtered.

5 Bootstrap Model Filtering

In all the prior filtering techniques, we find that by having more than one model to evaluate the instances, they extend the model space as compared to a single filter. We also observe that in all the three techniques, a crude approximation of *model averaging* is performed.

Consider the typical view of mining where a single training set of size n is available from the underlying distribution that generated the data F. However this view masks the underlying uncertainty in the data, namely that the training data we have is one of many that could have been chosen. If we were to build a model for each possible data set we would have a probability distribution over the model space. However, typically we do not have the luxury of many different data sets drawn independently of the same size so that we can compute the uncertainty over the model space from them. To approximate this distribution using a single dataset, Efron [8] created the bootstrapping approaches. Non-parametric bootstrapping is an example of attempting to simulate draws from F, where no knowledge of its form is known. Formally, the computation performed would be:

$$Pr(y|\vec{x}, \Theta) = \sum_{D', \theta \in \Theta} Pr(y|x, \theta).Pr(D'|D) \tag{10}$$

where the datasets D' are the different bootstrap samples generated from the original dataset D.

The standard bootstrap procedure creates each simulated data set by drawing observations from x with replacement. In any given resample, each observation may occur 0, 1 or more times according to $Binomial(n, \frac{1}{n})$. And since the total number of observations is constrained to be n, the counts are jointly $Multinomial(n, \frac{1}{n}, \ldots, \frac{1}{n})$. We also note that:

$$\lim_{n \to \infty} Binomial(n, \frac{1}{n}) = Poisson(1) \tag{11}$$

Therefore for each training instance, we choose that instance $K \sim Poisson(1)$ times independently. We refer to this as *Poisson Bootstrap Sampling*.

Our Bootstrap Model Filtering algorithm as shown in Figure 1 begins with n Poisson bootstrapped samples of our dataset E (step 1) and an empty output set A of detected noisy examples (step 2). The main loop (steps 3-12) is repeated for each bootstrap sample E_i. E_i is used as an input for the k inductive learning algorithms to generate models k models $\theta_{i,1}, \theta_{i,2} \cdots \theta_{i,k}$. In step 7, we form a set E_t which includes all the examples from E except E_i. The set E_t is evaluated by our k models in steps 8-11. If more than half of the models misclassify an instance, then it is treated as noise and eliminated.

Algorithm. BootstrapModelFiltering (E, n, k)

Input: E (training set),
n (number of Poisson bootstrap samples),
k (number of inductive learning algorithms)
Output: A (a detected noisy subset of E)
1: Form n Poisson bootstrap samples $E_{1,\ldots,n}$ of E
2: $A \leftarrow \emptyset$
3: **for** $i = 1, \ldots, n$ **do**
4: **for** $j = 1, \ldots, k$ **do**
5: $\theta_{i,j} \leftarrow$ model built from bootstrap sample E_i and inductive algorithm j
6: **end for**
7: $E_t \leftarrow E \backslash E_i$
8: **for** every $e \in E_t$ **do**
9: If e is misclassified by more than half of the $\theta_{i,j}$ models built, then it is noisy and needs to be eliminated.
10: $A \leftarrow A \cup \{e\}$
11: **end for**
12: **end for**

Fig. 1 Bootstrap Model Filtering Algorithm

6 Empirical Results

In this section, we discuss experiments using J48, RandomForest and RandomTree as the base classifier models for the bootstrap model filtering approach. We tested our approach on several UCI [12] and other datasets of varying (small, medium and large) sizes. For each dataset, we compare the accuracies after filtering using the following techniques:

1. Single Model: We used decision trees (J48) as our single filtering base model.
2. Partition Filtering: Partition Filtering is chosen as it outperforms Classification and Ensemble Filtering techniques [6]. We use Decision Tree as the base model.
3. Online Bagging: We implemented online bagging as illustrated by Oza [10] using Naïve Bayes as the base model.

Table 1 Sizes of Small Datasets

Dataset	# of instances	# of attributes	# of classes
Diabetes	768	9	2
Liver Disorder	345	7	2
Hepatitis	155	20	2
Dengue	846	18	2

6.1 Small Datasets

We compare our approach on small UCI datasets (in the area of medicine). Medicine is an area where the quality of the data is essential for accurate prediction. The datasets selected are: Diabetes, Hepatitis, Liver Disorders and Dengue [11] disease as shown in Table 1.

For each dataset, for bootstrap model filtering we use 30 Poisson bootstrap samples with 3 different classifiers (J48, RandomForest and RandomTree) on each sample. The results of our experiments on small datasets are shown in Table 2.

Table 2 Classification Accuracy after eliminating noisy instances

Dataset	Decision Tree (J48)	Partition Filtering	Online Bagging	Bootstrap Model Filtering
Diabetes	0.73	0.75	0.79	**0.88**
Liver Disorder	0.68	**0.74**	**0.76**	**0.79**
Hepatitis	0.83	0.84	0.88	0.89
Dengue	0.74	0.77	0.81	**0.92**

6.2 Medium and Large Datasets

In this section, we used our approach on several medium and large datasets as shown in Table 3. In each of the medium datasets (Abalone, CMC, Car Evaluation, Multiple Features, Waveform and Wine Quality) we use 50 Poisson bootstrap samples. For the massive datasets (Nursery, Letter, Electricity board, and Localization) we used 100 Poisson bootstrap samples. The results are as shown in Table 4.

The boldface entries in Table 2 and Table 4 represent cases when the ensemble technique significantly (t-test, $\alpha = 0.05$) outperformed a single filtering technique. As we can see from the results, for small (e.g. Diabetes) and very large sized datasets (e.g. Localization, Electricity Board), we see that our technique outperforms the other techniques significantly. For medium sized, the ensemble techniques show a very small improvement in accuracy.

Table 3 Sizes of the Medium and Large datasets

Dataset	# of instances	# of attributes	# of classes
CMC	1473	10	3
Car Evaluation	1728	7	4
Multiple Features	2000	48	10
Abalone	2924	9	28
Waveform	3500	41	3
Wine Quality	3429	12	11
Nursery	12960	9	5
Letter Recognition	20000	17	26
Electricity Board	45781	5	31
Localization	164860	8	11

Table 4 Classification Accuracy after Filtering

Dataset	Decision Tree (J48)	Partition Filtering	Online Bagging	Bootstrap Model Filtering
CMC	0.52	0.65	**0.69**	**0.75**
Car Evaluation	0.92	0.94	0.95	0.97
Multiple Features	0.72	0.74	0.76	**0.78**
Abalone	0.20	0.34	**0.38**	**0.48**
Waveform	0.75	0.76	0.77	**0.82**
Wine Quality	0.56	0.58	0.61	**0.69**
Nursery	0.97	0.97	0.98	0.98
Letter Recognition	0.87	0.88	0.88	0.89
Electricity Board	0.65	0.66	0.68	**0.73**
Localization	0.76	0.73	0.79	**0.89**

6.3 Classification Accuracy vs Noise

Figure 2 shows the effect of filtering on classification accuracy on the test set by each of the four techniques. We introduce noise levels from 10% to 40% on each of the datasets by adding mislabeled instances in the test set. At 10% noise level we see that there is no significant difference for any of the ensemble techniques used, however there is significant difference when compared to the single filter. For noise levels up to 20%, we see that online bagging and bootstrap model filtering significantly outperform the other two techniques. For noise levels between 30% and 40%, bootstrap model filtering outperforms the other techniques significantly. We also see that bootstrap model filtering approach does well on very large datasets with higher noise levels.

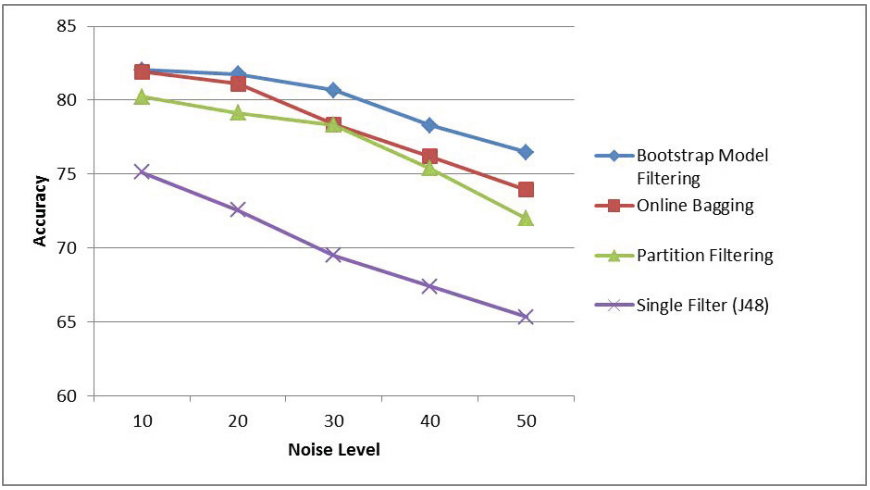


Fig. 2 Simulation results for the network.

6.4 Learning Curve Comparison

To illustrate the learning curve convergence in all the techniques, we tried the car evaluation dataset, which has about 1728 instances. We used 528 as the test set, and used 200, 400, 600, 800 and 1000 instances as streams of training data. We see in Figure 3 that bootstrap averaged ensemble technique performs better than the other two ensemble technique in terms of the number of instances needed for accuracy convergence, as the inductive learning algorithm learns from several small bootstrapped samples.

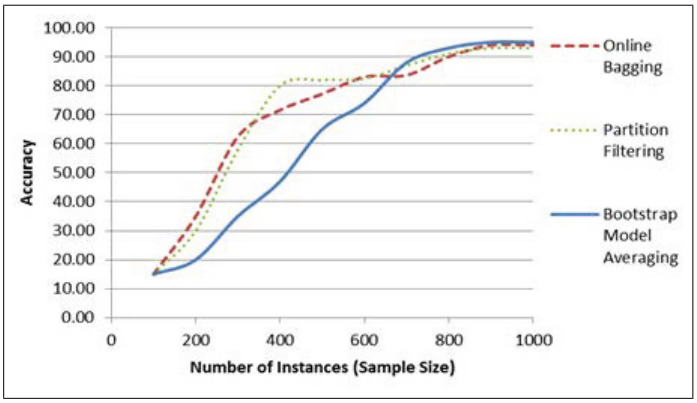


Fig. 3 Learning curve convergence of ensemble filtering techniques for the car evaluation dataset.

7 Conclusion

Resolving data quality issues in stream mining is often one of the biggest efforts in data mining. Prior filtering techniques required multiple passes of data. The main challenge in applying prior filtering techniques to data streams is that a stream is theoretically infinite and that means that data has to be processed in a single pass using little memory. Also, using all the available data is prohibitive due to memory and time constraints. In this paper, we attempt to address this problem for different sized datasets by using Poisson bootstrap samples with multiple base classifiers.

We analyzed prior filtering techniques by using Bayesian computation analysis. We then used that knowledge and extended the model space by using 1) several bootstrap samples and 2) multiple classifiers. Our empirical results show that although online bagging performs well for medium sized datasets in terms of computational size required, the bootstrap model filtering approach outperforms other approaches in the case of small and massive datasets.

Acknowledgment We would like to thank the Emerging Scholars undergraduate research program at New York City College of Technology for supporting Rosemary Chinchilla in this research. Ashwin Satyanarayana's research was supported by PSC CUNY grant # 68180-00 46 of the Research Foundation of City University of New York (CUNY).

References

1. Quinlan, J.R.: Induction of decision trees. Machine Learning **1**(1), 81–106 (1986)
2. Brodley, C.E., Friedl, M.A.: Identifying and eliminating mislabeled training instances. In: AAAI/IAAI, vol. 1, pp. 799–805 (1996)
3. Brodley, C.E., Friedl, M.A.: Identifying mislabeled training data. Journal of Artificial Intelligence Research, 131–167 (1999)
4. Gamberger, D., Lavrac, N., Groselj, C.: Experiments with noise filtering in a medical domain. In: ICML, pp. 143–151 (1999)
5. Weisberg, S.: Applied linear regression, vol. 528. John Wiley and Sons (2005)
6. Zhu, X., Wu, X., Chen, Q.: Eliminating class noise in large datasets. In: ICML, vol. 3, pp. 920–927, August 2003
7. Davidson, I., Fan, W.: When efficient model averaging out-performs boosting and bagging. In: Knowledge Discovery in Databases: PKDD 2006, pp. 478–486 (2006)
8. Efron, B., Tibshirani, R.J.: An introduction to the bootstrap. CRC Press (1994)
9. Satyanarayana, A.: Data mining for large datasets: intelligent sampling and filtering. Doctoral Dissertation, State University of New York, Albany (2006)
10. Oza, N.C.: Online bagging and boosting. In: IEEE International Conference on Systems, Man and Cybernetics, vol. 3, pp. 2340–2345, October 2005
11. Shakil, K.A., Anis, S., Alam, M.: Dengue disease prediction using weka data mining tool, arXiv preprint arXiv:1502.05167 (2015)
12. Lichman, M.: UCI Machine Learning Rep. University of California, School of Information and Computer Science, Irvine, CA (2013). http://archive.ics.uci.edu/ml

Mining Persistent and Dynamic Spatio-Temporal Change in Global Climate Data

Jie Lian and Michael P. McGuire

Abstract The potential impacts of climate change on natural and man-made systems can have a drastic effect on life on Earth. The application of data mining algorithms on global climate data can result in a better understanding of the climate system. Of which, change detection has proven to be a very useful approach when mining climate data. Understanding spatio-temporal change can give insight to interesting patterns that can be used to predict climate events. This paper proposes a method to generate spatial homogeneous regions that uses a novel indexing structure for the analysis of spatial change including homogeneous change and heterogeneous change. The resulting regions are then used to analyze persistent and dynamic regions at longer time scales. The efficacy of the approach was demonstrated on a real-world climate dataset and the results suggest interesting patterns that are explained by known climate phenomena.

1 Introduction

Scientific results have suggested that there is little doubt that the Earth's climate is in a period of change. Climate change can have many impacts ranging from sea level rise, severe storms, severe droughts, and natural disasters (Barros et al. 2015). These impacts can have a drastic effect on vulnerable populations resulting in both degraded ecological and socioeconomic conditions and in severe cases, loss of life. Because of this, a massive amount of data has been produced regarding Earth's climate system.

Data mining and machine learning have been used to uncover interesting patterns in climate data (Faghmous and Kumar 2014). In particular, the biggest question in climate science is to verify and quantify global climate change (Shekhar et al. 2015). The analysis of change in climate data can generally be classified into two categories:

J. Lian · M.P. McGuire(✉)
Towson University, Towson, MD 21252, USA
e-mail: mmcguire@towson.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_76

spatial change and temporal change (Zhou et al. 2014). Spatial change is to detect change in space at only one point in time, such as the analysis of air temperature change at different latitudes. Temporal change is concerned with change associated variable over time such as the change in global air temperature from 1990 to 2015. The detection of change in both space and time simultaneously can be used to find more interesting patterns in climate data where regions can be characterized by the type and extent of spatio-temporal change occurring for a given variable.

This paper introduces a novel approach to find spatio-temporal change in climate data. The pattern of spatio-temporal change characterizes a change in processes occurring in spatial regions during a time interval (Shekhar et al. 2015). In particular, the method presented in this paper finds both homogeneous change, represented by a spatially uniform change over whole region, and heterogeneous change, represented by a spatially variable change over a region. The proposed approach first finds homogeneous regions in gridded climate data for each time period. Then, a novel indexing structure is used to characterize spatial change over time. The results of which are then used to analyze high change dynamic regions and high change persistent regions in this time period. The efficacy of the approach was demonstrated on a real-world climate dataset where the resulting spatio-temporal change indicate conditions of known global climate regions. The rest of the paper is organized as follows: Section 2 reviews literature that is related to this work, Section 3 presents the method for detecting and classifying changes in spatial and analyzing spatial change over long time scales, Section 4 presents the experimental results, and Section 5 summarizes the results and offers suggestions for future research.

2 Related Work

Over the last decade, data mining techniques have been applied to climate datasets. Ganguly and Steinhaeuser (2008) presented that even a small data mining algorithm can bring about interesting insight in climate science. Clustering algorithm is a data mining approach used to detect spatial and temporal patterns in climate data (Hoffman et al. 2005). One such study clusters typhoon images by taking advantage of k-means and self organizing maps (SOM) (Kitamoto 2002). Other methods have focused on discovering climate indices using a clustering-based methodology (Steinbach et al. 2003), (Steinhaeuser et al. 2011). Furthermore, artificial neural networks and decision trees have been used to predict weather forecasts (Olaiya and Adeyemo 2012). The research presented in this paper adds to the existing literature by utilizing a quadtree-based clustering method to divide spatial areas into multiple homogeneous regions based on its internal variation.

Another related research area is the analysis of spatial and temporal cluster change. For example, Gunnemann et al. (2012) introduced an approach that traced subspace cluster behavior in temporal climate data. Also, Rinsurongkawong and Eick (2009) presented clustering-based approaches to perform change analysis in spatial earth-

quake datasets. The approach presented in this paper builds on the existing literature where a novel indexing method is used to track homogeneous and heterogeneous change between time periods.

Spatio-temporal change detection is another research field that is relevant to the work presented in this paper. For example, Kleynhans et al. (2014) provided a method to evaluate land cover change detection by estimating the spatial covariance of a hyper-temporal time series at each point in time for a specific pixel neighborhood and MODIS band combination. However, the methods that use spatio-temporal data mining can dedicate to improving the analysis of large spatio-temporal databases (Andrienko et al. 2006). Pekerskaya et al. (2006) demonstrated a decision tree based approach to detect changing regions between two spatial snapshots. While this approach identifies changes efficiently and accurately by comparing two cluster-embedded decision trees, it only detects change between two consecutive days rather than over long time scales. Finally, McGuire et al. (2014) provided a method based on spatial autocorrelation to find dynamic regions in space across time periods and also find dynamic time periods for each location in space.

The research presented in this paper builds on the existing literature focused on spatio-temporal change detection in climate data. The contributions of this paper can be summarized as below. First, a novel efficient quadtree-based index structure was developed to detect spatio-temporal change. Second, an algorithm was developed not only to detect high change regions, also classify changes to heterogeneous change and homogeneous change. Third, spatio-temporal change is analyzed at long time scales to find high change persistent regions and high change dynamic regions. Finally, the approach is validated using real-world climate data which reveals climate pattern in known climate zones.

3 Methods

3.1 Preliminaries

Given a three dimensional spatio-temporal array $X(i, j, k)$ where the ith and jth dimensions of the array represent the spatial locations, Loc=$\{loc_1, loc_2, ..., loc_n\}$, such that each $loc_i \in Loc$ is defined by a set of spatial coordinates in two dimensional Euclidean space (loc_{ix}, loc_{iy}) and each loc_i has a set of spatial neighbors in the x and y dimensions. The kth dimension of the spatio-temporal array represents an ordered set of timestamps T=$\{t_1, t_2, ..., t_m\}$ where $t_1 < t_2 < ... < t_m$. Each cell in array $X(i, j, k)$ is a measurement value for a particular variable of interest (VOI).

3.2 Objectives

Given the $X(i, j, k)$, the first objective of the approach is to find homogeneous regions $\zeta = \{C_1, C_2, ..., C_n\}$ of a VOI in a timestamp t_j. Then given the homogeneous regions, the next objective is to characterize spatial change of the VOI over time. The homogeneous region $C_i \in \zeta$ has three possible spatial changes in detail; which are high change heterogeneous, high change homogeneous and high change stable. Based on this characterization of spatial change, the overall goal is to detect high change dynamic regions and high change persistent regions over long time scales in gridded climate data.

3.3 Approach

The proposed approach shown in Fig. 1, is comprised of two parts. The first part is to find spatially homogeneous regions for each $t_j \in T$ for the VOI in a three dimensional spatio-temporal array $X(i, j, k)$. The second part of the approach analyzes spatial change over long time scales to characterize high change dynamic and high change persistent regions.

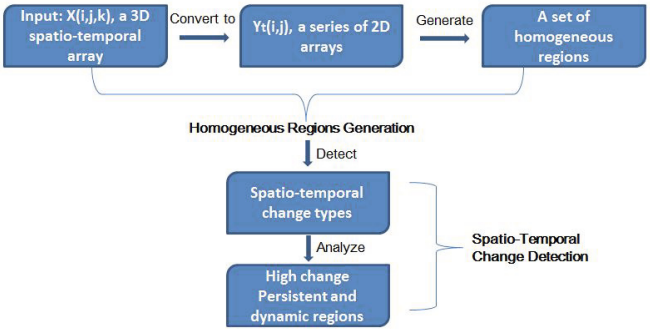


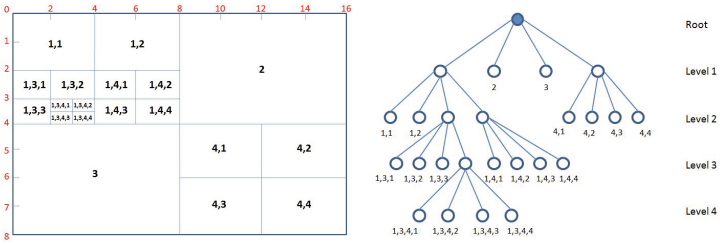
Fig. 1 Overall approach

Homogeneous Regions. The approach uses a quadtree-based algorithm to find homogeneous regions in the gridded dataset. A quadtree is a tree data structure where the root node is split into four internal nodes or quadrants, such that Q=$\{q_1, q_2, q_3, q_4\}$, and each internal node can be recursively subdivided into four internal nodes again until each internal node reaches a variance threshold and then becomes a leaf node(Samet 1984). In this approach, we define a homogeneous region as follows:

Definition 1 (Homogeneous Region). *Given a timestamp t_j, where $t_j \in T$ and a variable of interest VOI, a set of homogeneous regions $\zeta = \{C_1, C_2, ..., C_n\}$ is a set*

of sets generated by t_j, for each $t_j \in T$ based on the value of a VOI and a region variance threshold θ, where the region variance (σ^2)<θ. And each $C_i \in \zeta$ denoted by an index, covers a set of spatial locations, $SLoc=\{loc_m, loc_{m+1}, ..., loc_{m+n}\}$, additionally, one spatial location loc_m can only belong to one particular homogeneous region, C_i. The specific index of a homogeneous region C_i is given an index value V={1,2,3,4} where 1 represents the top-left region, 2 the top-right region, 3 the bottom-left region, and 4 the bottom-right region.

The presented approach first converts the 3 dimensional spatio-temporal array $X(i, j, k)$ into a series of 2 dimensional spatial arrays for each time period $Y_t(i, j) \in Y(i, j)=\{Y_{t_1}(i, j), Y_{t_2}(i, j), ..., Y_{t_m}(i, j)\}$. And then, a quadtree is built for each 2d spatial array $Y_t(i, j)$ where the stopping criteria for the quadtree is the variance threshold θ. Fig. 2 (a) shows an example of the quadtree structure, given the input



(a) Quadtree Example (b) Tree Structure

Fig. 2 A quadtree example

array Y(0:16, 0:8). First the variance of array Y is calculated, if the variance exceeds θ, the original space will be partitioned into four subspaces; Y(0:8, 0:4), Y(8:16, 0:4), Y(0:8, 4:8) and Y(8:16, 4:8) in the first level; which is denoted by index {1,2,3,4} respectively. The algorithm runs recursively to further divide each resulting region until the variance for each quadrant is less than or equal to the threshold θ. Ultimately, one index will signify one area of the whole space, e.g. the index (1,4,1) indicates array Y(4:6, 2:3). The corresponding tree data structure of the example is shown in Fig. 2 (b).

Spatio-Temporal Change Detection. Given a set of homogeneous regions $\zeta_i = \{C_1, ..., C_a\}$, the second part of the approach is to detect spatio-temporal change between two consecutive timestamps. In this section, we first define the change types and explain the algorithm used to find them. Then, we describe an approach to analyze the high change dynamic and high change persistent regions at longer time scales.

Two types of spatio-temporal change are detected. First, homogeneous change is characterized by regions that have a decrease in variance from timestamp 1 to timestamp 2. This indicates that change in these regions is uniform across the entire

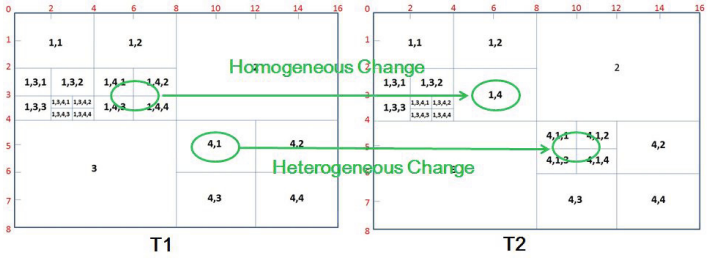


Fig. 3 Spatio-temporal change types

region such that separate adjacent regions in timestamp 1 will be merged into a single region in timestamp 2. Conversely, heterogeneous change is characterized by regions that have increasing variance between timestamp 1 and timestamp 2. Therefore, the region in timestamp 1 is divided into multiple regions in timestamp 2. Fig. 3 shows examples of heterogeneous and homogeneous change. The following gives formal definitions for each type of spatio-temporal change:

Definition 2 (Homogeneous Change). *Given two homogeneous regions* ζ_i *and* ζ_j *at time* t_p *and* t_q *respectively, where* $(p < q)$ *and region* $C_m \in \zeta_i = \{C_1, ..., C_a\}$ *and region* $D_n \in \zeta_j = \{D_1, ..., D_b\}$*, the homogeneous change regions are the regions* C_m*,* C_{m+1}*,* C_{m+2}*,* C_{m+3} *in* ζ_i *indexed by* $(a, ..., b, 1)$*,* $(a, ..., b, 2)$*,* $(a, ..., b, 3)$ *and* $(a, ..., b, 4)$*, where* $(a, b) \in \{1, 2, 3, 4\}$ *merged to region* D_n *in* ζ_j*, indexed by* $(a, ..., b)$*.*

Definition 3 (Heterogeneous Change). *Given two homogeneous regions* ζ_i *and* ζ_j *at time* t_p *and* t_q *respectively, where* $(p < q)$ *and region* $C_m \in \zeta_i = \{C_1, ..., C_a\}$ *and region* $D_n \in \zeta_j = \{D_1, ..., D_b\}$*, the heterogeneous change regions are the region* C_m *in* ζ_i *indexed by* $(a, ..., b)$*, where* $(a, b) \in \{1, 2, 3, 4\}$*, split to the regions* D_n*,* D_{n+1}*,* D_{n+2}*,* D_{n+3} *in* ζ_j *indexed by* $(a, ..., b, 1)$*,* $(a, ..., b, 2)$*,* $(a, ..., b, 3)$ *and* $(a, ..., b, 4)$*.*

Once the type of spatio-temporal change is characterized, the next step is to find high change regions. The high change heterogeneous regions are the regions in heterogeneous change but which absolute difference of timestamp 1 and timestamp 2, that is Δ; exceeds the changing threshold λ. The high change homogeneous regions are the regions in homogeneous change but the Δ of which is larger than λ also. In the same way, high change stable regions are calculated based on its Δ and λ thresholds. The threshold λ is determined by using 1.5 times interquartile range (IQR), which is a measure of statistical dispersion. In our methodology, the IQR is based on the difference of timestamp 1 and timestamp 2. Therefore, if the absolute difference is greater than 1.5*IQR, which means the region changes significantly at this time period, then it will be categorized as high change region.

The algorithm for spatio-temporal change detection is shown in Algorithm 1. The algorithm takes the input ζ_i and ζ_j for the set of homogeneous regions for t_i and t_j, where $(i < j)$. First, the stable regions can be generated when C_p is equal

to D_q. Then compare the absolute difference of this region from t_i to t_j with the threshold λ, if Δ is larger than λ, this region will be classified as a high change stable region, shown in line 7-8. If a region is not a stable region, it must belong to homogeneous change or heterogeneous change. In the next step, get the index length l_p of region C_p, and the index length l_q of region D_q. If l_p is greater than l_q, that means the index length of region C_p is longer than the index length of region D_q. Then define a variable MergeCandidate; which is the C_p in length l_q shown in line 10. The MergeCandidate means the after-merged index structure of C_p if C_p merges from t_1 to t_2. If the MergeCandidate is the same as D_q, that indicates C_p merges to region D_q. If Δ is larger than λ then this region is a high change homogeneous region. Likewise, if the l_p is smaller than l_q, and define SplitCandidate; which is the D_q in length l_p shown in line 14. The SplitCandidate means the index structure of after-splitted C_p if C_p splits from t_1 to t_2. If SplitCandidate equals C_p that means C_p is the heterogeneous region that split to multiple regions, and one of which is D_q. Also compare the Δ with λ to determine if it is a high change heterogeneous region. The time complexity is $O(n^2)$ because there are two loops in this algorithm where one loop is nested. Given the resulting set of changing regions, we can then analyze

Algorithm 1. Spatio-Temporal Change Algorithm

Require: The homogeneous regions ζ_i, where $C_m \in \zeta_i = \{C_1, ..., C_a\}$ and the homogeneous regions ζ_j, where $D_n \in \zeta_j = \{D_1, ..., D_b\}$ of time i and time j;
Ensure: High change stable; High change heterogeneous; High change homogeneous
1: HC_stable=$\emptyset$;HC_heterogeneous=$\emptyset$;HC_homogeneous=$\emptyset$;
2: **for** $C_p \in \zeta_i$ **do**
3: l_p=length(C_p);
4: **for** $D_q \in \zeta_j$ **do**
5: l_q=length(D_q);
6: **if** C_p==D_q
7: **if** $\Delta_{CD} > \lambda$
8: HC_stable+=C_p;
9: **else if** $l_p > l_q$
10: MergeCandidate=$C_p(1{:}l_q)$;
11: **if** D_p==MergeCandidate && $\Delta_{CD} > \lambda$
12: HC_homogeneous+=C_p;
13: **else if** $l_p < l_q$
14: SplitCandidate=$D_q(1{:}l_p)$;
15: **if** C_p==SplitCandidate && $\Delta_{CD} > \lambda$
16: HC_heterogeneous+=C_p;
17: **end for**
18: **end for**

change over longer time scales. This analysis is done at the grid cell level where a high change stable grid cell is a grid cell that belongs to stable quadrants for a long period of time but the value changes significantly from day to day. And high change dynamic grid cells are the grid cells that belong to high change heterogeneous or high change homogeneous quadrants for long time periods. In particular, we will describe the algorithm of how to analyze the two types grid cells in Algorithm 2. Two thresholds $\alpha 1$, $\alpha 2$, where $0 \leq \alpha 1, \alpha 2 \leq 1$, are chosen to determine if the grid cell is a high change persistent or high change dynamic grid cell respectively. Given the quadrant change types $R(i, j, k-1)$ of a continuous timestamp generated from

array $X(i, j, k)$, each grid cell is assigned a number that represents quadrant change type, such that 0 means stable, 1 means high change stable, 2 means high change homogeneous and 3 means high change heterogeneous. The first step of this method is to retrieve the high change survival rate and high change dynamic rate of each grid cell in line 6,7. The high change survival rate is the percentage of the time period that the grid cell is a high change stable quadrant. And high change dynamic rate is the percentage that the grid cell is high change homogeneous or high change heterogeneous change type. If the high change survival rate of a grid cell is more than $\alpha 1$, it will be signified as a high change persistent grid cell, shown in line 13. And to determine if a grid cell is a high change dynamic grid cell, the high change dynamic rate of which must be larger than $\alpha 2$.

Algorithm 2. High Change Dynamic and Persistent Grids Analysis Algorithm

Require: The spatio-temporal change type results $R(i, j, k-1)$ that generated from the 3 dimensional spatio-temporal array X(i,j,k);
Ensure: hc_persistGrid; hc_dynamicGrid
1: **for** $a \in i$ **do**
2: **for** $b \in j$ **do**
3: len=length(R(a,b,:)==0)+length(R(a,b,:)==1);
4: HCS_len=length(R(a,b,:)==1);
5: HCD_len=length(R(a,b,:)==2)+length(R(a,b,:)==3);
6: hc_survivalRate(a,b)=HCS_len/(k-1);
7: hc_dynamicRate(a,b)=HCD_len/(k-1);
8: **end for**
9: **end for**
10: **for** $a \in i$ **do**
11: **for** $b \in j$ **do**
12: **if**hc_survivalRate(a,b)$\geq \alpha 1$
13: $hc\_persistGrid \leftarrow (a, b)$
14: **if**hc_dynamicRate(a,b)$\geq \alpha 2$
15: $hc\_dynamicGrid \leftarrow (a, b)$
16: **end for**
17: **end for**

4 Experimental Results

To demonstrate the efficacy of the approach presented in this paper, experiments were performed to detect high change persistent and high change dynamic regions in the NOAA Reanalysis I Dataset (Kalnay et al. 1996). The dataset consists of a series of daily values of 2.5°by 2.5°grids from 1948 to present at 17 atmospheric pressure levels. The dataset contains a number of variables such as air temperature, relative humidity, surface level pressure. In this experiment, high change persistent and high change dynamic regions are shown for the variables precipitable water and surface level pressure for the year 2012.

First, Algorithm 1 was used to generate homogeneous regions and characterize heterogeneous and homogeneous change between each time step. The value of θ when generating homogeneous regions was determined by using a sensitivity analysis and selecting the value where any increase in θ results in a marginal decrease in quadrant variance. As was mentioned in section 3, the value of λ for Algorithm

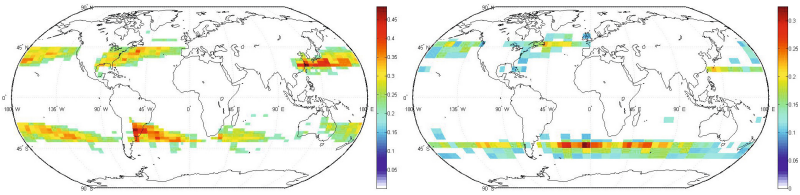


Fig. 4 High change persistent regions (left) and high change dynamic regions (right) for precipitable water 2012

1 was determined by using 1.5 times interquartile range (IQR). this resulted in an acceptable set of homogeneous regions for both variables. Once the spatial change was characterized for each time step, Algorithm 2 was used to produce high change persistent and high change dynamic regions. In Algorithm 2, $\alpha 1$ and $\alpha 2$ were chosen by a heuristic approach that increase α from a small value until a reasonable number of high change persistent and high change dynamic grid cells are found. Fig. 4 depicts high change persistent and high change dynamic regions of precipitable water for the year 2012. The color map in Fig. 4 represents the high change persistent ratio and high change dynamic ratio respectively calculated in Algorithm 2.

The pattern in Fig. 4 left shows that the high change persistent regions of precipitable water are primarily located in the mid latitudes with the highest values located on the east coast of North America, South America, and Asia. This pattern coincides with warm, temperate, moist climate regions on land combined with warm ocean currents. This pattern is likely caused by the Coriolis effect of land masses. The pattern of high change dynamic regions shown in Fig. 4 right shows the most dynamic high change regions occurring over areas related to subtropic and temperate ocean climate regions.

Fig. 5 shows the yearly high change dynamic and high change persistent regions for surface level pressure in 2012. The color maps indicate the high change dynamic and high change persistent ratios similar to that shown in Fig. 4. The pattern of high change persistent and dynamic regions of surface level pressure are primarily

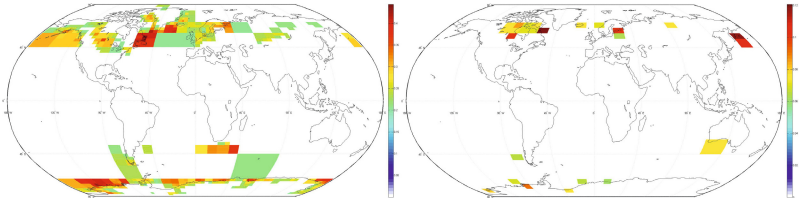


Fig. 5 High change persistent regions (left) and high change dynamic regions (right) for surface level pressure 2012

located in the subpolor regions in the northern and southern hemispheres. These pattern indicate the influence of polar fronts along the boundary between polar and temperate regions.

5 Conclusion and Discussion

The discovery and exploration of patterns in climate data can offer a better understanding of the global climate system. The analysis of spatio-temporal change can offer new insights to the discovery of climate patterns. In this paper, a method was proposed to detect high change persistent and dynamic changing regions in global climate data and was demonstrated on global precipitable water and surface level pressure data for the year 2012. The results show the efficacy of the approach in the detection of persistent and dynamic regions that are verified by known climate regions. The preliminary results suggest that the patterns of dynamic and persistent changing regions can be used to further characterize climate patterns for a period of time.

There are a number of ways in which this research could be extended. First, the approach could be extended to find multi-variate change patterns where overlapping changing regions are analyzed for multiple climate variables. Second, the analysis could be extended to longer time periods and periods of known climate extremes such as El Niño and La Niña. Furthermore, both of these extensions will result in a big data problem and therefore, extending the algorithm to be efficient on big data is another potential area for further research. Finally, this paper is concerned primarily with change over static regions. Another obvious extension would be to include the analysis of moving changing regions.

References

Barros, V., Field, C., Dokke, D., Mastrandrea, M., Mach, K., Bilir, T., Chatterjee, M., Ebi, K., Estrada, Y., Genova, R., et al.: Climate change 2014: impacts, adaptation, and vulnerability. part b: regional aspects. A contribution of working group ii to the fifth assessment report of the intergovernmental panel on climate change (2015)

Faghmous, J., Kumar, V.: Spatio-temporal data mining for climate data: advances, challenges, and opportunities. In: Data Mining and Knowledge Discovery for Big Data, pp. 83–116. Springer (2014)

Shekhar, S., Jiang, Z., Ali, R., Eftelioglu, E., Tang, X., Gunturi, V., Zhou, X.: Spatiotemporal data mining: A computational perspective. ISPRS International Journal of Geo-Information **4**, 2306–2338 (2015)

Zhou, X., Shekhar, S., Ali, R.: Spatiaotemporal change footprint pattern discovery: an interdisciplinary survey. WIREs Data Mining Knowl. Discov. **4** (2014)

Ganguly, A., Steinhaeuser, K.: Data mining for climate change and impacts. In: International Conference on Data Mining Workshops. IEEE (2008)

Hoffman, F., Hargrove, W., Erickson, D., Oglesby, R.: Using clustered climate regimes to analyze and compare predictions from fully coupled general circulation models (2005)

Kitamoto, A.: Spatio-temporal data and mining for typhoon and image collection. Journal of Intelligent Information Systems **19**(1) (2002)
Steinbach, M., Tan, P., Kumar, V., Klooster, S., Potter, C.: Discovery of climate indices using clustering. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 446–455 (2003)
Steinhaeuser, K., Chawla, N., Ganguly, A.: Comparing predictive power in climate data: clustering matters. In: Advances in Spatial and Temporal Databases, pp. 39–55. Springer (2011)
Olaiya, F., Adeyemo, A.: Application of data mining techniques in weather prediction and climate change studies. International Journal of Information Engineering and Electronic Business **4**, 51–59 (2012)
Gunnemann, S., Kremer, H., Laufkotter, C., Seidl, T.: Tracing evolving subspace clusters in temporal climate data. Data Mining and Knowledge Discovery **24**(2), 387–410 (2012)
Rinsurongkawong, V., Eick, C.: Change analysis in spatial datasets by interestingness comparison (2009)
Kleynhans, W., Salmon, B., Wessels, K.: A novel spatiotemporal change detection approach using hyper-temporal satellite data, pp. 4208–4211. IEEE (2014)
Andrienko, G., Malerba, D., May, M., Teisseire, M.: Mining spatiotemporal data. Journal of Intelligent Information Systems **27**, 187–190 (2006)
Pekerskaya, I., Pei, J., Wang, K.: Mining changing regions from access-constrained snapshots: a cluster-embedded decision tree approach. Journal of Intelligent Information Systems **27**, 215–242 (2006)
McGuire, M., Janeja, V., Gangopadhyay, A.: Mining trajectories of moving dynamic spatio-temporal regions in sensor datasets. Data Mining and Knowledge Discovery **28**(4), 961–1003 (2014)
Samet, H.: The quadtree and related hierarchical data structures (1984)
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The ncep/ncar 40-year reanalysis project. Bulletin of the American Meteorological Society (1996)

Open Source Data Quality Tools: Revisited

Venkata Sai Venkatesh Pulla, Cihan Varol and Murat Al

Abstract High data quality is defined as the reliability and application efficiency of data present in a system. Maintaining high data quality has become a key feature for most organizations. Different data quality tools are used for extracting, cleaning, and matching data sources. In this paper, we first introduce state of the art open source data quality tools, specifically Talend Open Studio, DataCleaner, WinPure, Data Preparator, Data Match, DataMartist, Pentaho Kettle, SQL Power Architect, SQL Power DQguru, and DQ Analyzer. Secondly, we compare these tools based on their key features and performance in data profiling, integration, and cleaning. Overall, DataCleaner scores highest among the considered tools.

Keywords Data cleaning · Data integration · Data profiling · Data quality · Data quality tools

1 Introduction

Much of modern life is reliant on computers. With the rapid growth of technology and usage of computers in various fields, large amounts of data are generated day-by-day. In the meantime, losing important information from the continuously generated big data has become an emerging problem for many organizations. However, in an ideal scenario, the data present in the system should be reliable, useful, and complete. Data quality tools can be used to gain such high quality data.

In an early step, collected data may need to be reduced to relevant data by organizing and removing unwanted elements. Extracting key information, finding

V.S. Venkatesh Pulla · Cihan Varol(✉)
Department of Computer Science,
Sam Houston State University, Huntsville, TX 77341, USA
e-mail: {vvp001,cvarol}@shsu.edu

M. Al
School of Computing, University of North Florida, Jacksonville, FL 32224, USA
e-mail: muratal.usa@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_77

patterns, analyzing tables, and managing metadata are some of the other key aspects in maintaining high data quality. The many features of the data quality tools are categorized into three core functionality areas, namely data profiling, data integration, and data cleansing. Tables 2, 3, and 4 present the comparison of the investigated tools in regards to these core functionalities.

In this paper, we re-visit the work of Pushkarev, Neumann, Varol, and Talburt [1] by evaluating and comparing ten state of the art data quality tools, specifically — (*Talend Open Studio, DataCleaner, WinPure, Data Preparator, Data Match, DataMartist, Pentaho Kettle, SQL Power Architect, SQL Power DQguru* and *DQAnalyzer*). The comparison is done based on key features such as user interface, dataset management, and performances in data access, data propagation, metadata management, data extraction, data transformation, parsing and standardization, data de-duping, pattern discovery, table structure analysis, frequency analysis, and domain analysis.

The organization of the rest of the paper is as follows. In the Background section, research conducted on data quality and factors leading to poor quality are discussed. Methodology section explains the data quality tools considered and their evaluation criteria. Results section provides the performance of the tools in detail. Finally, the Conclusion section discusses the findings from the work.

2 Background

As generated data has become an important aspect in many organizations, maintaining high data quality has become a key task. From pizza chains over web services to NASA space shuttle, organizations have suffered for not maintaining high data quality [1]. For example, low data quality has cost more than $600 billion for excess postage [1]. The complexity of addressing data misinterpretation and improving data quality rises when composing multiple services developed by independent providers that are distributed throughout the world [2].

Data integration concept is becoming increasingly important as organizations strive to integrate an increasing quantity of external and internal data [3]. The success of data integration, among other aspects, is dependent on data quality. Lack of accuracy and credibility in statistical data endanger statistical work and influence the reasonable understanding of the real situation, making it more difficult to formulate corresponding policies including in data integration [4].

Overall, with the advent of big data, data quality management has become more important than ever. Without proper data quality management, even minor errors can result in revenue loss, process inefficiency, and failure to comply with industry and government regulations [5]. The data quality can be improved if there is an adequate knowledge about data, right from the beginning of the production process [6]. Moreover, rather than simply changing data values manually, processes, specifically data quality tools, should be adopted to increase its quality automatically.

There are four facets of data quality. These are *data profiling, data integration, data cleansing,* and *data monitoring*. *Data profiling's* main purpose is to explore the

relationships among values across datasets, providing frequency distribution of the different values present in the tables and to check for anomalies that exist in datasets [1]. All data profiling tools should have features that can satisfy the mentioned requirements. Analysis features, such as value frequency analysis, cross-table analysis, domain analysis, etc. are categorized under profiling. Pattern discovery is another key aspect of profiling, which is useful for data mining purposes.

Data integration involves combining data residing in different sources and providing users with a unified view of data [1]. Some of the features that are categorized under data integration are data extraction and transformation, data propagation, managing metadata, and data access. Data extraction and transformation are vital to consolidate data in one common format.

Data cleansing is generally used for removing unnecessary data present in the tables or files. Parsing and standardization, data transformation, record matching and identity resolution, and data de-duping are some of the features that are categorized under data cleansing. Achieving high data quality via cleansing will yield to more accurate outcomes for data related activities such as data mining, prediction, etc.

Data monitoring's purpose is to enforce data quality standards and business rules across the enterprise. It can be used to maintain compliance with either internal best practices or regulatory imperatives. Data monitoring provides the assurance that once data problems are fixed, they will remain within tight limits. When data gets out of control, users know immediately and can address problems before the quality of the data declines.

These four features play an important role in improving data quality. Hence, special importance is given to them when developing or selecting data quality tools.

3 Methodology

This work focuses on three aspects of data quality, namely data profiling, data integration, and data cleaning. Since the data monitoring feature is not provided by any of the tools presented in this paper, it will not be part of the analysis. In addition to the latest versions of the data quality tools evaluated by [1], new tools have been added to this study. This paper analyzes data quality tools, namely, Talend Open Studio v6.0, DataCleaner v4.0.9, WinPure v6.2.8, Data Preparator v1.7, Data Match v4.3, DataMartist v1.7.1, Pentaho Kettle v5.4, SQL Power Architect v1.0.7, SQL Power DQguru v0.9.7, and DQ Analyzer v8.0. Each tool is elucidated concisely as below.

Talend Open Studio is an open source data quality tool used for data profiling and integration, which has the capability of maintaining a metadata repository [7]. It is developed based on Java and Perl; and the SQL connectivity is used for connecting almost all the data sources available. The tool has an eclipse-based graphical workspace, which enables drag and drop functionality. It facilitates an easy access to major databases, applications, and accepts different file formats as input. Different tasks, such as statistical profiling, analysis of text and numeric

fields, and validation against standard patterns can be performed. A variety of tables, datasets, and graphs can be generated.

The latest version of this tool has included features for both data profiling and data integration. Data profiling tasks including connection analysis, catalog analysis, schema analysis, table analysis, and redundancy analysis can be conducted effortlessly. Business modeler can be used for building business-oriented views using the predefined shapes available in the tool. *Talend Open Studio* is a metadata based tool. Metadata manager is used for organizing metadata present in all the modules, which in turn is helpful for the integration process.

Data Cleaner is another open source data quality tool used for achieving data profiling, data cleaning, and data integration. The tool has an excellent graphical interface and contains many features. The drag and drop component enhances the usability. The transformation function provides an extensive number of features such as composition, conversion, filtering specific data, matching and standardization, finding patterns, etc.

The generated data can be exported to different formats such as csv, text, and others. The results can also be exported to html and can be posted in a server. The tool comes with a built in csv and a database file for demonstration purposes. Additionally, it includes some web-based features and a user manual. The tool is developed in Java and the JDBC connectivity is used for connecting the data [8].

WinPure is an open source data quality tool used for data cleaning [9]. Its interface is very simple and consists of features related to data cleaning. It includes well-written tutorials and a start guide, which most tools do not provide. Essentially, there are three components in this tool, namely, data, clean, and match. The data component is used for importing data into the tool and creating and opening projects. Furthermore, the imported data can be edited in this section. The clean component consists of a wide range of data cleansing tools that help to populate, standardize, and correct the lists present in a database. Match component is a data matching module that is used to identify, remove, and merge all duplications present in a database.

The tool supports different file formats such as csv, xls, and txt. There are different features that support the cleaning of files, which include statistics, text cleaner, case converter, word manager, column splitter, column merger, and email cleaner. Each feature in turn contains many sub-features. For example, text cleaner contains additional features including numeric only column cleaning, removing dots, commas, spaces, and digits separately. The tool is developed in Java and uses JDBC connectivity for connecting to data sources. Usually, the output is generated in the form of tables or datasets and can be exported into different file formats.

Data Preparator is an open source data quality tool used to assist with the common tasks of data preprocessing in data analysis and data mining [10]. It has a very simple interface. It provides a variety of techniques for data cleaning, transformation, and exploration. For example, the tool assists in exploring and preparing data, among others using the operators cleaning, numeration, scaling, attribute selection, statistics, visualization, balancing. In addition, the tool allows the chaining of preprocessing operators into a flow graph.

The availability of a plethora of output formats is a strong characteristic of this tool. The output tab of the tool lists various options, such as table, file database, and Excel. Output can also be presented in graphical forms, such as pie charts, histograms, sequences, and scatter plots. The tool was developed in Java.

Data Match is an open source data quality tool used for data cleaning [11]. It is useful for data de-duplication, standardizing, and fuzzy matching. It is mainly a data cleaning tool that removes unwanted information present in files, disks, etc. Its interface is very simple, which consists of opening projects and adding files and cleaning them. Supported encoding formats include ASCII, Unicode, UTF32, UTF7, and UTF8. Once a file is added, different tabs become available in the software. Some of these are data profile, clean, match, and word definition.

The clean tab offers many sub-features for data cleaning, such as removing spaces, removing letters, removing numbers, and replacing characters and numbers separately. The data profile tab is used for obtaining statistics related to the file or data. These statistics can be depicted in graphical form using pie charts, bar charts, and other visual presentation techniques. The tool is developed in Java and can be used to connect to any kind of data source available.

DataMartist is an open source data quality tool used for visual data profiling and transformation [12]. It has a typical user interface consisting of data sources, block library, canvasses, and output. All the data added to the tool will be categorized under data sources. Block library contains all of the important features for data profiling and transformation. Filter, summarize, segment, de-duplicate, and categorize are some of the important features. There is an easy drag and drop option into the main screen that makes the tool very easy to operate.

Most of the tool is designed to show the connections graphically. For example, if a particular project consisting of a database is added to the tool, the connections between all the tables are shown in a graphical way. Once the project is established, different features can be executed on the data. However, the tool cannot create an output using graphical formats; instead, the output can be generated in text or csv format delimited by a specific delimiter. The tool is developed in Java and can connect a variety of data sources.

Pentaho Kettle is a data integration tool that uses the extraction, transformation, and loading techniques. Kettle is a recursive acronym for "Kettle Extraction, Transport, Transformation and Loading Environment" [1]. This tool is metadata-oriented and supports a vast array of input and output formats including text files, data sheets, and commercial and free database engines. There are four major sub-applications in Pentaho Kettle data integration, namely, Spoon, Pan, Kitchen, and Carte. Spoon is a graphical user interface that allows the design of transformations and jobs that can be run with the Kettle tools – Pan and Kitchen. Pan is a data transformation engine that performs common functions such as reading, manipulating, and writing data to and from various data sources. Kitchen is a program that executes jobs designed by spoon in xml or in a database repository. Carte is a simple web server that allows users to execute transformations and jobs [13]. The new version 5.4 supports Hadoop configurations MapR 4.0.1, HDP 2.2, and CDH 5.3 [14].

SQL Power Architect is a data modeling tool offering many unique features specifically for the data warehouse architect [15]. It allows users to reverse-engineer existing databases, perform data profiling on source databases, and auto-generate ETL metadata. Source database connectivity is established using JDBC. It can connect to multiple source databases concurrently, as well as compare data models and database structures and identify discrepancies. It remembers the origin of each column, allowing it to generate source-to-target visual mapping reports. Source data structure snapshots can be stored in the project itself.

SQL Power DQguru is a data cleaning tool used to clean data, validate and correct addresses, identify and remove duplicates, and build cross-references between source and target tables [16]. It works with all major databases, e.g. Oracle, MySQL, and Sybase. Data de-duping can be achieved on input data based on the rules specified by the user. There are different transformation and matching functions available in the tool. Some of them are concatenation, case conversion, string substitution, space trimming, word sorting, and address correction. Phonetic coding for words such as Soundex, Metaphone and Double Metaphone are also available. The tool is developed in Java and generates output in both graphical and table or dataset formats.

DQ Analyzer is a data profiling tool developed by Ataccama [17]. In addition to profiling, it can perform different types of analysis such as standard analysis, domain analysis, and mask analysis. After profiling, different statistics become available under the various tabs. The Basic tab contains general summary statistics. Frequency tab shows the number of times a value occurs in the data. Domain tab lists the domains to which the data belongs. Mask tab shows the pattern analysis of the data. Quantiles tab shows slices of data when placed in order. Finally, the Groups tab shows the number of times each frequency count is repeated. It generates output in both the table and graphical format.

This paper considers following features and characteristics for the evaluation of the tools, namely, data source connectivity, data set management, metadata repository, interface, development platform, and report creation. Data source connectivity is used to evaluate the compatibility of the tool with different available data sources. Data set management considers the techniques for extracting and managing data. The evaluation of metadata repository determines whether the tool has a designated location for storing the generated metadata as well as how it is managed inside the tool. Interface is one of the key features that are used for the comparison of the tools, with particular importance given to graphical representation capabilities of the tool. Furthermore, certain criteria such as learning curve, usability, speed, consistency, and aesthetic design are considered under interface. Development platform refers to the programming language in which the application is developed. Report creation refers to the format in which the output is generated under different circumstances. Tables and data sources as well as graphical representations are the two different output formats. All the core functionality related features including table structure analysis, domain analysis, data extraction, and data transformation fit into the categories of data profiling, integration, and cleaning.

4 Results

Talend Open Studio has a decent graphical interface and connects to all data sources, as reflected in Table 1. It can generate both graphical and table or file format output. Its biggest use is in the area of data profiling where different analyses can be performed on the available data in the database. *Data Cleaner* has an excellent interface with numerous features. It can generate outputs in table or file formats. It belongs to four of the ten tools investigated that can perform all of the core functionalities. It is the best performing tool in data profiling, as shown in Table 2. *WinPure*, *Data Match,* and *SQL Power DQguru* are designed for data cleaning, which have satisfying performance in terms of parsing and standardization, data transformation, record matching, identity resolution, and data de-duping. While *WinPure* and *SQL Power DQguru* share the first place in Table 3, *WinPure* has a better user interface.

Data Preparator is one of the four tools that cover all of the core functionalities. However, its feature set does not stand out in any of the areas depicted in the Tables 2 through 4. *Data Match* is one of the few tools that offer connectivity of all data sources. Its strength is in data cleaning, where it offers three of the four features, as shown in Table 4. *DataMartist* is similar to *Data Preparator* in terms of functionalities. While it offers features in all the categories, it is not a strong contender in any of them. *Pentaho Kettle* does not offer any of the functionalities in data profiling, however, it clearly dominates the core functionality area of data integration, as shown in Table 3. It is the only tool that offers all features in this category. *SQL Power Architect* makes second place in Table 2 and Table 3. Its feature set, however, is limited to these core functionalities since it offers none of the functions making up the core area of data cleaning. The output can be easily distinguished by the graphical representation present in the tool.

Table 1 Performance Criteria

Tools	Data Source Connectivity	Dataset Mng.	Metadata Repository	GUI	Development Platform	Reporting
Talend Open Studio	All	ELT	Y	G	Perl/Java	T/Gr
Data Cleaner	JDBC	ETL	Y	G	Java	T/Gr
WinPure	JDBC	ETL	N	G	Java	T
Data Preparator	N	ELT	N	G	Java	T/Gr
Data Match	All	ELT	N	G	Java	T/Gr
DataMartist	All	ETL	N	G	Java	T
Pentaho Kettle	JDBC	ETL	Y	G	Java	T
SQL Power Architect	JDBC	ETL	Y	G	Java	T/Gr
SQL Power DQguru	JDBC	ELT	Y	G	Java	T/Gr
DQ Analyzer	JDBC	N	Y	G	Java	T/Gr

(All) Ability to connect to different data sources. (ELT) Extract, Load and Transform. (ETL) Extract, Transform and Load. (Y) Yes. (N) No. (G) Graphical user Interface. (T) Report can be generated in table or data sets. (Gr) Report can be generated in graphical format.

Table 2 Core Functionality (Data Profiling)

Tools	Column Value Frequency Analysis & Statistics	Table Structure Analysis	Domain Analysis	Cross Table Redundancy Analysis	Data Pattern Discovery	Drill-through Analysis
Talend Open Studio	Y	Y	N	Y	N	N
Data Cleaner	Y	N	Y	Y	Y	Y
WinPure	N	N	N	N	N	N
Data Preparator	Y	Y	N	N	N	N
Data Match	N	N	N	N	N	N
DataMartist	N	Y	N	N	Y	N
Pentaho Kettle	N	N	N	N	N	N
SQL Power Architect	Y	Y	N	Y	N	N
SQL Power DQguru	N	N	N	N	N	N
DQ Analyzer	Y	N	Y	N	Y	N

SQL Power DQguru has a limited feature set. It does not support any of the features in Table 2 and Table 3. However, it offers all features required for data cleaning and dominates Table 4 along with *WinPure*. *DQ Analyzer* is even more limited since it offers features in only one of the core areas, namely, data profiling. From the total of 15 features distributed across the Tables 2, 3, and 4, it provides only three, which can be seen in Table 2. In addition, Table 1 lists it as the only tool lacking the dataset management feature.

Table 3 Core Functionality (Data Integration)

Tools	Data Extraction, Transformation, & Consolidation	Metadata Manage	Data Propagation	Business Modeler & Job Designer	Data Access
Talend Open Studio	Y	Y	Y	N	Y
Data Cleaner	Y	Y	Y	N	N
WinPure	Y	N	Y	N	Y
Data Preparator	Y	N	Y	N	N
Data Match	Y	N	Y	N	Y
DataMartist	Y	N	Y	N	Y
Pentaho Kettle	Y	Y	Y	Y	Y
SQL Power Architect	Y	Y	Y	N	Y
SQL Power DQguru	N	N	N	N	N
DQ Analyzer	N	N	N	N	N

Table 4 Core Functionality (Data Cleaning)

Tools	Parsing and Standardization	Data Transformation	Record Matching and Identity Solution	Data de-duping
Talend Open Studio	Y	Y	N	N
Data Cleaner	Y	Y	N	N
WinPure	Y	Y	Y	Y
Data Preparator	N	Y	N	N
Data Match	Y	Y	N	Y
DataMartist	N	Y	N	N
Pentaho Kettle	Y	Y	N	N
SQL Power Architect	N	N	N	N
SQL Power DQguru	Y	Y	Y	Y
DQ Analyzer	N	N	N	N

5 Conclusion

Collection of digital data by organizations is in the rise. This data may be related to financial reports, user behavior, or product specifications; the areas are endless. Maintaining quality of such data and extracting useful information from it has become crucial. To accomplish this and support their decision making process, organizations make use of data quality tools. However, most of the available tools are very expensive; ultimately increasing the budgets of the projects as data quality has to be maintained on daily basis.

At this tick, open source/free data quality tools are given the principal priority in achieving data quality. These open source/free data quality tools have distinct features for conducting data profiling, data integration, and data cleaning. Each tool mentioned in the paper has its own advantages and features. All the tools have a decent user interface that facilitates the user with features necessary for maintaining data quality.

In contrast to some commercially available expensive solutions, the tools evaluated here may not fully satiate to address the problems of data quality, since they may not offer the wide range of sub-features within data profiling, data integration, and data cleansing. However, in most cases, they are sufficient for maintaining the data quality at an acceptable level. Some of the tools are splendid in terms of data cleaning, allowing organizations to clean terabytes of data generated in the system daily. Some tools have metadata repository for maintaining metadata, which improves performance and is useful for forensic purposes.

These open source data quality tools are especially beneficial for projects with low budget. While most of today's businesses have a use for data quality tools, the number of small businesses that cannot afford high-priced tools is by far larger than corporations that can afford them. Most of the tools have been developed concentrating on individual features.

The salient characteristics of the tools have been discussed in the results section. The results show, with ten out of 15 features, *Data Cleaner* offers the broadest feature set across the three core functionality areas. With nine features, *Talend Open Studio* ranks as the second most versatile tool. The third place is shared by *WinPure*, *Pentaho Kettle*, and *SQL Power Architect*, each with seven features, respectively. *DQ Analyzer* and *SQL Power DQguru* offer the least amount of features. While *SQL Power DQguru* may not qualify as an all-round tool, along with *WinPure* it achieves a perfect score in data cleaning.

Depending on the tasks at hand, one could employ different tools, i.e. the best suitable tool for the task. However, this would require learning each of the used tools. Therefore, an all-round tool that consolidates as many functions as possible would be desirable. A combination of *Talend Open Studio*, *Data Cleaner*, and *WinPure* may convince as a model tool for requisite features.

References

1. Pushkarev, V., Neumann, H., Varol, C., Talburt, C.: An overview of Open source data quality tools. In: International Conference on Information & Knowledge Engineering (IKE), Las Vegas (2010)
2. Li, X., Madnick, S., Zhu, H., Fan, Y.: Improving Data Quality for Web Services Composition. In: 10th International Workshop on Quality in Databases (QDB), Trento, Italy (2009)
3. Lin, M., Hua, Z.: A Method for Measuring Data Quality in Data Integration. In: International Seminar on Future Information Technology and Management Engineering, Leicestershire, UK (2008)
4. Liu, H.: Analysis of Statistical Data Quality. In: Seventh International Joint Conference on Computational Sciences and Optimization, Beijing, China (2014)
5. Saha, B., Srivastava, D.: Data Quality: The other Face of Big Data. In: IEEE 30th International Conference on Data Engineering (ICDE), Chicago (2014)
6. Glowalla, P., Balazy, P., Basten, D., Sunyaev, A.: Process-driven Data Quality management-An Application of the combined conceptual Life Cycle Model. In: 47th Hawaii International conference on System Science, Hawaii (2014)
7. Talend Open Studio. https://www.talend.com/resource/data-quality-tools.html
8. Data Cleaner. http://datacleaner.org/docs
9. WinPure clean & Match. http://www.winpure.com/cleanmatch.html
10. Data Preparator. http://www.datapreparator.com/
11. Data Match. http://dataladder.com/data-matching-software/
12. DataMartist. http://www.datamartist.com/
13. Pentaho Kettle. http://wiki.pentaho.com/display/EAI/Latest+Pentaho+Data+Integration+%28aka+Kettle%29+Documentation
14. Pentaho Kettle. https://help.pentaho.com/Documentation/5.4/0T0/040
15. SQL power architect. http://www.sqlpower.ca/page/architect
16. SQL Power DQguru. http://www.sqlpower.ca/page/dqguru
17. DQ Analyzer. https://www.ataccama.com/products/dq-analyzer

Part IX
New Trends in Wavelet and Numerical Analysis

Some Outflow Boundary Conditions for the Navier-Stokes Equations

Yoshiki Sugitani, Guanyu Zhou and Norikazu Saito

Abstract In numerical simulation of real-world flow problems, we often encounter some issues concerning artificial boundary conditions. An important example is the blood flow problem in the large arteries and the simulation is highly dependent on the choice of artificial boundary conditions posed on the outflow boundary. In the present paper, we examine some outflow boundary conditions including our new condition from the view-point of mathematical analysis.

Keywords Outflow boundary condition · Stokes equations · Finite element approximation · Numerical simulation

1 Introduction

The Navier-Stokes (NS) equations

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = \nabla \cdot \sigma(u, p) + f, \quad \nabla \cdot \mathbf{u} = 0 \tag{1}$$

are well-known as the equations of motion of viscous incompressible fluid. Here, $\mathbf{u} = (u_1, \ldots, u_d)$ denotes the velocity and p denotes the pressure with the density $\rho = 1$ and the kinetic viscosity ν . $\sigma(u, p)$ is the stress tensor defined by $\sigma(u, p) = -pI + 2\nu D(u)$ with the identity matrix I and the deformation-rate tensor $D(u) = \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^T\right)$. Therein, $\tau(u, p) = \sigma(u, p) \cdot n$ denotes the traction vector, where n is the outward normal vector. Finally, f is outer force to fluid.

Y. Sugitani(✉) · G. Zhou · N. Saito
Graduate School of Mathematical Sciences, The University of Tokyo,
3-8-1 Komaba, Meguro-ku, Tokyo, Japan
e-mail: sugitani@ms.u-tokyo.ac.jp

This work is supported by JST, CREST, and JSPS KAKENHI Grant Number 15H03635, 15K13454.

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_78

NS equations are one of the most standard mathematical modeling in fluid problems. For example, in numerical simulations of blood flow at large arteries, one needs to solve (1) over a branched pipe domain Ω such as Fig. 1. The boundary of Ω is

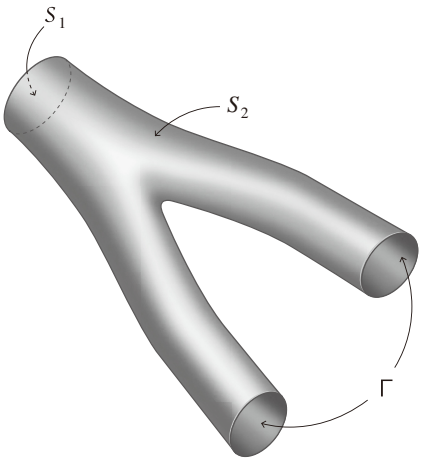


Fig. 1 An example of blood vessels.

composed of three parts S_1, S_2 and Γ, which stand for the inflow boundary, the wall, and the outflow boundary, respectively. When you want to know the velocity and pressure profiles in Ω with the inflow $\mathbf{b}$ on S_1, you can impose the velocity profiles as boundary conditions on S_1 and S_2 (i.e., $\mathbf{u} = \mathbf{b}$ on S_1, $\mathbf{u} = \mathbf{0}$ on S_2). On the other hand the choice of the outflow boundary conditions is nontrivial. In such a case, one of the most common outflow boundary conditions (BCs) is the free-traction BC

$$\tau(u, p) = 0 \text{ on } \Gamma. \tag{2}$$

This is also called the do-nothing BC for ease implementation. Under this outflow BC, however, you may not obtain the expecting flow distribution because of the backward flow at one branch. Moreover, the do-nothing BC has a disadvantage of lack of the energy inequality. To describe it more specifically, we take a reference flow (g, π) which is the solution of the Stokes system

$$\nabla \cdot \sigma(g, \pi) = 0, \quad \nabla \cdot g = 0 \quad \text{in } \Omega, \tag{3a}$$

$$g = b \text{ on } S_1, \quad g = 0 \text{ on } S_2, \tag{3b}$$

for all $t \in [0, T]$. g is nothing but a lifting function of $\mathbf{b}$. By using this, we will find (v, q) instead of (u, p) of the form $u = v + g, \quad p = q + \pi$.

The problem holds the energy inequality if you have

$$\sup_{t\in[0,T]} \|v(t)\|^2_{L^2(\Omega)^d} + 2\nu \int_0^T \int_\Omega D_{ij}(v) : D_{ij}(v) \le C, \tag{4}$$

where C is a positive constant depending only on $f, \mathbf{b}, T$ and the initial condition v_0. This inequality is of use because it plays a crucial role in the construction of a solution of the NS equations. Moreover, it is connected with the stability of numerical schemes from the viewpoint of numerical computation. The do-nothing BC, unfortunately, does not hold the energy inequality. For these kind of reasons, a proper setting of the outflow BC at artificial boundaries is one of the main issues in simulations of blood flow problem and studied so far.

This paper is composed of 6 sections. Section 2 is devoted to presentation of several outflow BCs including our new idea named the unilateral BC. In section 3, we state a variational formulation and well-posedness for a model unilateral boundary value problem. In section 4 we consider the finite element approximation for our problem. The well-posedness and error estimates are proved. We confirm our results by numerical experiments in section 5. Finally in section 6, we remark a relationship between our proposed unilateral BC and another nonlinear outflow BC. All mathematical results will be stated without the proofs; most of the complete proofs are shown in Saito et al. [6].

2 Various Outflow Boundary Conditions

In 2002, Formaggia et al. [3] proposed a method to obtain the velocity satisfying

$$\int_{\Gamma_i} \mathbf{u} \cdot \mathbf{n} \, d\Gamma_i = Q_j \quad (j = 1, 2) \tag{5}$$

for given flux Q_j at each branch Γ_j. Here $u_n = u \cdot n$ denotes the normal component of a vector while $u_T = u - (u \cdot n)n$ denotes the tangential one. Vignon-Clementel et al. [6] also proposed another outflow BC named the resistance BC

$$\tau_n + p_0 + R_j \int_{\Gamma_j} u_n \, d\Gamma_j = 0, \quad \tau_T = 0 \text{ on } \Gamma. \tag{6}$$

for given resistance of each branch R_j and a level pressure p_0. These outflow BCs are proposed in order to controll each flux at artificial boundaries Γ_i.

Of course the energy preserving outflow BCs are studied. For example, among some of nonlinear outflow BCs proposed by Bruneau and Fabrie [1] to hold the energy inequality (4), the typical one is

$$\tau(u, p) = -\frac{1}{2}[u_n]_- v + \tau(g, \pi) \text{ on } \Gamma, \tag{7}$$

where $[s]_\pm = \max\{0, \pm s\}, \quad s = [s]_+ - [s]_- \quad (s \in R)$.

Bazilevs et al. [2] proposed another nonlinear energy preserving outflow BC by introducing the regularized traction

$$\tilde{\tau}(u, p) = \tau(u, p) + [u_n]_- \mathbf{u}. \tag{8}$$

and by imposing

$$\tilde{\tau}_n + p_0 + R_j \int_{\Gamma_j} u_n \, d\Gamma_j = 0, \quad \tilde{\tau}_T = 0 \text{ on } \Gamma. \tag{9}$$

This is nothing but the combination of (6) and (8). If we employ these nonlinear BCs we need to calculate the reference flow for computation.

We recently proposed a new outflow BC named a unilateral BC

$$\begin{cases} u_n \geq 0, \quad \tau_n(u, p) \geq 0 \\ u_n \tau_n(u, p) = 0, \quad \tau_T(u) = 0 \end{cases} \quad \text{on } \Gamma \tag{10}$$

without any reference flow. This is indeed a generalization of the free-traction condition (2). Namely,

$$\text{if } u_n > 0 \text{ on } \omega \subset \Gamma, \text{ then } \tau_n(u, p) = 0 \text{ on } \omega; \tag{11}$$

$$\text{if } u_n = 0 \text{ on } \omega \subset \Gamma, \text{ then } \tau_n(u, p) \geq 0 \text{ on } \omega. \tag{12}$$

An advantage of employing (10) is that (1) admits energy inequality (4), while the condition (10) is described in terms of inequalities so that it does not fit numerical calculations. However, we can utilize its penalty approximation

$$\tau_n = \frac{1}{\varepsilon}[u_n]_-, \quad \tau_T = 0 \text{ on } \Gamma \tag{13}$$

with $0 < \varepsilon << 1$. We also obtain energy inequality with (13) for a sufficiently small ε. Since the condition (13) is a kind of nonlinear Robin BC, after introducing a C^1 regularization of $[\cdot]_-$, we can solve (1) with (13) by Newton's iteration.

3 Model Unilateral Boundary Value Problem

Our final aim is to develop the theory for the initial-boundary value problems for the Navier-Stokes equations (1) with (10) or with (13) from the standpoint both of analysis and numerical computations. In this paper, as a primary step, we report the

well-posedness and the finite element application of a model problem for the Stokes equations.

Let us consider the Stokes equation with (10). Thanks to the reference flow $u = v + g \quad p = q + \pi$, we will change variables (still denoted by (u, p)). That is, finding the velocity u and the pressure p such that

$$
\begin{aligned}
&-\nu\Delta u + \nabla p = f, \quad \nabla \cdot u = 0 && \text{in } \Omega, && (14a)\\
&u = 0 && \text{on } S_1 \cup S_2, && (14b)\\
&u_n + g_n \geq 0, && \text{on } \Gamma, && (14c)\\
&\tau_n(u, p) + \alpha_n \geq 0 && \text{on } \Gamma, && (14d)\\
&(u_n + g_n)(\tau_n(u, p) + \alpha_n) = 0 && \text{on } \Gamma, && (14e)\\
&\tau_T(u) + \alpha_T = 0 && \text{on } \Gamma && (14f)
\end{aligned}
$$

where $\alpha = \tau(g, \pi)$.

Remark 1. As a matter of fact, (14) themselves are not new problems. In a classical monograph, Kikuchi and Oden [5], Chapter 7 is devoted to similar problems. However, their problem contains the traction condition $\tau(u, p) = h$. More precisely, they suppose that S_2 is divided into two parts S_{21}, S_{22} and consider

$$u = 0 \text{ on } S_{21}, \quad \tau_T(u) = 0 \text{ on } S_{22} \tag{15}$$

instead of (14b). Then you can prove an inequality called the coupled Babuska-Brezzi condition and the well-posedness and error estimates of the corresponding penalty problem are direct consequences of this result. In contrast, we are interested in establishing a formulation without the traction boundary condition so that we have to develop a totally new method of analysis in this paper.

Before the weak formulation of (14), we introduce some function spaces. We use the standard Lebesgue and Sobolev spaces like $L^2(\Omega)$, $H^1(\Omega)$.The basic function spaces of our consideration are

$$V = \{v \in H^1(\Omega)^d \mid v = 0 \text{ on} S_1 \cup S_2\} \text{ and } Q = L^2(\Omega). \tag{16}$$

They are Hilbert spaces equipped with the norms $\|v\|_1$ and $\|q\|$, respectively. Closed convex subset

$$K = \{v \in V \mid v_n + g_n \geq 0 \text{ on } \Gamma\} \tag{17}$$

also plays an important role. Finally we introduce so-called Lions-Magenes space $M = H_{00}^{\frac{1}{2}}(\Gamma)$, which is actually the trace space of V (see [4, §11.5,Ch.1]).

From now on, we always assume

$$f \in Q^d, \quad b \in M^d, \quad \beta \equiv -\int_{S_1} b \cdot n > 0. \tag{18}$$

Then the weak formulation of our problem is written as;
(PDE) Find $(u, p) \in V \times Q$ such that

$$a(u, v) + b(p, v) = \int_\Omega f \cdot v \, dx \qquad (\forall v \in H_0^1(\Omega)^d), \tag{19a}$$

$$b(q, u) = 0 \qquad (\forall q \in Q), \tag{19b}$$

$$u_n + g_n \geq 0 \qquad \text{a.e. on } \Gamma, \tag{19c}$$

$$[\tau_n(u, p) + \alpha_n, \mu] \geq 0 \qquad (\forall \mu \in M, \ \mu \geq 0), \tag{19d}$$

$$[\tau_n(u, p) + \alpha_n, u_n + g_n] = 0 \tag{19e}$$

$$[[\tau_T(u) + \alpha_T, \mu]] = 0 \qquad (\forall \mu \in M^d, \ \mu_n = 0). \tag{19f}$$

where $[\cdot, \cdot]$ denote the duality pairing over M (so does $[[\cdot, \cdot]]$ over M^d), and

$$a(u, v) = 2\nu \int_\Omega D(u) : D(v) \, dx, \quad b(p, u) = -\int_\Omega p(\nabla \cdot u) \, dx. \tag{20}$$

Problem (PDE) is equivalent to the following variational inequality.
(VI) Find $(u, p) \in K \times Q$ such that

$$a(u, v - u) + b(p, v - u) \geq (f, v - u) - [[\alpha, v - u]] \qquad (\forall v \in K), \tag{21a}$$

$$b(q, u) = 0 \qquad (\forall q \in Q). \tag{21b}$$

Theorem 1. *Problems (VI) and (PDE) are equivalent.*

Theorem 2. *There exists a unique solution $(u, p) \in K \times Q$ of (VI) and it holds*

$$\|u\|_1 + \|p\| \leq C \tag{22}$$

where C is a positive constant depending only on f, g, α and Ω.

4 The Finite Element Approximation

As mentioned above, since the problem (PDE) and (VI) are not directly applicable for numerical computation, we propose the penalty approximation for (VI). Let us consider the Stokes equation with (13), that is, finding the velocity u and the pressure p such that

$$-\nu\Delta u + \nabla p = f, \quad \nabla \cdot u = 0 \qquad \text{in } \Omega, \tag{23a}$$

$$u = 0 \qquad \text{on } S_1 \cup S_2, \tag{23b}$$

$$\tau_n(u, p) + \alpha_n = \frac{1}{\varepsilon}\phi_\delta(u_n + g_n), \qquad \text{on } \Gamma, \tag{23c}$$

$$\tau_T(u, p) + \alpha_T = 0, \qquad \text{on } \Gamma, \tag{23d}$$

where ϕ_δ is C^1 regularization to $[\cdot]_-$ defined as

$$\phi_\delta(s) = \begin{cases} 0 & (s \geq 0) \\ (\sqrt{s^2 + \delta^2} - \delta) & (s < 0). \end{cases} \tag{24}$$

For penalty parameter $\varepsilon \in (0, 1]$, weak formulation of this problem reads as;
(PE$_{\varepsilon,\delta}$) Find $(u, p) \in V \times Q$ such that

$$a(u, v) + b(p, v) - \frac{1}{\varepsilon}\int_\Gamma \phi_\delta(u_n + g_n)v_n \, d\Gamma = (f, v) - [[\alpha, v]] \quad (\forall v \in V), \tag{25a}$$

$$b(q, u) = 0 \qquad (\forall q \in Q). \tag{25b}$$

Now let us discretize (PE$_{\varepsilon,\delta}$) by the finite element approximation. In order to avoid unessential difficulties concerning "curved boundary", we consider only the case of polyhedral (polygonal) domain. Consequently, the unit outer normal vector n to Γ is a constant vector over Γ. We use so-called MINI (P1 bubble/P1) elements for discretization. Let $\{\mathcal{T}_h\}_h$ be a regular family of triangulations of Ω. As the granularity parameter, we have employed $h = \max\{h_T \mid T \in \mathcal{T}_h\}$, where h_T denotes the diameter of each triangle T. We introduce the following function spaces:

$$V_h = \{v_h \in C^0(\overline{\Omega})^d \mid v_h = 0 \text{ on } S,\ v_h|_T \in \mathcal{P}_1^{(d)} \oplus \text{span}\{\varphi_T\}\ (\forall T \in \mathcal{T}_h)\},$$
$$Q_h = \{q_h \in C^0(\overline{\Omega}) \mid q_h|_T \in \mathcal{P}_1^{(d)}\ (\forall T \in \mathcal{T}_h)\},$$
$$M_h = \{\mu_h = v_{hn}|_\Gamma \mid v_h \in V_h\}.$$

where $\mathcal{P}_k^{(d)}$ denotes the set of all polynomials in $x_1, \ldots, x_d$ of degree $\leq k$, and $\varphi_T = \prod_{i=1}^{d+1} \lambda_{T,i}$, with $\lambda_{T,1}, \ldots, \lambda_{T,d+1}$ the barycentric co-ordinates of T.

We are now ready to state the finite element approximation for (PE) as follows:
(PE$_h$) Find $(u_h, p_h) \in V_h \times Q_h$ such that for all $v_h \in V_h$ and $q_h \in Q_h$,

$$a(u_h, v_h) + b(p_h, v_h) - \frac{1}{\varepsilon}\int_\Gamma \phi_\delta(u_{hn} + g_n)v_{hn} = (f, v_h) - [[\alpha, v_h]], \tag{26a}$$

$$b(q_h, u_h) = 0. \tag{26b}$$

Theorem 3. *There exists a unique solution* $(u_h, p_h) \in V_h \times Q_h$ *of* (PE_h)*, and we have*

$$\|u_h\|_1 + \|\hat{p}_h\| + \left\| \frac{1}{\varepsilon}\phi_\delta(u_{hn} + g_n) + k_h \right\|_{M_h'} \le C_*, \tag{27}$$

where $\hat{p}_h = p_h - k_h$, $k_h = \frac{1}{|\Omega|}\int_\Omega p_h \, dx$, *and* M_h' *is the dual space of* M_h.

In order to obtain the error estimates we need the stability result of ϕ_δ over M', the smaller space than M_h. Let us denote by $\mathcal{S}_h$ the $d-1$ dimensional triangulation of Γ inherited from $\mathcal{T}_h$.

Theorem 4. *Assume*

(A1) *the family* $\{\mathcal{S}_h\}_h$ *is quasi-uniform;*
(A2) *there exists* $\Gamma_1 \subset \Gamma$ *with* $|\Gamma_1| > 0$ *which is independent of* h, ε, δ *and* Ω *such that* $u_{hn} + g_n > 0$ *on* Γ_1.

Then, the solution $(u_h, p_h) \in V_h \times Q_h$ *of* $(\mathrm{PE}_{\varepsilon,\delta,h})$ *admits the following estimates:*

$$\|u_h\|_1 + \|p_h\| + \left\| \frac{1}{\varepsilon}\phi_\delta(u_{hn} + g_n) \right\|_{M_h'} \le C_*; \tag{28a}$$

$$\left\| \frac{1}{\varepsilon}\phi_\delta(u_{hn} + g_n) \right\|_{M'} \le C_* \left(1 + \frac{h}{\varepsilon}\right); \tag{28b}$$

$$\frac{1}{\sqrt{\varepsilon}}\|[u_{hn} + g_n]_-\|_\Gamma \le C_* \left(1 + \frac{\delta}{\varepsilon}\right). \tag{28c}$$

Theorem 5. *Assume that* (A1) *and* (A2) *are satisfied. Let* (u, p) *and* (u_h, p_h) *be solutions of* (PDE) *and* (PE_h)*, respectively, and suppose that* $(u, p) \in H^2(\Omega)^d \times H^1(\Omega)$ *and* $\tau_n(u, p) + \alpha_n \in M$. *Moreover, assume that* h, ε, δ *are sufficiently small and* $h \le c_1\varepsilon$ *with a constant* $c_1 > 0$, *then we have*

$$\|u - u_h\|_1 + \|p - p_h\| \le C_{**}\left(h + \varepsilon + \sqrt{\frac{\delta^2}{\varepsilon}}\right) \tag{29}$$

and

$$\left\| \tau_n(u, p) + \alpha_n - \frac{1}{\varepsilon}\phi_\delta(u_{hn} + g_n) \right\|_{M'} + \sqrt{\varepsilon}\left\| \tau_n(u, p) + \alpha_n - \frac{1}{\varepsilon}\phi_\delta(u_{hn} + g_n) \right\|_\Gamma \le C_{**}\left(h + \varepsilon + \sqrt{\frac{\delta^2}{\varepsilon}}\right), \tag{30}$$

where C_{**} *denotes a positive constant depending only on* c_1, Ω, $|u|_2$, $|p|_1$, $\|\tau_n(u, p) + \alpha_n\|_M$, $\|f\|$, $\|g\|_1$ *and* $\|\alpha\|_{(M^d)'}$. *Particularly, if taking as* $c_2\varepsilon \le h$ *and* $\delta \le c_3 h^{\frac{3}{2}}$ *with constants* $(c_1 >)c_2, c_3 > 0$, *we have the optimal-order error estimate*

$$\|u - u_h\|_1 + \|p - p_h\| \le C_{**}h.$$

5 Numerical Experiments

We present some results of numerical experiments in order to confirm our theoretical results. Let we consider a model Stokes problem with a Robin condition

$$-\nu\Delta v + \nabla q = f, \quad \nabla\cdot v = 0 \qquad \text{in } \Omega, \tag{31a}$$

$$v = b \qquad \text{on } S_1, \tag{31b}$$

$$v = 0 \qquad \text{on } S_2, \tag{31c}$$

$$\tau_n(v,q) = \frac{1}{\varepsilon}\phi_\delta(v_n), \quad v_T = 0 \qquad \text{on } \Gamma, \tag{31d}$$

Here, we prefer the original setting without any reference flow because it is not obvious that the reference flow (g, p) is always available in actual computations whereas it plays an important role in theoretical considerations.

The finite element approximation for (31) reads as follows.

(PE'$_h$) Find $(u_h, p_h) \in W_h \times Q_h$ such that $u_h = b$ on S_1 and

$$a(u_h, v_h) + b(p_h, v_h) - \frac{1}{\varepsilon}\int_\Gamma \phi_\delta(u_{hn})v_{hn}\,d\Gamma = \int_\Omega f\cdot v_h, \quad (\forall v_h \in V_h), \tag{32a}$$

$$b(q_h, u_h) = 0, \qquad (\forall q_h \in Q_h). \tag{32b}$$

where $W_h = \{v_h \in C^0(\overline{\Omega})^d \mid v_h|_{S_2} = 0,\ v_h|_T \in \mathcal{P}_1^{(d)} \oplus \mathrm{span}\{\varphi_T\},\ v_{hT}|_\Gamma = 0\}$.

We deal with two examples: The rectangle domain $\Omega = [0, L]\times[-R, R]$, ($S_1 = \{0\}\times[-R, R]$, and $\Gamma = \{L\}\times[-R, R]$), and a two-dimensional branched pipe domain Ω similar to Fig. 1.

In the first case, we impose $b(x, y) = (C_0(R^2 - y^2), 0)$, and $f \equiv 0$ with $C_0 > 0$. Then, (31) has the exact solution (the well-known Poiseuille flow) given as

$$u(x, y) = (C_0(R^2 - y^2), 0),\ \ p(x, y) = 2\nu C_0 L\left(1 - \frac{x}{L}\right). \tag{33}$$

In the second case we cannot obtain the (explicit) exact solution, but we make use of numerical solutions with extra fine mesh instead of it. We do not explain the detail of our computation here, but we choose parameters as $\varepsilon = \delta^{\frac{2}{3}} = h/20$. Hence, it is ensured by Theorem 5 that

$$\|v - v_h\|_1 + \|q - q_h\| \le Ch. \tag{34}$$

To verify this point, we set

$$E_h^{(1)} = \|v - v_h\|, \quad E_h^{(2)} = \|v - v_h\|_1, \quad E_h^{(3)} = \|q - q_h\|,$$

and observe that

$$\rho_h^{(i)} = \frac{\log E_{h'}^{(i)} - \log E_h^{(i)}}{\log h' - \log h} \qquad (i = 1, 2, 3)$$

with $h' \approx 2h$.

The results are reported in Tab. 1 for the first case and in Tab. 2 for the second case, respectively. Both of them support our theoretical result (34).

Table 1 Numerical convergence rates of (PE'$_h$) for rectangle.

h	$E_h^{(1)}$	$\rho_h^{(1)}$	$E_h^{(2)}$	$\rho_h^{(2)}$	$E_h^{(3)}$	$\rho_h^{(3)}$
1.0743	13.9	–	$1.20 \cdot 10^{2}$	–	$2.07 \cdot 10^{-1}$	–
0.5371	3.47	2.001	$5.95 \cdot 10^{1}$	1.010	$6.57 \cdot 10^{-2}$	1.656
0.2685	0.86	2.000	$2.97 \cdot 10^{1}$	1.003	$2.17 \cdot 10^{-2}$	1.594
0.1342	0.21	2.000	$1.48 \cdot 10^{1}$	1.001	$7.42 \cdot 10^{-3}$	1.553
0.0665	0.052	2.000	7.17	1.000	$2.56 \cdot 10^{-3}$	1.527

Table 2 Numerical convergence rates of (PE'$_h$) for branched pipe.

h	$E_h^{(1)}$	$\rho_h^{(1)}$	$E_h^{(2)}$	$\rho_h^{(2)}$	$E_h^{(3)}$	$\rho_h^{(3)}$
0.69279	$2.497 \cdot 10^{-1}$	—	5.940	—	$1.786 \cdot 10^{-1}$	—
0.33353	$7.766 \cdot 10^{-2}$	1.552	3.358	0.780	$5.909 \cdot 10^{-2}$	1.513
0.17571	$2.044 \cdot 10^{-2}$	2.083	1.767	1.001	$3.069 \cdot 10^{-2}$	1.022

6 Model Nonlinear BC

Finally, we also study the model Stokes problem with the Bazilevs outflow BC (9), or equivalently,

$$\begin{cases} \tau_n = -[u_n]_- u_n - p_0 - R_j \int_{\Gamma_j} u_n \, d\Gamma_j \\ \tau_T = [u_n]_- u_T \end{cases} \quad \text{on } \Gamma_j \quad (j = 1, 2) \tag{35}$$

If $\tau_T = u_T = 0$, we can show the well-posedness of this variational problem as similarly to (VI). In addition, for a rectangle domain Ω, since there is only one branch for outflow, we can put $p_0 = R = 0$ by scale transformations. Thus, (35) is written as

$$\tau_n = -[u_n]_- u_n \text{ and } \tau_T = [u_n]_- u_T = 0 \text{ on } \Gamma. \tag{36}$$

On the other hand, we know the penalty BC (13) holds energy inequality (4) for a sufficiently small ε, especially, $\varepsilon = 1/(\theta[u_n]_-)$ with $\theta \geq \frac{1}{2}$. Here, ordinally constant ε is regarded as a function. With this ε we can actually deduce (4) when u_T is small enough. Consequently, (10) is written as

$$\tau_n = \theta[u_n]_-^2 \text{ and } \tau_T = 0 \text{ on } \Gamma, \tag{37}$$

which coincides with (36) since $-[u_n]_- u_n = [u_n]_-^2$ for $\theta = 1$. That is, the regularized traction introduced by Bazilevs can be interpreted as a penalty approximation of the unilateral BC.

Acknowledgement We thank Professor K. Takizawa and H. Suito who brought the subject to our attention. This work is supported by JST, CREST, and JSPS KAKENHI Grant Number 15H03635, 15K13454.

References

1. Bruneau, C.H., Fabrie, P.: Effective down-stream boundary conditions for imcompressible Navier-Stokes equations. Internat. J. Numer. Methods Fluids **19**, 693–705 (1994)
2. Bazilevs, Y., Gohean, J.R., Hughes, T.J.R., Moser, R.D., Zhang, Y.: Patient-specific isogeometric fluid-structure interaction analysis of thoracic aortic blood flow due to implantation of the Jarvik 2000 left ventricular assist device. Comput. Methods Appl. Mech. Engrg. **198**(2009), 3534–3550 (2000)
3. Formaggia, L., Gerbeau, J.F., Nobile, F., Quarteroni, A.: Numerical treatment of defective boundary conditions for the Navier-Stokes equations. SIAM J. Numer. Anal. **40**, 376–401 (2002)
4. Lions, J.L., Magenes, E.: Non-homogeneous Boundary Value Problems and Applications, I, Springer (1972)
5. Kikuchi, N., Oden, J.T.: Contact Problems in Elasticity, SIAM (1988)
6. Vignon-Clementel, I.E., Figueroa, C.A., Jansenc, K.E., Taylor, C.A.: Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries. Comput. Methods Appl. Mech. Engrg. **195**, 3776–3796 (2006)
7. Saito, N., Sugitani, Y., Zhou, G.: Unilateral problem for the Stokes equations:the well-posedness and finite element approximation, UTMS Preprint Series 2015-3. http://www.ms.u-tokyo.ac.jp/preprint/2015/

Successive Projection with B-spline

Yuki Ueda and Norikazu Saito

Abstract Space-time finite element methods often construct the approximate solution by time discontinuous Galerkin methods. Successive projection technique (SPT) with B-spline functions allows us to convert this approximate solution into time-continuous representation. We present the stability and error estimate of SPT.

1 Introduction

The space-time finite element formulation is widely applied for analysis of time-dependent phenomena. The idea of discretization is based on discontinuous Galerkin methods, and governing equations are solved on the space-time slabs, which is the slice of space-time domain between two time steps. We can use high order polynomials for time basis function, and it provide higher accuracy of approximate solution.

The basis functions are continuous in space-time slab, but discontinuous between different slabs. The discontinuity of the basis functions allow us to compute the system gradually. If we use continuous basis functions for solving the variational problem on the space-time domain, it is necessary to solve $n + 1$ dimensional finite element methods, where n the dimension of spatial domain, and that takes extremely a lot of computational costs to solve very large system of linear equations when $n = 3$. Therefore, time discontinuous Galerkin methods are often applied to space-time computation for remove this difficulty.

However, time discontinuous Galerkin methods needs the large number of basis functions for representing the approximate solution. Therefore, it takes some computer storage costs to store the computed data. In addition, the discontinuity may give rise to a problem when we simulate some phenomena, because the simulation behave discontinuous in the time steps.

Y. Ueda(✉) · N. Saito
Graduate School of Mathematical Sciences, The University of Tokyo, Tokyo, Japan
e-mail: ueda@ms.u-tokyo.ac.jp

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_79

The successive projection technique (SPT) is proposed to convert the time discontinuous solution into the representation by B-spline [4]. B-spline is piecewise polynomial with some smoothness on the breakpoints, and the degrees of freedom is less than the less smooth piecewise polynomials. The SPT can replace discontinuous function with smooth one gradually, then we can execute the SPT computation in parallel with the time discontinuous Galerkin methods, and the computed data on each space-time slab can be changed with B-spline function immediately.

In this paper, we present the stability and error estimate of SPT projection. The results are given when the discontinuous function is real-valued, but it can be extended in the case of Hilbert space-valued function.

2 B-spline Functions

2.1 Definition and Properties

First, we introduce the (univariate) B-spline basis function. For given positive integer p, we define the knot vector

$$\Xi := \{\xi_1, \xi_2, \ldots, \xi_m\}, \tag{1}$$

where we assume $\xi_1 \leq \xi_2 \leq \cdots \leq \xi_m$, and without loss of generality, we also assume $\xi_1 = 0$ and $\xi_m = 1$. We note the repetition of the knots are allowed. Related to this knot, we introduce the representation

$$\Xi = \{\underbrace{\zeta_1, \ldots, \zeta_1}_{m_1 \text{ times}}, \underbrace{\zeta_2, \ldots, \zeta_2}_{m_2 \text{ times}}, \ldots, \underbrace{\zeta_r, \ldots, \zeta_r}_{m_r \text{ times}}\}, \tag{2}$$

where $0 = \zeta_1 < \zeta_2 < \cdots < \zeta_r = 1$. We denote by m_i the multiplicity of ζ_i, and we assume $m_i \leq p + 1$ for all knots.

For the knot vector Ξ, the B-spline basis functions of degree p are given by the Cox-de Boor algorithm [3]:

if $q = 0$,

$$B_{i,0}(t) := \begin{cases} 1 & \text{if } \xi_i \leq t < \xi_{i+1}, \\ 0 & \text{else,} \end{cases} \tag{3}$$

for $i = 1, 2, \ldots, m - 1$, and if $q = 1, 2, \ldots, p$, then

$$B_{i,q}(t) := \frac{t - \xi_i}{\xi_{i+q} - \xi_i} B_{i,q-1}(t) + \frac{\xi_{i+q+1} - t}{\xi_{i+q+1} - \xi_{i+1}} B_{i+1,q-1}(t), \tag{4}$$

for $i = 1, 2, \ldots, m - q - 1$. We note it is assumed that $0/0 = 0$ in this definition.

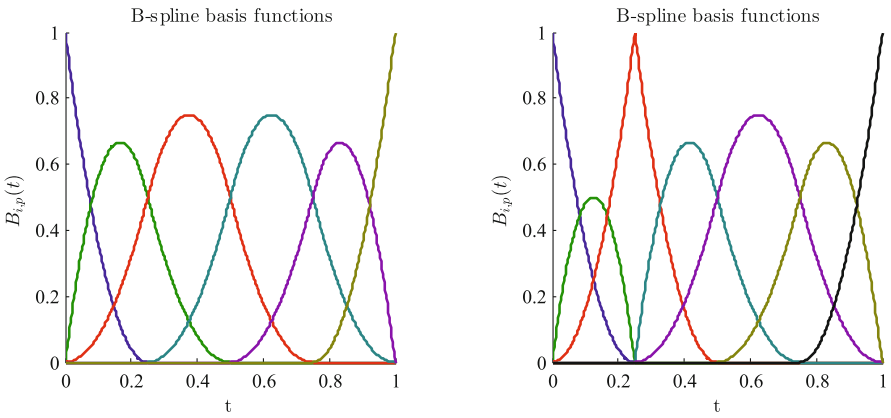


Fig. 1 The B-spline basis function of degree 2 for the knot vector {0, 0, 0, 1/4, 1/2, 3/4, 1, 1, 1} and {0, 0, 0, 1/4, 1/4, 1/2, 3/4, 1, 1, 1}.

For the knot vector Ξ, we let the (univariate) B-spline space $S_p(\Xi)$ is the space of p-th degree piecewise polynomials on $I := [0, 1]$ that are $p - m_i$ continuously-differentiable at ζ_i for $i = 2, 3, \ldots, r - 1$. Then the B-spline basis functions $B_{i,p}$ are in the space $S_p(\Xi)$.

Particularly, we assume the knot vector Ξ is open, that is, the multiplicities of external knots $m_1 = m_r = p + 1$, then the B-spline basis functions $B_{i,p}$ form a basis of $S_p(\Xi)$.

3 Basic Error Estimate for B-spline Approximations

Let $I = [0, 1]$ and $I_i := [\zeta_i, \zeta_{i+1}]$ for $i = 1, 2, \ldots, r - 1$. For the approximation error estimate, we introduce the bent Sobolev space [1], that is, the broken Sobolev space with some continuous derivative. We define the bent Sobolev space on I by

$$\mathcal{H}^s(I) := \left\{ \begin{array}{l} f \in L^2(I) : f|_{I_i} \in H^s(I_i) \text{ for } i = 1, 2, \ldots, r-1, \text{ and} \\ f \text{ has } \min\{s-1, p-m_i\} \text{ continuous derivative at } \zeta_i \end{array} \right\}. \quad (5)$$

This is a Hilbert space equipped with the semi-norm and norm

$$|v|^2_{\mathcal{H}^s(\Xi)} := \sum_{i=1}^{r-1} |v|^2_{H^s(I_i)}, \quad \|v\|^2_{\mathcal{H}^s(\Xi)} := \sum_{j=0}^{s} |v|^2_{\mathcal{H}^j(\Xi)}. \quad (6)$$

We have the following approximation error estimate [2].

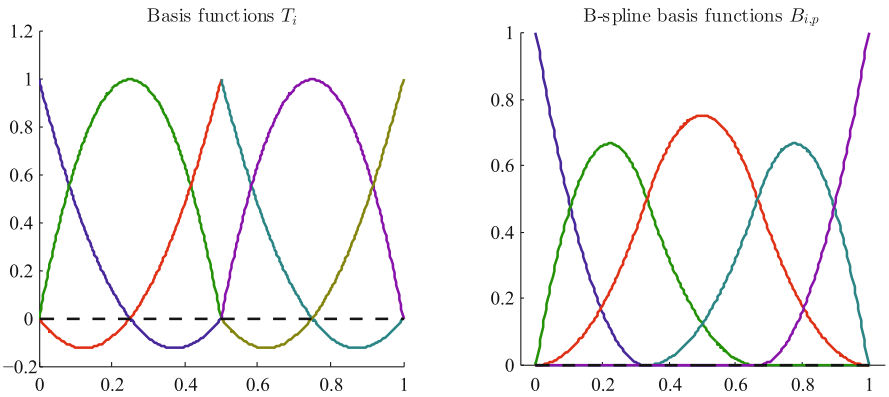


Fig. 2 Example of basis functions. Left: discontinuous piecewise polynomials T_i and right: B-spline basis functions $B_{i,p}$.

Lemma 1. *There exists a quasi-interpolant operator*

$$\Pi_{p,\Xi} : L^2(I) \to S_p(\Xi) \tag{7}$$

and a positive constant C such that, for all $s \in \mathbb{N}$, $s \leq p+1$,

$$\|f - \Pi_{p,\Xi}(f)\|_{L^2(I)} \leq Ch^s |f|_{\mathcal{H}^s(\Xi)}, \quad \forall f \in \mathcal{H}^s(\Xi) \tag{8}$$

where $h := \max\{\text{diam } I_i\}$.

4 Successive Projection to Spline Space

4.1 The Solution of Space-time Finite Element Methods

We assume that the governing equation is discretized by space-time finite element methods. Moreover, let the space-time mesh satisfy the following conditions ([5] and [6]): the space-time domain is divided in some space-time slabs Q_n, which is the slice of the domain between the time steps t_n and t_{n+1}. Then, all nodes of the space time slab are on the time plane t_n or t_{n+1}, and the number of nodes on each space-time slabs are equal. In other words, we construct the space-time mesh by transform the mesh that is structured in time. This means the approximate solution of space-time FEM can be written by the linear combination of $T_i(t)N_j(x)$ $(i = 1, 2, ..., n_t, j = 1, 2, ..., n_s)$, where $T_i(t)$ is the temporal basis function, $N_j(x)$ is the basis function of referential space, n_t and n_s are the number of temporal and spatial basis functions. By the time discontinuous Galerkin method,

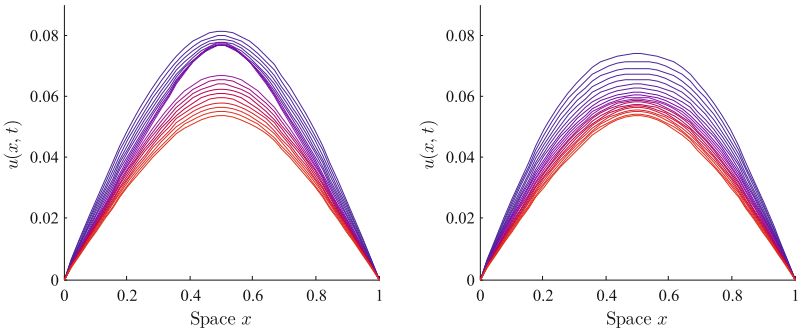


Fig. 3 An example of SPT. The solution of space-time FEM (left) jumps for some time steps, and SPT canverts the data into representation with B-spline (right).

the temporal basis functions $T_i(t)$ are piecewise polynomial that is discontinuous on each time step t_n. The successive projection technique (SPT) converts the approximate solution into representation with time spline function,

$$\sum_{i=1}^{n_t}\sum_{j=1}^{n_s} a_{i,j}T_i(t)N_j(x) \approx \sum_{i=1}^{m}\sum_{j=1}^{n_s} b_{i,j}B_{i,p}(t)N_j(x). \tag{9}$$

This reduces the degrees of freedoms, and the data can be expressed more efficiently. We note the coefficients $a_{i,j}$ are given gradually by space-time FEM. Therefore, SPT provide the way to compute the projection also gradually, and the projection can be constructed in parallel with the space-time FEM computation.

4.2 Successive Projection Technique (SPT)

For simplicity, we consider the projection from $L^2(I)$ to B-spline space. We let the open knot vector

$$\Xi = \{0 = \underbrace{t_0, \ldots, t_0}_{p+1 \text{ times}}, t_1, t_2, \ldots, t_{N-1}, \underbrace{t_N, \ldots, t_N}_{p+1 \text{ times}} = 1\} \tag{10}$$

where $N \in \mathbb{N}$ is partition number, and $h := 1/N$, $t_n := nh$. We note that the number of B-spline basis function for this knot vector Ξ is smaller because of the regularity. Then, we denote by

$$\Xi_n := \{\underbrace{t_0, \ldots, t_0}_{p+1 \text{ times}}, t_1, \ldots, t_{n-1}, \underbrace{t_n, \ldots, t_n}_{p+1 \text{ times}}\} \tag{11}$$

the successive knot vector. This is also an open knot vector. Furthermore, we denote by

$$B_{i,p}^{n}(t) \quad (i = 1, 2, \ldots, n + p) \tag{12}$$

the B-spline basis functions for knot vector Ξ_n, and define

$$S_n := \mathrm{span}\{B_{i,p}^{n} : i = 1, 2, \ldots, n + p\}, \tag{13}$$

$$\tilde{S}_n := \mathrm{span}\{B_{i,p}^{n} : i = n, n + 1, \ldots, n + p\}. \tag{14}$$

Then, we note the following properties holds because of the definition of B-spline basis functions.

Lemma 2. *$B_{i,p}^{n} = B_{i,p}^{n+1}$ for all $i = 1, 2, \ldots, n$*

Lemma 3. *$B_{i,p}^{n+1}(t) = B_{i-1}^{n}(t - h)$, $(t_1 \le t \le t_{n+1},\ \max\{n, p + 1\} \le i \le n + p)$*

Lemma 4. *$B_{i,p}^{n+1}(t) \neq 0$ for $t \in I_j$ if and only if $i = j + 1, \ldots, j + p + 1$.*

Now, we introduce the algorithm of SPT. For the given function $f \in L^2(I)$, we define $\Pi_1(f)$ by the L^2-projection of $f|_{t_0 \le t \le t_1}$ to S_1. Then, we construct the projection $\Pi_N(f) \in S_N = S_p(\Xi)$ with the following procedure: For given

$$\Pi_n(f) := \sum_{i=1}^{n+p} a_i^n B_{i,p}^{n} \in S_n \tag{15}$$

and $f|_{t_n \le t \le t_{n+1}}$, find

$$\Pi_{n+1}(f) := \sum_{i=1}^{n+p+1} a_i^{n+1} B_{i,p}^{n+1} \in S_{n+1} \tag{16}$$

such that

$$a_i^{n+1} = a_i^n, \ \text{for } i = 1, 2, \ldots, n \tag{17}$$

$$\int_{t_m}^{t_n} v_{n+1} \left(\Pi_{n+1}(f) - \Pi_n(f)\right) dt + \int_{t_n}^{t_{n+1}} v_{n+1} \left(\Pi_{n+1}(f) - f\right) dt = 0, \quad \forall v_{n+1} \in \tilde{S}_{n+1} \tag{18}$$

where $m := \max\{0, n - p\}$. We note the equation (17) is corresponding to the Lemma 2. It is means that for the same basis functions, we use the same coefficients for construct the projection. We also note that the data u on $[t_n, t_{n+1}]$ are necessary for one step, and they can be discard after we get $\Pi_{n+1}(f)$. Because of these features, the

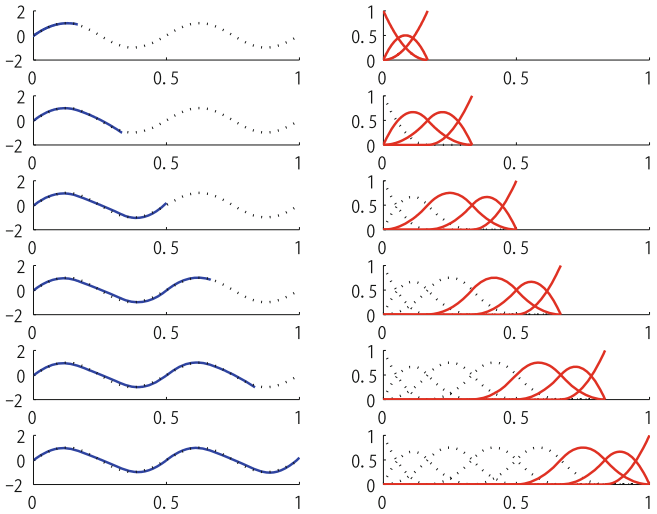


Fig. 4 Each step of successive projection and basis functions. The red lines represent the basis of $\tilde{S}_n$

SPT computation can be executed in parallel with the space-time FEM computation, and it is no need to store the data by space-time FEM after they are converted in a step of SPT.

4.3 Stability Analysis

First, we transfer the SPT algorithm in the matrix form. We define

$$\boldsymbol{M}_{n+1} := \left(\int_{t_m}^{t_{n+1}} B_{i+n,p}^{n+1} B_{j+n,p}^{n+1} dt \right)_{1 \le i,j \le p+1}, \tag{19}$$

$$\boldsymbol{L}_{n+1} := \left(\int_{t_m}^{t_n} B_{i+n,p}^{n+1} B_{j+n-1}^{n} dt \right)_{1 \le i,j \le p+1}, \tag{20}$$

$$F_{n+1} := \left(\int_{I_n} B_{i+n,p}^{n+1} f \, dt \right)_{1 \le i \le p+1}, \tag{21}$$

$$\tilde{\boldsymbol{a}}_{n+1} := \left(a_{i+n}^{n+1} \right)_{1 \le i \le p+1}, \tag{22}$$

where $I_n := [t_n, t_{n+1}]$ for $n = 0, 1, \ldots, N-1$. Then, the SPT scheme is equivalent to

$$\begin{cases} \tilde{\boldsymbol{a}}_1 = (\boldsymbol{M}_1)^{-1} F_1, \\ \tilde{\boldsymbol{a}}_{n+1} := \boldsymbol{K}_{n+1}\tilde{\boldsymbol{a}}_n + (\boldsymbol{M}_{n+1})^{-1} F_{n+1}, \quad (n = 1, 2, \ldots, N-1) \end{cases} \tag{23}$$

where $\boldsymbol{K}_{n+1} := (\boldsymbol{M}_{n+1})^{-1} \boldsymbol{L}_{n+1} LU$, and the matrix $U := (\delta_{i+1,j})_{1 \le i,j \le p+1}$ and $L := (\delta_{i,j+1})_{1 \le i,j \le p+1}$ are the shift matrices. Related to above matrix form, we obtain the following properties.

Lemma 5. *For all $n = 1, 2, \ldots, N-1$, the matrix $\boldsymbol{K}_{n+1}$ depends only on the degree p, and for all $n = 0, 1, \ldots, N-1$, there exists a matrix $\boldsymbol{C}_{n+1}$ depends only on the degree p such that $(\boldsymbol{M}_{n+1})^{-1} = h^{-1}\boldsymbol{C}_{n+1}$.*

Lemma 6. *For all $n \le p$, $\boldsymbol{M}_{n+1} = \boldsymbol{M}_{p+1}$ and $\boldsymbol{K}_{n+1} = \boldsymbol{K}_{p+1}$.*

Lemma 7. *For all $n = 0, 1, \ldots, N-1$, there exists a positive constant C that depends only on p such that*

$$\| (\boldsymbol{M}_{n+1})^{-1} F_{n+1} \|_\infty \le C h^{-\frac{1}{2}} \| f \|_{L^2(I_n)}. \tag{24}$$

These lemmas gives the following estimate for a^N_{n+1}.

Proposition 1.

$$\left(a^N_{n+1}\right)^2 \le C h^{-1} \sum_{i=0}^{\min\{n, N-1\}} \left\| \left(\sqrt{2}\boldsymbol{K}\right)^{\min\{n,N-1\}-i} \right\|^2_\infty \| f \|^2_{L^2(I_i)}, \tag{25}$$

where $\boldsymbol{K} := \boldsymbol{K}_{p+1}$, and we recall it depends only on the degree p.

Proof. Let $\tilde{n} := \min\{n, N-1\}$. First, SPT algorithm (17) leads $a^N_{n+1} = a^{N-1}_{n+1} = \cdots = a^{n+1}_{n+1}$ for $n = 0, 1, \ldots, N-1$. This gives $|a^N_{n+1}| \le \|\tilde{\boldsymbol{a}}_{\tilde{n}+1}\|_\infty$. Regarding to the right side,

$$\tilde{\boldsymbol{a}}_{n+1} = K_{n+1}\tilde{\boldsymbol{a}}_n + (\boldsymbol{M}_{n+1})^{-1} F_{n+1} = \sum_{i=1}^{n+1} \left(\prod_{j=i+1}^{n+1} \boldsymbol{K}_j \right) (\boldsymbol{M}_i)^{-1} F_i. \tag{26}$$

follows from the equation (23). Therefore, we can get the estimate

$$|a^N_{n+1}| \le \|\tilde{\boldsymbol{a}}_{\tilde{n}+1}\|_\infty \le C h^{-\frac{1}{2}} \sum_{i=1}^{\tilde{n}+1} \left\| \prod_{j=i+1}^{\tilde{n}+1} \boldsymbol{K}_j \right\|_\infty \| f \|_{L^2(I_{i-1})}. \tag{27}$$

Here, lemma 5 and lemma 6 implies that the estimate

$$\left\| \prod_{j=i+1}^{\tilde{n}+1} \boldsymbol{K}_j \right\|_\infty \le C \|\boldsymbol{K}^{\tilde{n}-i+1}\|_\infty \tag{28}$$

for some positive constant C that depends only on p. Therefore,

$$\begin{aligned}
\left(a_{n+1}^N\right)^2 &\le Ch^{-1} \left(\sum_{i=0}^{\tilde{n}} \|\boldsymbol{K}^i\|_\infty \|f\|_{L^2(I_{\tilde{n}-i})} \right)^2 \\
&\le Ch^{-1} \left(2\|\boldsymbol{K}^0\|_\infty^2 \|f\|_{L^2(I_{\tilde{n}})}^2 + 2 \left(\sum_{i=1}^{\tilde{n}} \|\boldsymbol{K}^i\|_\infty \|f\|_{L^2(I_{\tilde{n}-i})} \right)^2 \right) \\
&\le 2Ch^{-1} \sum_{i=0}^{\tilde{n}} \left\| \left(\sqrt{2}\boldsymbol{K}\right)^i \right\|_\infty^2 \|f\|_{L^2(I_{\tilde{n}-i})}^2
\end{aligned}$$

so that (25) follows. □

This proposition shows the sufficient condition of the stability.

Theorem 1. *If there exists a natural number q that depends only on p such that*

$$\left\| \left(\sqrt{2}\boldsymbol{K}\right)^q \right\|_\infty < 1, \tag{29}$$

then, there exists a positive constant $C = C(p)$ such that

$$\|\Pi_N(f)\|_{L^2(I)} \le C \|f\|_{L^2(I)} \tag{30}$$

for all $f \in L^2(I)$.

Proof. We note the each supports of B-spline basis functions are local (lemma 4), it makes

$$\|\Pi_N(f)\|_{L^2(I)}^2 \le \sum_{j=0}^{N-1} \left\| \sum_{i=j+1}^{j+p+1} a_i^N B_{i,p}^N \right\|_{L^2(I_j)}^2. \tag{31}$$

The right side can be bounded by $Ch \sum_{i=1}^{N+p} \left(a_i^N\right)^2$, therefore

$$\begin{aligned}
\|\Pi_N(f)\|_{L^2(I)}^2 &\le C \sum_{i=1}^{N+p} \sum_{j=0}^{\min\{i-1,N-1\}} \left\| \left(\sqrt{2}\boldsymbol{K}\right)^{\min\{i-1,N-1\}-j} \right\|_\infty^2 \|f\|_{L^2(I_j)}^2 \\
&\le C \sum_{i=0}^{N-1} \|f\|_{L^2(I_j)}^2,
\end{aligned}$$

because

$$\sum_{i=0}^{\infty} \left\| \left(\sqrt{2}\boldsymbol{K}\right)^i \right\|_\infty^2 \leq \sum_{j=0}^{\infty} \left(\left\| \left(\sqrt{2}\boldsymbol{K}\right)^q \right\|_\infty^2 \right)^j \left(\sum_{i=1}^{q-1} \left\| \left(\sqrt{2}\boldsymbol{K}\right)^i \right\|_\infty^2 \right) \leq C(p) \quad (32)$$

follows from $\left\| \left(\sqrt{2}\boldsymbol{K}\right)^q \right\|_\infty < 1$ and lemma 5. □

We have confirmed the condition $\left\| \left(\sqrt{2}\boldsymbol{K}\right)^q \right\|_\infty < 1$ for some $q = q(p)$ when the degree $p = 1, 2, 3$, because the matrix $\boldsymbol{K}$ depends only on p and it can be calculated directly.

4.4 The Error Estimate of SPT Projection

It is obvious that the projection Π_N satisfies the spline preserving property.

Lemma 8. *$\Pi_N(u) = u$ for all $u \in S_p(\Xi)$.*

This lemma, lemma 1 and the stability result imply the error estimate.

Theorem 2. *If $\left\| \left(\sqrt{2}\boldsymbol{K}\right)^q \right\|_\infty < 1$ for some $q = q(p) \in \mathbb{N}$, there exists a positive constant $C = C(p)$ such that, for all $s \in \mathbb{N}$, $s \leq p + 1$,*

$$\|f - \Pi_N(f)\|_{L^2(I)} \leq Ch^s |f|_{\mathcal{H}^s(\Xi)}. \quad (33)$$

We note that the order of mesh size h of SPT error estimate is equal to one of the ordinary L^2-projection.

5 Application to the Space-time FEM

The error estimate of SPT can be applied in the conversion of time discontinuous approximate solution to the representation with B-spline function as follows: we denote by u and u_h the exact solution and the approximate solution that is discontinuous in time variable, and we assume the exact solution has enough regularity. Then we can estimate

$$\|u - \Pi_N(u_h)\|_{L^2(I)} \leq \|u - \Pi_N(u)\|_{L^2(I)} + C\|u - u_h\|_{L^2(I)}. \quad (34)$$

This estimate implies that if the SPT error can be bounded enough small, space-time FEM's error control the error of entire.

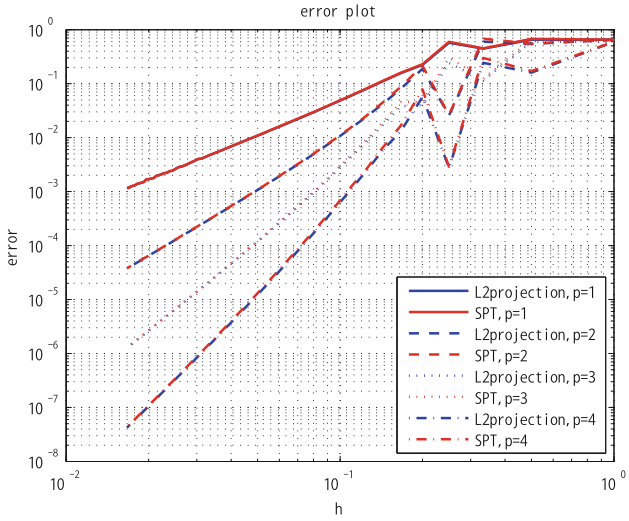


Fig. 5 L^2-error of SPT and ordinary L^2-projection with $f(t) = \sin(4\pi s)$

6 Conclusions

In this paper we have described the mathematical formulation of the SPT, and presented the stability analysis and the error estimate. The construction of the open knot vector leads some useful properties, then we can find the stability estimate hold if the norm of a matrix that depends only on the degree is small, and we have confirmed the condition hold when $p = 1, 2, 3$. We show the numerical example when the function is real-valued. These results can be extended to the Hilbert-valued function, thus the idea of SPT is available in many result of space-time finite element methods.

Acknowledgement We thank Professor K. Takizawa who brought the subject to our attention. The authors are supported by JST, CREST, and JSPS KAKENHI Grant Number 15H03635, 15K13454. Yuki Ueda also supported by Program for Leading Graduate Schools, MEXT, Japan.

References

1. Bazileves, Y., Beirão da Veiga, L., Cottrell, J.A., Hughes, T.J.R., Sangalli, G.: Isogeometric analysis: approximation, stability and error estimate for h-refinement meshes. Mech. Models Methods Appl. Sci. **16**(7), 1031–1090 (2006)
2. Beirão da Veiga, L., Buffa, A., Sangalli, G., Vázquez, R.: Mathematical analysis of variational isogeometric methods. Acta Numerica **23**, 157–287 (2014)
3. Piegl, L., Tiller, W.: The NURBS Book. Springer (1977)
4. Takizawa, K., Tezduyar, T.E.: Space-time computation technique with continuous representation in time (ST-C). Comput. Mech. **53**, 91–99 (2014)

5. Tezduyar, T.E., Behr, M., Lions, J.: A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain / space-time procedure: I. The concept and the preliminary numerical tests. Comput. Methods Appl. Mech. Engrg. **94**, 339–351 (1992)
6. Tezduyar, T.E., Behr, M., Mittal, S., Lions, J.: A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain / space-time procedure: II. Computation of free-surface flows, two-liquid flows, and flows with drifting cylinders. Comput. Methods Appl. Mech. Engrg. **94**, 353–371 (1992)

Computer-Aided Diagnosis Method for Detecting Early Esophageal Cancer from Endoscopic Image by Using Dyadic Wavelet Transform and Fractal Dimension

Ryuji Ohura, Hajime Omura, Yasuhisa Sakata and Teruya Minamoto

Abstract We propose a new computer-aided method for diagnosing early esophageal cancer from endoscopic images by using the dyadic wavelet transform (DYWT) and the fractal dimension. In our method, an input image is converted into HSV color space, and a fusion image is made from the S (saturation) and V (value) components based on the DYWT. We apply the contrast enhancement to produce a grayscale image in which the structure of abnormal regions is enhanced. We can obtain binary images composed of multiple layers by low-gradation processing. We visualize abnormal regions by summing these fractal dimensions by computing the complexity of these images. We describe a process for enhancing, detecting and visualizing abnormal regions in detail, and we present experimental results demonstrating that our method gives visualized images in which abnormal regions in endoscopic images can be located and that contain data useful for actual diagnosis of early esophageal cancer.

Keywords Computer-aided diagnosis · Early esophageal cancer · Endoscopic image · Fractal dimension · Dyadic wavelet transform

1 Introduction

In Japan, the incidence of esophageal cancer is more than 20,000 patients per year, and the number of deaths is more than 10,000 patients per year. Although diagnosis support systems are widely used in medical institutions, the number of patients and deaths are increasing year by year. In addition, it is difficult to decrease the risk of

R. Ohura(✉) · H. Omura · T. Minamoto
Department of Information Science, Saga University, Saga, Japan
e-mail: {ohura,ohmurah,minamoto}@ma.is.saga-u.ac.jp

Y. Sakata
Department of Internal Medicine, Saga University, Saga, Japan
e-mail: sakatay@cc.saga-u.ac.jp

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_80

death primarily caused by esophageal cancer to less than 50%. The goal of our work is to develop a computer-aided diagnosis (CAD) system for assisting and training doctors to allow them to accurately diagnose the presence or absence of cancer.

In order to concentrate on treatment, it is very important to find cancer as early as possible. However, it is difficult for inexperienced doctors to perform diagnosis from endoscopic images without overlooking early cancer, and doctors have to spend a lot of time to learn the skills necessary to detect early cancer from endoscopic images. As shown in Fig. 1, in the case of early esophageal cancer, abnormal regions appear very similar to normal regions, and visual differences, such as the color information and structure of the surface, cannot be observed. In image-based diagnosis for detecting early cancer using an endoscope system, the viewpoints for diagnosing a normal region and an abnormal region differ from doctor to doctor.

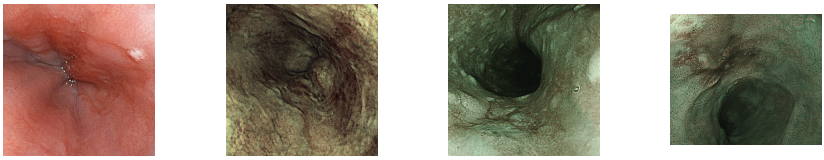


Fig. 1 An endoscopic image including a region of early esophageal cancer: Normal, FICE, BLI and NBI modes (from left to right).

Doctors use a number of methods to diagnose early cancer, such as the method of spraying indigo carmine dye or Lugol's iodine dye, as well as endoscopy or endoscopic ultrasonography. The method of spraying indigo carmine dye and Lugol's iodine dye makes it easier for doctors to find an early cancer during endoscopy. However, many patients complain of symptoms such as chest pain. In addition, there is a disadvantage that the symptoms last from a few hours to one day. There is a demand to establish a new diagnostic criteria for diagnosing the presence of cancer by developing a new endoscope system that does not require dyeing to avoid such side effects. In the case of early esophageal cancer, there is no established diagnostic criteria, such as classification by pit patterns used in the case of colorectal cancer.

In recent years, research on medical image analysis with a focus on the fractal dimension has been actively carried out. Russel et al. [3] described the box-counting method, which has been developed as a way of easily computing the fractal dimension using a computer, and as a computing method to measure the complexity of images. The box-counting method has been applied to a variety of medical images. For instance, Mavroforakis et al. performed textural classification of mammography images [5].

We propose a new computer-aided method of diagnosing early esophageal cancers from endoscope images that emphasizes an abnormal region from S (saturation) and V (value) components in HSV color space and binarizes the enhanced image into multiple layers by low-gradation processing. For the binary image corresponding to each layer, the fractal dimension is computed by using a box-counting method

proposed in Ref. [9]. We then sum the fractal dimensions for all the layers to improve the detection accuracy of abnormal regions.

The remainder of this paper is organized as follows. In Section 2, we briefly explain our new laser endoscope system. In Section 3, we provide a simple explanation of the dyadic wavelet transform (DYWT), and in Section 4, we briefly describe the fractal dimension and introduce a method of computing it, specifically, the box-counting method. In Section 5, we propose a new computer-aided method for diagnosing early esophageal cancer from endoscopic images by using the DYWT and the fractal dimension. Experimental results are presented in Section 6, and Section 7 concludes the paper.

2 New Laser Endoscope System

There are no diagnostic criteria for endoscope diagnosis of the presence of cancer that does not require dyeing to avoid side effects. Endoscope systems that solve this problem have been developed, as described in Refs. [1, 7]. In cancer screening using the current endoscope system of Fujifilm Corp., an endoscope specialist observes the organ under examination while switching among 4 modes: Normal (White light), FICE (Flexible spectral Imaging Color Enhancement), BLI (Blue LASER Imaging), and BLI-bright modes. Similarly, in an endoscope system that Olympus Corp has developed, an endoscope specialist observes the organ under examination while switching between 2 modes: Normal and NBI (Narrow Band Imaging). See Refs. [1, 7] for details regarding features of the modes.

The cancer screening is dependent on the subjective judgment of doctors, and standardized diagnostic criteria do not exist. If we can develop certain objective (computable) criteria that take into consideration doctors' specialized knowledge and accumulated experience to accurately classify abnormal and normal regions, then we could detect the differences between early cancer and non-cancer regions from an endoscopic image by quantifying and visualizing the abnormal regions based on such criteria. As a result, we wish to develop a new system that can support doctors in the diagnosis of esophageal lesions by realizing both bright images, as in the Normal mode, and high-contrast observation, as in the FICE and BLI modes.

3 Dyadic Wavelet Transform (DYWT)

Let us represent the original image by samples of a normalized discrete image defined on a two-dimensional lattice of finite extent, $C^0[m, n]$. It is well-known from Refs. [4, 6] that the DYWT for this image is:

$$C^{j+1}[m,n] = \sum_k \sum_l h[k]h[l]C^j_{k,l}[m,n],\ D^{j+1}[m,n] = \sum_k \sum_l g[k]h[l]C^j_{k,l}[m,n],$$
$$E^{j+1}[m,n] = \sum_k \sum_l h[k]g[l]C^j_{k,l}[m,n],\ F^{j+1}[m,n] = \sum_k \sum_l g[k]g[l]C^j_{k,l}[m,n].$$

Here, $C^j_{k,l}[m,n] = C^j[m + 2^j k, n + 2^j l]$, h is a low-pass filter, and g is a high-pass filter. More precisely, $C^j[m,n]$, $D^j[m,n]$, $E^j[m,n]$, and $F^j[m,n]$ are the low-frequency components, the horizontal high-frequency components, the vertical high-frequency components, and the high-frequency components in both directions, respectively. The indices m and n indicate the locations in the horizontal and vertical directions, respectively.

4 Fractal Dimension

The fractal dimension is an index for characterizing a target figure by using a self-similar figure. It can be used as a measure of complexity by comparing how the detail in a self-similar figure changes with the scale. Let S be an input image and $N(s)$ the minimum number of circles of diameter s needed to cover S.

$$D = -\lim_{s \to 0} \frac{\log N(s)}{\log s}.$$

The fractal dimension differs from Euclidean dimension and is a non-integer value. Von Koch curves are often used as an example to measure the fractal dimension. Fig. 2 illustrates Von Koch curves in the case of a well-known fractal object.

Fig. 2 Von Koch curves.

The fractal dimension of the Von Koch curves in Fig. 2 is computed as follows:

$$D = -\frac{\log 4^3}{\log 1/3^3} = \frac{\log 64}{\log 27} \approx 1.262.$$

Because an endoscope image is typically composed of complicated shapes and textures, it is difficult to define the fractal dimension of the endoscopic images. We applied the box-counting method to measure the complexity of the image. This is explained in the following subsection.

4.1 Box-Counting Method

The box-counting method is often used to compute the fractal dimension of an image. In this method, we cover a binarized image with square boxes of side length r, and the fractal dimension is computed as follows:

$$FD = -\lim_{r \to 0} \frac{\log N(r)}{\log r}.$$

Here, $N(r)$ is number of boxes needed to completely cover the binarized image. The algorithm of the box-counting method is as follows:

1. Divide the binarized image into non-overlapping square boxes of side length r.
2. Count the number of boxes, $N(r)$, containing the value 1, which represents a black square.
3. Repeat Step 1 and 2 while changing the side length r of square boxes.
4. Plot the logarithm of the side length r of square boxes on the horizontal axis and the logarithm of $N(r)$ on the vertical axis.
5. Obtain the straight line that passes through the plotted points, which has a negative slope. The absolute value of the slope is the fractal dimension, FD.

5 Algorithm for Visualizing Abnormal Regions

In this section, we present an algorithm for visualizing abnormal regions based on the DYWT and the fractal dimension for CAD from endoscopy images.

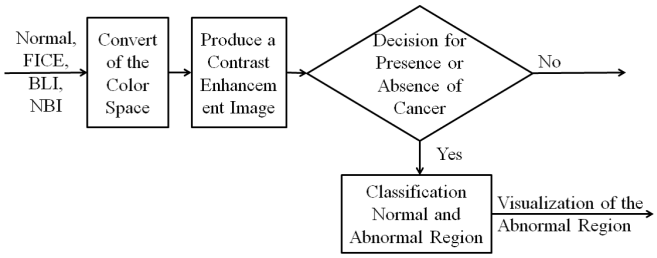


Fig. 3 System structure and features.

Fig. 3 shows a diagram of the system structure for detecting early esophageal cancer with the proposed CAD.

The structure enhancement procedure is as follows:

1. Normalize the input image in RGB color space so that the closed interval [0,255] becomes the closed interval [0,1].

2. Convert the normalized RGB image with the closed interval [0,1] to HSV color space, and obtain the H (hue), S (saturation) and V (value) components with the closed interval [0,1].
3. Apply the DYWT once to each of the S and V components. In Ref. [9], only low-frequency components were used as the analysis target. Therefore, the low-frequency components S_{Low} and V_{Low} are used here.
4. Compute the fusion image SV using the low-frequency components S_{Low} and V_{Low}:
$$SV[i, j] = \sqrt{(S_{Low}[i, j])^2 + (V_{Low}[i, j])^2}.$$
5. Normalize the fusion image to the closed interval [0,255].
6. Apply conventional contrast enhancement (such as MATLAB histeq and imadjust commands) to the fusion image SV and obtain the structure-enhanced image SV_{hist}.

We experimented with the idea of false positives, that is, the detection of normal regions as abnormal regions. The detecting and visualizing process is as follows:

1. Set the number of layers, Q, using low-gradation processing, and obtain multiple binarized images. Specifically, apply the structure-enhanced image SV_{hist} to the following:
$$I_N[i, j] = \begin{cases} 1 & R(N-1) \leq SV_{hist}[i, j] < RN \\ 0 & otherwise, \end{cases}$$
where P is the number of gradation levels of the structure-enhanced image SV_{hist}, $N(1 \leq N \leq Q)$ is the layer number, and $R = \frac{P}{Q}$ is the number of elements in each layer.
2. Segment each layer into 16×16-pixel non-overlapping blocks and compute the fractal dimension in each small block using the box-counting method.
3. Sum up the values of these fractal dimensions in each block and quantify and visualize the abnormal regions.
$$FD_{sum} = \sum_{N=1}^{Q} FD(I_N).$$
Here, FD is the fractal dimension of each layer. FD_{sum} is visualized as a grayscale image.

For the enhanced image SV_{hist}, low-gradation processing is performed from 2 to 255 gradations, and the fractal dimension at each layer is obtained. Finally, obtain the sum of the fractal dimensions to improve the detection accuracy, and highlight abnormal regions.

In Step 1, subject $P = 256$ grayscale images to low-gradation processing by using the number of layers $Q = 4$. Since $R = 256/4 = 64$, in the first layer of a binary image, the values of the pixels included in the closed interval [0,63] are all set to 1

(black box), and the others are set to 0 (white box). In other words, the values of the pixels included in the closed interval [64,255] are all set to 0. In the second layer of a binary image, the values of the pixels included in the closed interval [64,127] are all set to 1, and the others are set to 0. In other words, the values of the pixels included in the closed intervals [0,63] and [128,255] are all set to 0. The same process is applied to obtain binary images with $Q = 4$ layers.

6 Experimental Results

In the experiment, we used endoscopic images of 23 early esophageal cancer patients, and there were 135 images in total. These images are not illustrated in the present paper due to space limitations. To evaluate the performance of the proposed method, we compared abnormal regions detected by the proposed method with the regions identified as early esophageal cancer when diagnosed by medical experts. In the actual computations, we chose the parameter value $Q = 16$. Fig. 4 shows the results of applying our method using the original image S and V instead of the low-frequency components S_{Low} and V_{Low}, respectively, in Step 4 of the structure enhancement procedure. Applying contrast enhancement to the fusion image using the original image S and V directly, we obtain the enhancement image with added noise on the smooth areas in the original image. Consequently, this phenomenon often leads to reduce the precision of the image analysis. To avoid this problem, we adopted the image fusion based on the DYWT.

Fig. 4 Endoscopic image obtained in Normal mode, the fractal dimension visualized with our method, the fractal dimension with our method using the original image, and the region of esophageal cancer marked by a doctor (from left to right).

Fig. 5 shows the results of applying our method in the Normal mode of the Fujifilm endoscope system. As shown in Fig. 5, the white-light image shows a slightly reddish lesion. Although inexperienced doctors may be unable to detect the abnormal region with only white-light endoscopy, the abnormal region can be perceived by using our proposed method.

Fig. 6 shows the results of applying our method in the FICE mode of the Fujifilm endoscope system. As shown in Fig. 6, the lesion can be seen as a dark brown region in the FICE image. Although an experienced endoscopist would hardly miss the lesion in this mode, our proposed method further enhanced the abnormal region, making

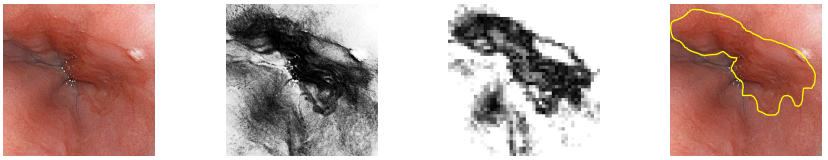


Fig. 5 Endoscopic image obtained in Normal mode, original image, structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

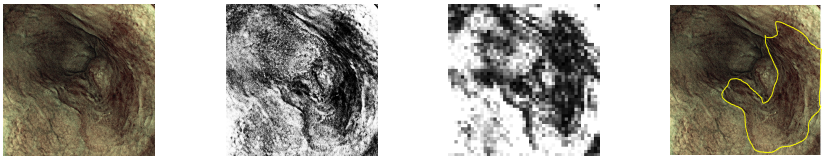


Fig. 6 Endoscopic image obtained in FICE mode, original image, the structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

it easy to detect the presence or absence of a cancer. Thus, even an inexperienced doctor can perceive the abnormal region.

Figs. 7 and 8 show the results of applying our method in the BLI and BLI-bright modes of the Fujifilm endoscope system. As shown in Figs. 7 and 8, the lesion is visible as a brownish area in the BLI and BLI-bright modes. Since the system can provide high-contrast, bright images, surface structures and blood vessels can be observed more clearly than in the white-light image. However, our proposed method further enhances abnormal regions and is effective for identifying the presence or absence of a cancer and its approximate location.

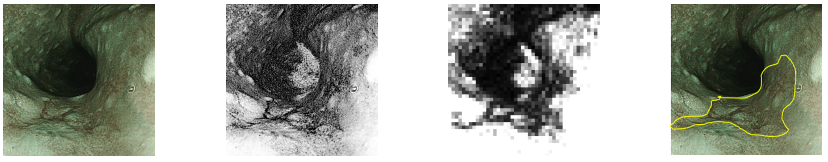


Fig. 7 Endoscopic image obtained in BLI mode, original image, the structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

Fig. 9 shows the results of applying our method in the Normal mode of the Olympus endoscope system. As shown in Fig. 9, the white-light image shows a slightly reddish lesion. Although an inexperienced doctor may be unable to detect the abnormal region with only white-light endoscopy, the abnormal region can be perceived by using our proposed method.

Fig. 10 shows the results of applying our method in the NBI mode of the Olympus endoscope system. As shown in Fig. 10, the lesion is visible as a brownish area

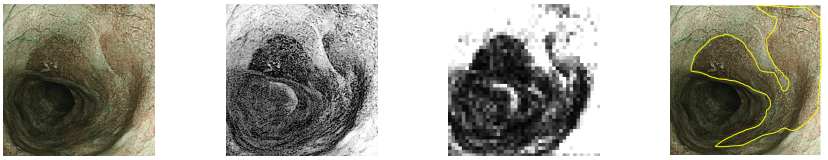


Fig. 8 Endoscopic image obtained in BLI-bright mode, original image, the structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

Fig. 9 Endoscopic image obtained in Normal mode, original image, the structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

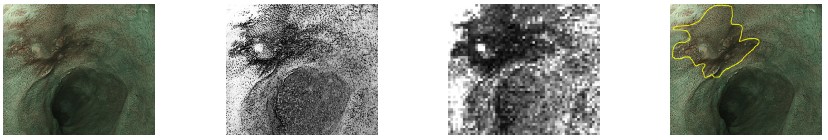


Fig. 10 Endoscopic image obtained in NBI mode, original image, the structure-enhanced image, the fractal dimension visualized with our method, and the region of esophageal cancer marked by a doctor (from left to right).

in the NBI mode. Since our method can provide a high-contrast image, surface structures and blood vessels can be observed more clearly than in the white-light image. However, our proposed method further enhances abnormal regions and is effective for identifying the presence or absence of a cancer and its approximate location.

As shown in Figs. 5–10, thus, even an inexperienced doctor can perceive abnormal regions by using our proposed method in any mode. In addition, we are able to visualize and highlight the region in which the doctor diagnosed the cancer.

7 Conclusion

We have proposed a computer-aided method for the diagnosis of early esophageal cancer from endoscopic images by using the DYWT and the fractal dimension. Experimental results showed that we were able to provide diagnostic information that can assist medical experts by using a CAD system. A CAD system for diagnosing the presence of cancer must be based on new diagnostic criteria that do not require the use of dyeing to avoid the associated side effects.

Our method can visualize abnormal regions such as early esophageal cancers, which are easy to overlook by inexperienced doctors. The method is independent of the particular endoscope system and the observation mode used. Therefore, if our method could be implemented in an endoscope system, to help in the early detection and treatment of early esophageal cancer, it would be possible to provide an endoscope diagnosis system having a lower risk of side effects and that can contribute to a reduction in fatalities. Since the size of each frequency component obtained by the DYWT is the same as the size in the original image, we can use a smaller block in comparison with the method based on the DWT described in Ref. [9] when analyzing the image. This property make it easier to detect and visualize the abnormal regions in the endoscopic image with high precision. Extraction of better new features using the high-frequency component, development of automatic detection, system implementation and clinical trials remain the topics of a future study.

Acknowledgments Thanks are due to staff in the Department of Internal Medicine, Saga University, Japan, for helpful suggestions and comments. Their medical images and specialist advice were the motivation for us to attempt this study. This work was partially supported by JSPS KAKENHI Grant Number 26400206.

References

1. Gono, K., Obi, T., Yamaguchi, M., Ohyama, N., Machida, H., Sano, Y., Yoshida, S., Hamamoto, Y., Endo, T.: Appearance of enhanced tissue features in narrow-band endoscopic imaging. Journal of Biomedical Optics **9**(3), 568–577 (2004)
2. Kenneth, F.: Fractal geometry : mathematical foundations and applications. 3rd edition, Wiley (2014)
3. Lopes, R., Betrouni, N.: Fractal and multifractal analysis: A review. Medical Image Analysis **13**(4), 634–649 (2009)
4. Mallat, S.: A wavelet tour of signal processing. Academic Press (2009)
5. Mavroforakis, M., Georgiou, H., Dimitropoulos, N., Cavouras, D., Theodoridis, S.: Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers. Artificial Intelligence In Medicine **37**, 145–162 (2006)
6. Minamoto, T., Ohura, R.: A Blind Digital Image Watermarking Method Based on the Dyadic Wavelet Transform and Interval Arithmetic. Applied Mathematics and Computation **226**, 306–319 (2014)
7. Morimoto, Y., Kubo, M., Kuramoto, M., Yamaguchi, H. and Kaku, T.: Development of a New Generation Endoscope System with Lasers “LASEREO”. FUJIFILM Research & Development, No. 58-2013, p. 6
8. Russel, D., Hanson, J., Ott, E.: Dimension of strange attractors. Physical Review Letters **45**(14), 1175–1178 (1980)
9. Yamaguchi, J., Yoneyama, A. and Minamoto, T.: Automatic detection of early esophageal cancer from endoscope image using fractal dimension and discrete wavelet transform. ITNG ’15: Proceedings of the 2015 12th International Conference on Information Technology: New Generations, 317-322 (2015), IEEE Computer Society

Daubechies Wavelet-Based Method for Early Esophageal Cancer Detection from Flexible Spectral Imaging Color Enhancement Image

Hiroki Matsunaga, Hajime Omura, Ryuji Ohura and Teruya Minamoto

Abstract We propose a new method for detecting early esophageal cancer from Flexible spectral Imaging Color Enhancement (FICE) mode images based on the Daubechies wavelet transform. In our method, we convert an image into CIEL*a*b* color space and use the a* components. Next, we divide the a* components into small blocks, apply two types of Daubechies wavelet transforms to each block, and then obtain the low- and high-frequency components at each block. The histogram of the low-frequency components tends to be positioned at the right and the left of a particular value for abnormal and normal regions, respectively. The histogram of the high-frequency components for an abnormal region has a longer tail. We describe the detection procedure of the abnormal regions in detail and present experimental results demonstrating that our method is able to detect early esophageal cancer from FICE images based on these features.

Keywords Frequency analysis · Early esophageal cancer · CIEL*a*b* color space · Discrete wavelet transform · Daubechies wavelets

1 Introduction

In health screening and cancer screening, a doctor inserts an endoscope equipped with a small camera at the tip of the tube into the esophagus and looks for esophageal cancer by observing the surface structure of mucosa or lesions. Doctors typically use different types of endoscopic color images in RGB color space, such as Normal, FICE, BLI (BLI-bright) and NBI mode images. Progression of cancer is determined from the position, size, number, surface structure and color of lesions. Furthermore, a specialist observes the esophagus in detail using a dye, agent and cancer

H. Matsunaga(✉) · H. Omura · R. Ohura · T. Minamoto
Department of Information Science, Saga University, Saga, Japan
e-mail: {matsunah,ohmurah,ohura,minamoto}@ma.is.saga-u.ac.jp

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_81

is diagnosed by pathological examination after collecting a tissue sample from the site in question and checks whether the tissue includes cancer cells. Therefore, endoscope systems provided with FICE, BLI (BLI-bright), and NBI modes and that can highlight abnormal regions within an organ have been under development.

Olympus' endoscope system has the largest with a 43.2% share in Japan, and Fujifilm's endoscope system has the second largest with a 15.2% share in Ref. [5]. As we will describe in Section 2, Olympus Corp. and Fujifilm Corp. have developed NBI and FICE modes, respectively. Hence, the number of research papers based on NBI larger than the number of papers based on FICE. using the NBI mode for digestive system cancer in Ref. [7]. Of the optical diagnostics using the NBI mode for the rectosigmoid in Ref. [8]. On the other hand, cancer screening using the FICE mode depends on the doctor's knowledge and experience, and there are no studies that have evaluated the usefulness of the FICE mode from the point of view of medical engineering.

However, there are no examples involving contour and texture analysis of images, as far as we know. That is, the features represented by high-frequency components are not used in computer-based cancer detecting systems. in this study, we intend to clarify the usefulness, in terms of medical engineering, of cancer screening from an FICE image by finding new features of abnormal regions from the obtained frequency components by applying a Daubechies wavelet transform to the FICE image. In the case of early esophageal cancer, there are no established diagnostic criteria, such as classification based on the pit patterns in the case of colorectal cancer using the NBI mode in Ref. [6]. Observation modes that will assist in cancer screening are being developed, but a definitive scheme has not yet been established. An experienced doctor is able to distinguish between cancer and non-cancer sites based on his or her experience and knowledge by observing the organs using the modes described above. This would open the way to developing schemes for high-accuracy cancer detection.

The remainder of this paper is organized as follows. In Section 2, we briefly explain a new-generation endoscope system based on lasers. In Section 3, we provide a simple explanation of the Daubechies wavelet transform, and in Section 4, we define the features of normal and the abnormal regions based on the Daubechies wavelet transform. In Section 5, of early esophageal cancer, and we propose a new method for detecting early esophageal cancer. Experimental results are presented in Section 6, and Section 7 concludes the paper.

2 New-Generation Endoscope System with Lasers

There are no diagnostic criteria for diagnosing the presence of cancer using an endoscope system that does not require dyeing to avoid associated side effects. Endoscope systems that can solve this problem have been under development in Refs. [1, 2]. In cancer screening using the current endoscope system of Fujifilm Corp., an endoscope specialist observes an organ while switching among four modes:

Normal (White light), FICE (Flexible spectral Imaging Color Enhancement), BLI (Blue LASER Imaging), and BLI-bright. In the endoscope system that Olympus Corp. has developed, an endoscope specialist observes an organ while switching between two modes: Normal and NBI (Narrow Band Imaging).

In the Normal mode, as shown in Fig. 1, the organ is irradiated with mainly intense white laser light (wavelength: $450nm \pm 10nm$), and is irradiated with weak narrow-band-imaging laser light (wavelength: $410nm \pm 10nm$), and the system displays this information on the monitor. This mode is very effective in a wide area of observation because it can represent the structure of the organ as a very bright image. However, since the output data has no color information other than red, an inexperienced doctor has difficulty in performing diagnosis without overlooking an early cancer. Of course, an inexperienced doctor cannot distinguish between cancer and non-cancer, and it is impossible to find abnormal regions. Therefore, cancers such as early esophageal cancer and poorly differentiated gastric cancer are sometimes overlooked in cancer screening.

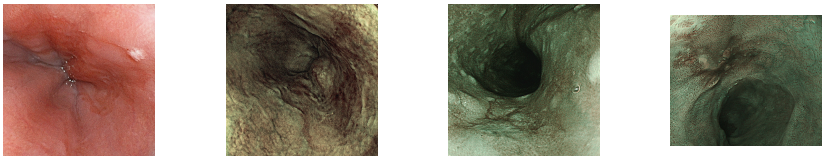


Fig. 1 An endoscopic image including a region of early esophageal cancer: Normal, FICE, BLI and NBI modes (from left to right).

As shown in Fig. 1, the FICE mode makes it easy to see the shape of blood vessels and the surface structure of the mucosa by performing color enhancement processing by extracting a spectral image composed of any wavelength based on the information obtained in the Normal mode. The surface structure of the gastric mucosa can be visualized by clearly emphasizing the abnormal regions but not normal regions. However, particularly when the magnification is high, this approach is not effective for visualizing microvessels in the tumor surface.

In the BLI and BLI-bright modes, as shown in Fig. 1, the intensity setting of the laser beam is different between the Normal and FICE modes. In this case, the site under observation is irradiated mainly with intense narrow-band-imaging laser light, to focus on a region of interest in the organ being examined using a short wavelength and a narrow spectral width. There is some work involved in characterizing abnormal regions. In addition, white laser light plays an important role as white-light illumination having a wide spectral width suitable for Normal observation because of its weak irradiation intensity. In other words, we can observe bright, high-contrast information such as the shape of blood vessels and the surface structure of the mucosa. Endoscopy specialists diagnose the presence of cancer by observing the video and images obtained in these modes.

As shown in Fig. 1, the NBI mode can highlight the capillaries in the mucosal surface and fine patterns in the mucosa by irradiating the mucosa with two narrow-band light beams having wavelengths that are easily absorbed in hemoglobin in the blood. This mode utilizes the following properties in order to observe blood vessels with high contrast. Blue-narrow band light (wavelength: 390–445nm) is strongly absorbed in blood, and this property is useful for observing the capillaries in the mucosal surface. Green narrow-band light (wavelength: 530–550nm) is strongly reflected and scattered at the mucosal surface, and this property is useful for enhancing the contrast between deep, thick blood vessels and the capillaries in the mucosal surface. See Refs. [1, 2] for more details regarding the features of these modes.

3 Daubechies Wavelet Transform

Let us represent the original image by $I[m, n]$, which are samples of a normalized discrete image defined on a two-dimensional lattice of finite extent. It is well-known from Ref. [4] that the usual discrete wavelet transform (DWT) with Daubechies wavelets for images with support width $2N$ is given by:

$$C[i,\, j] = \sum_{m,n=0}^{2N-1} p_m p_n I[m+2i, n+2j],\;\; D[i,\, j] = \sum_{m,n=0}^{2N-1} p_m q_n I[m+2i, n+2j],$$

$$E[i,\, j] = \sum_{m,n=0}^{2N-1} q_m p_n I[m+2i, n+2j],\;\; F[i,\, j] = \sum_{m,n=0}^{2N-1} q_m q_n I[m+2i, n+2j],$$

where p_n and q_n are real parameters which have the relation $q_n = (-1)^n p_{2N-1-n}$. The indices i and j indicate the locations in the horizontal and vertical directions, respectively. More precisely, $C[i,\, j]$, $D_1[i,\, j]$, $D_2[i,\, j]$, and $D_3[i,\, j]$ are the low-frequency components, the horizontal high-frequency components, the vertical high-frequency components, and the high-frequency components in both directions, respectively. These Figures are not illustrated in the present paper due to space limitations.

4 Preliminary Experiment

As a preliminary experiment, we compared the features extracted from the histograms of low-frequency components and high-frequency components for normal and abnormal regions to distinguish these regions.

The histograms are graphs showing the distribution of the luminance values in the image. The horizontal axis of the graph represents the luminance value, and the vertical axis represents the number of pixels having that luminance value.

1. Convert the input image obtained in the FICE mode into CIEL*a*b* color space. In this experiment, we use the a* component.
2. Divide the a* component into small non-overlapping blocks with a size of 64×64 pixels.
3. Apply the DWT to each block.
4. Compare the features of the histograms at blocks including normal and abnormal regions.

We use the db3 to extract the low-frequency components which reflect the feature of the db3.

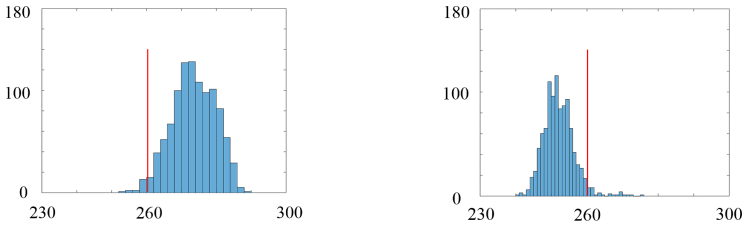


Fig. 2 Histograms of low-frequency components in blocks at the normal (left) and abnormal (right) regions.

Fig. 2 shows histograms of the low-frequency components in the blocks at normal and abnormal regions. As shown in Fig. 2, we found that the histogram of the normal regions tends to be positioned at the left of a particular value, and the histogram of the abnormal regions tends to be positioned at the right of a particular value. This fact means that there is a certain threshold which enable us to distinguish the normal and abnormal regions, and we describe how to determine a threshold in the next section.

We also found a difference with respect to the longer tails of the histograms of the high-frequency component at the normal and abnormal regions. We use the db9 to extract the high-frequency components which reflect the feature of the db9. This is because the chances of detecting a feature of cancer when the support length is short are low. In fact, we compared the preliminary experiment using db3, but obtained better the feature of the longer tails of the histogram when using db9.

From the preliminary experimental results, we found that the tails of the histograms of the abnormal regions tend to be longer than those of the normal regions. We define that longer tails of the histogram subtracts the minimum value from the maximum value of each blocks. In Fig. 3, the tail lengths of the histograms of the normal region and the abnormal region are eight and thirteen, respectively.

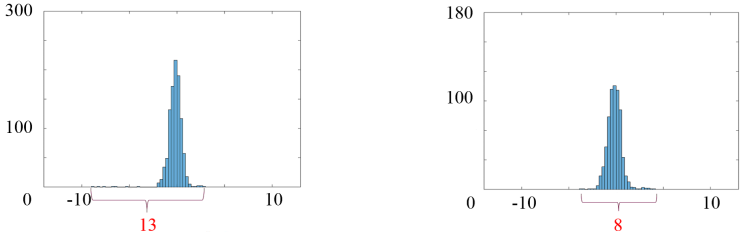


Fig. 3 Histograms of high-frequency components in blocks at the normal (left) and abnormal (right) regions.

5 Image Enhancement Algorithm for Cancer Detection

1. Convert the input image of FICE mode into CIEL*a*b* color space. In this experiment, we use a* component.
2. Clip the a* component to 640 × 640 pixels and divide the clipped a* component into 64 × 64-pixel blocks. (In this experiment, the image will be divided into 100 blocks, as shown in Fig. 4.)

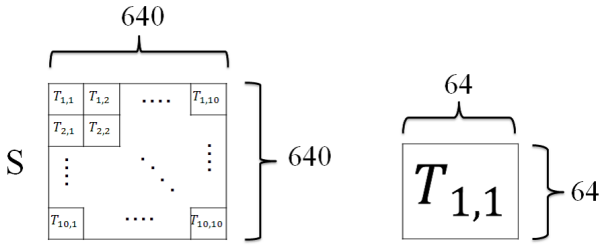


Fig. 4 Example of the block configuration in the image.

3. a. Apply the DWT with db3 to each block, and obtain the low-frequency components, which are 32 × 32 pixels.
   b. Apply the DWT with db9 to each block, and obtain the high-frequency component D and E, which are 32 × 32 pixels.
4. a. Using the test image that includes about half normal and half abnormal regions, compute the mean value μ_{test} for the pixel values of the test image and its standard deviation σ_{test}.
   b. Compute the representative value of the low-frequency component, $S^{low}[m, n]$, using the following relations (see Fig. 5):

$$S_{m,n}^{low} = \sum_{i=1}^{32} \sum_{j=1}^{32} \tilde{T}_{m,n}^{low}[i, j],$$

$$\tilde{T}^{low}_{m,n}[i,j] = \begin{cases} 0, & T^{low}_{m,n}[i,j] < \mu_{test} \\ 1, & \mu_{test} \le T^{low}_{m,n}[i,j] < \mu_{test}+2\sigma_{test} \\ 3, & \mu_{test}+2\sigma_{test} \le T^{low}_{m,n}[i,j], \end{cases}$$

where $T_{m,n}$ is a target block. The indices m and n indicate the locations in the horizontal and vertical directions of each blocks.

c. Compute representative value of the each block that $S^{high}_{m,n}$ to the following (component D and E):

$$S^{high}_{m,n} = \tilde{T}^{D,high}_{m,n} + \tilde{T}^{E,high}_{m,n},$$

where $\tilde{T}^{high}_{m,n}$ is obtained by subtracting the minimum value from the maximum value of each block.

5. Normalize the enhancement images S^{low} and S^{high} to the closed interval [0,255].
6. Compute the product of S^{low} and S^{high} in each block by the following equation, and normalize.

$$S_{m,n} = S^{high}_{m,n} \cdot S^{low}_{m,n}.$$

We can detect the abnormal region from endoscopic image using the enhancement image.

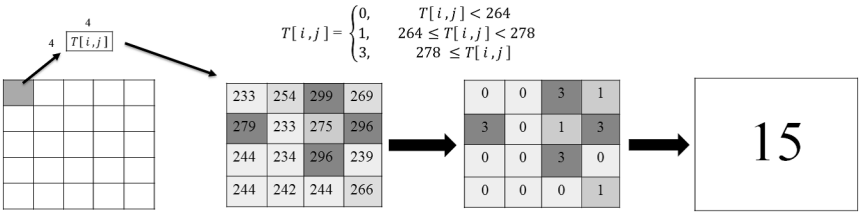


Fig. 5 An example of computing $S^{low}[m,n]$ for 4 × 4 -pixel low-frequency component.

We would like to describe the reason why the block size is 64 × 64. Fig. 6 shows an endoscopic image with the region of the esophageal cancer marked by a doctor, its enhancement images based on 32 × 32-pixel blocks, 64 × 64-pixel blocks and 128 × 128-pixel blocks. Adopting 32 × 32-pixel blocks, the false detection rate is relatively high. On the other hand, adopting 128 × 128-pixel blocks, the detection rate is getting worse. Therefore, we adopt the 64 × 64-pixel blocks.

6 Experimental results

700 × 700 pixels and 750 × 750 pixels, which are used in the Department of Internal Medicine, Saga University, Japan.

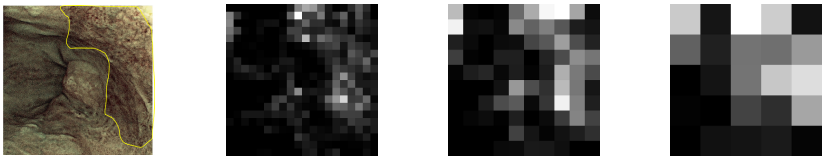


Fig. 6 An endoscopic image with the region of the esophageal cancer marked by a doctor, its enhancement images based on 32 × 32-pixel blocks, 64 × 64-pixel blocks and 128 × 128-pixel blocks (from left to right).

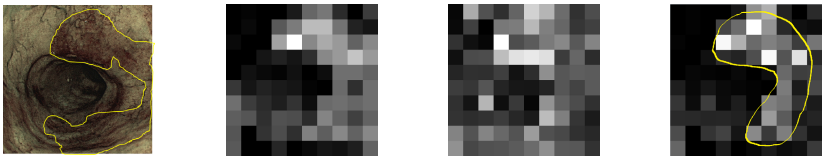


Fig. 7 Endoscopic image of patient A, with the region of the esophageal cancer region marked by a doctor, feature of low-frequency components, feature of high-frequency components, detection result with our method (from left to right).

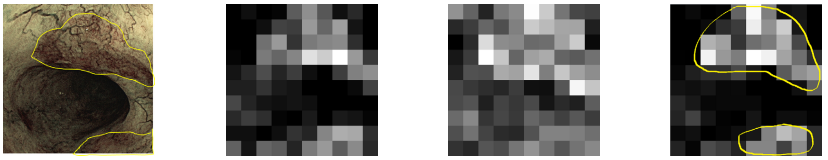


Fig. 8 Endoscopic image of patient A, with the region of the esophageal cancer region marked by a doctor, feature of low-frequency components, feature of high-frequency components, detection result with our method (from left to right).

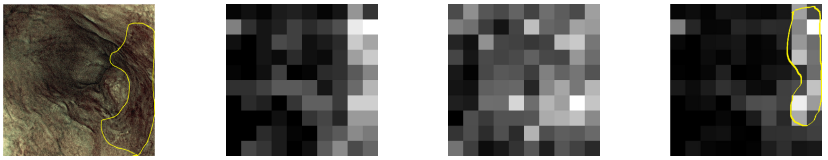


Fig. 9 Endoscopic image of patient B, with the region of the esophageal cancer region marked by a doctor, feature of low-frequency components, feature of high-frequency components, detection result with our method (from left to right).

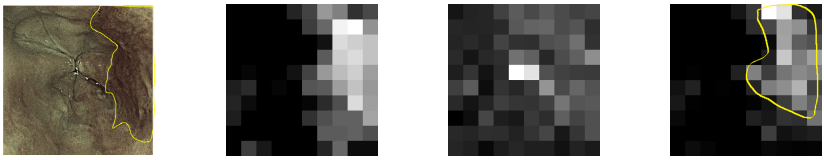


Fig. 10 Endoscopic image of patient C, with the region of the esophageal cancer region marked by a doctor, feature of low-frequency components, feature of high-frequency components, detection result with our method (from left to right).

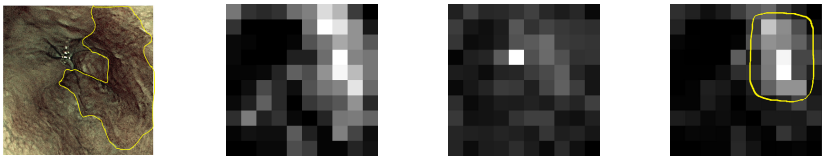


Fig. 11 Endoscopic image of patient C, with the region of the esophageal cancer region marked by a doctor, feature of low-frequency components, feature of high-frequency components, detection result with our method (from left to right).

In Figs. 7–11, we found that a common region exists by comparing the image with a cancer region marked by a doctor and the feature images of the low- and high-frequency components obtained by the proposed method. The detection result with our method is marked in yellow based on the difference in the luminance value. The abnormal region marked by a doctor was detected in a wide area by our method. In addition, we made sure that the marked region could be accurately detected in the image generated by taking the product of each feature image for the low- and high-frequency components. Therefore, early esophageal cancer is suspected in the case where the histogram of the low-frequency components is positioned at the right of a particular value and the tail of the histogram of the high-frequency component is longer, and it is thought that this can be detected due to the difference in surface structure. We could also accurately detect the region of the cancer by using two characteristics, that is, the characteristics of the low-frequency components and the characteristic of the high-frequency components.

7 Conclusion

From FICE endoscopic images based on the Daubechies wavelet. The experimental results show that our method can detect an abnormal region, and can be used for assisting experts in cancer diagnosis. Our method shows the usefulness of the FICE mode and can visualize abnormal regions of early esophageal cancer, allowing them to be detected by inexperienced doctors.

From the experimental result in Section 7, we want to develop a better feature of the high-frequency components using other wavelets, such as lifting wavelet and multiwavelets, which have different properties in terms of their wavelets.

Acknowledgements Thanks are due to Dr. Sakata Yasuhisa and other staff in the Department of Internal Medicine, Saga University, Japan, for helpful suggestions and comments. Their medical images and specialist advice were the motivation for us to attempt this study. This work was partially supported by JSPS KAKENHI Grant Number 26400206.

References

1. Gono, K., Obi, T., Yamaguchi, M., Ohyama, N., Machida, H., Sano, Y., Yoshida, S., Hamamoto, Y., Endo, T.: Appearance of enhanced tissue features in narrow-band endoscopic imaging. Journal of Biomedical Optics **9**(3), 568–577 (2004)
2. Morimoto, Y., Kubo, M., Kuramoto, M., Yamaguchi, H. and Kaku, T.: Development of a New Generation Endoscope System with Lasers “LASEREO". FUJIFILM Research & Development, No.58, 6 pages (2013)
3. Yamaguchi, J., Yoneyama, A. and Minamoto, T.: Automatic detection of early esophageal cancer from endoscope image using fractal dimension and discrete wavelet transform. In: ITNG 2015: Proceedings of the 2015 12th International Conference on Information Technology: New Generations, pp. 317–322. IEEE Computer Society (2015)
4. Daubechies, I. : Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics (1992)
5. Patent office: 2014 fiscal year Patent Application Technology Trends Survey Report. https://www.jpo.go.jp/shiryou/pdf/gidou-houkoku/26_1.pdf
6. Toru, T., Junki, Y., Misato, K., Bisser, R., Kazufumi, K., Shigeto, Y., Yoshito, T., Keiichi, O., Rie, M., Shinji, T.: Computer-aided colorectal tumor classification in NBI endoscopy using local features. Medical Image Analysis **17**, 78–100 (2013)
7. Horimatsu, T., Ezoe, Y., Morita, S. and Miyamoto, S.: Improving visualization techniques by narrow band imaging and magnification endoscopy. J Gastroenterol Hepatol, pp. 1333–1346 (2009)
8. Paggi, S., Rondonotti, E., Amato, A., Fuccio, L., Andrealli, A., Spinzi, G., Radaelli, F.: Narrow-band imaging in the prediction of surveillance intervals after polypectomy in community practice. Source of the Document Endoscopy **47**(9), 808–814 (2015)

A Refined Method for Estimating the Global Hölder Exponent

S. Nicolay and D. Kreit

Abstract In this paper, we recall basic results we have obtained about generalized Hölder spaces and present a wavelet characterization that holds under more general hypothesis than previously stated. This theoretical tool gives rise to a method for estimating the global Hölder exponent which seems to be more precise than other wavelet-based approaches. This work should prove helpful for estimating long range correlations.

Keywords Wavelets · Uniform Hölder exponent · Generalized Hölder spaces · Long range correlations · Brownian motion

1 Introduction

A continuous function $f \in L^\infty(\mathbf{R}^d)$ belongs to the uniform Hölder space $\Lambda^\alpha(\mathbf{R}^d) = \Lambda^\alpha$ with $\alpha > 0$ if there exists a constant C such that for each $x \in \mathbf{R}^d$, there exists a polynomial P_x of degree at most α for which

$$|f(x+h) - P_x(h)| \leq C|h|^\alpha. \tag{1}$$

Since these spaces are embedded, one can define the Hölder exponent as follows,

$$H_f = \sup\{\alpha : f \in \Lambda^\alpha\}.$$

The Hölder exponent of a function is a notion of global regularity: the larger the Hölder exponent, the more regular the corresponding function. In particular, if f is n

S. Nicolay(✉) · D. Kreit
Department of Mathematics, University of Liège,
12 Allée de la Découverte, B37, 4000 Liège, Belgium
e-mail: S.Nicolay@ulg.ac.be

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_82

times continuously differentiable then f belongs to Λ^α for any $\alpha \leq n$. For example, a sample path of a Brownian motion W belongs to $\Lambda^{1/2}$ almost surely and $H_W = 1/2$ almost surely. More generally, the sample path of a fractional Brownian motion of Hurst index H ($H \in (0, 1)$) belongs to Λ^H almost surely and the associated Hölder exponent is equal to H almost surely (a Brownian motion is a fractional Brownian motion with Hurst index $H = 1/2$) [11]. Figure 1 clearly illustrates the fact that the regularity increases with the Hölder exponent.

Among methods for estimating the Hölder exponent, the wavelet-based approach [9] is both fast and relatively efficient. It thus allows to test if whether or not a fractional Brownian motion displays long range correlations (i.e. is associated to a Hurst exponent $H > 1/2$). However, such a method for estimating the Hölder exponent cannot allow to make the distinction between say a Brownian motion and a process displaying the same Hölder exponent. Yet, the Brownian motion W exhibits a very specific behavior, since one has

$$|W(t+h) - W(t)| \leq C\sqrt{|h| \log |\log |h||}, \tag{2}$$

almost surely, for a constant $C > \sqrt{2}$ [5].

The purpose of this paper is to provide a method that could help to make the distinction between a process satisfying inequality (2) and a process belonging to $\Lambda^{1/2}$ for which inequality (2) is not satisfied. The basic idea is to generalize the spaces Λ^α by replacing the the expression $|h|^\alpha$ appearing in the right-hand side of (1) with something more general, covering the usual case. Such spaces are both inspired by generalized Besov spaces (see e.g. [10]) and moduli of continuity for Hölder spaces [4].

This paper is organized as follows: we first review the notion of generalized Hölder space based on admissible sequences, as introduced in [6, 7]. Next, we give a wavelet characterization of these spaces. Finally, we propose an application for better estimating the Hölder exponent of a Brownian motion, which should help to make the distinction between a Brownian motion and another process associated to the same Hölder exponent. This theorem as well as the application are new and generalize previous results (see [7, 9]).

Throughout the paper, B denotes the open unit ball and we use the letter C for generic positive constant whose value may be different at each occurrence.

2 Definitions

All the definitions and results presented in this section are taken from [6, 7]; the reader is referred to these articles for further details.

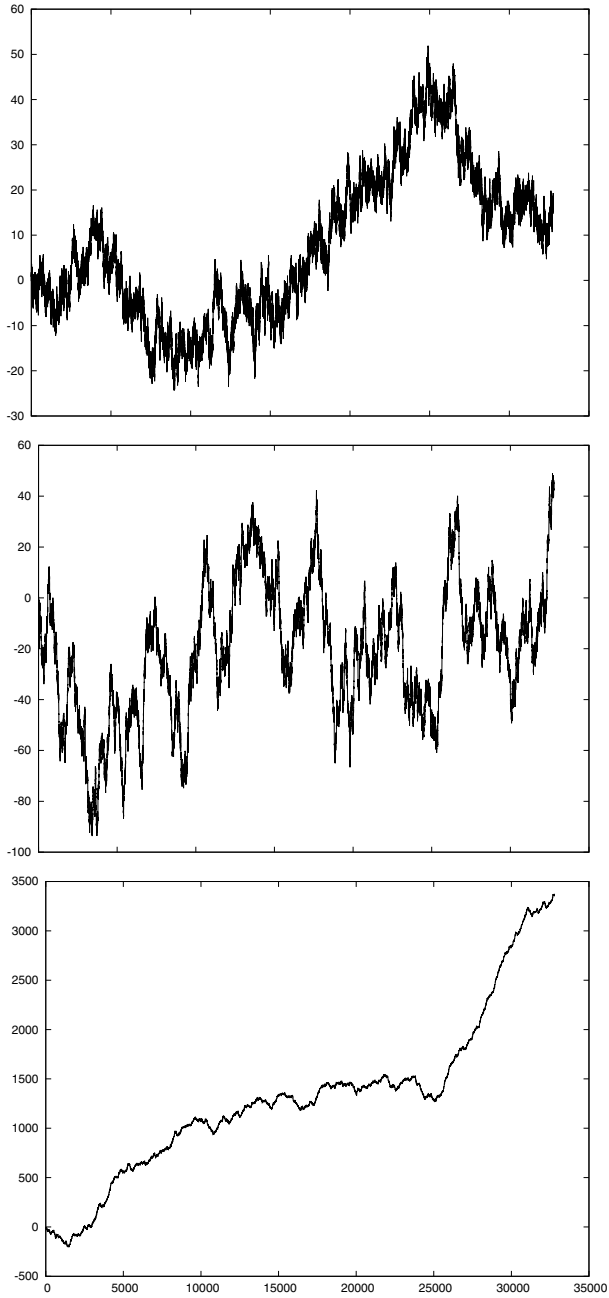


Fig. 1 A fractional Brownian motion with Hurst index $H = 0.3$ (top), a Brownian motion, i.e. a fractional Brownian motion with Hurst index $H = 0.5$ (middle) and a fractional Brownian motion with Hurst index $H = 0.7$ (bottom). One clearly sees that the regularity of the walk increases with the Hurst index.

2.1 Admissible Sequences

The notion of generalized Hölder space we want to introduce is based on the following definition.

Definition 1. A sequence $\sigma = (\sigma_j)_{j\in\mathbf{N}}$ of real positive numbers is called admissible if there exists a positive constant C such that

$$C^{-1}\sigma_j \leq \sigma_{j+1} \leq C\sigma_j,$$

for any $j \in \mathbf{N}$.

If σ is such a sequence, we set

$$\underline{\Theta}_j = \inf_{k\in\mathbf{N}} \frac{\sigma_{j+k}}{\sigma_k} \qquad \text{and} \qquad \overline{\Theta}_j = \sup_{k\in\mathbf{N}} \frac{\sigma_{j+k}}{\sigma_k}$$

and define the lower and upper Boyd indices as follows,

$$\underline{s}(\sigma) = \lim_j \frac{\log_2 \underline{\Theta}_j}{j} \qquad \text{and} \qquad \overline{s}(\sigma) = \lim_j \frac{\log_2 \overline{\Theta}_j}{j}.$$

Since $(\log \overline{\Theta}_j)_{j\in\mathbf{N}}$ is a subadditive sequence, such limits always exist. If σ is an admissible sequence, let $\varepsilon > 0$; there exists a positive constant C such that

$$C^{-1}2^{j(\underline{s}(\sigma)-\varepsilon)} \leq \underline{\Theta}_j \leq \frac{\sigma_{j+k}}{\sigma_k} \leq \overline{\Theta}_j \leq C2^{j(\overline{s}(\sigma)+\varepsilon)},$$

for any $j, k \in \mathbf{N}$.

The following result allows to generate admissible sequences from existing admissible sequences:

Proposition 1. If τ and υ are two admissible sequences and α is a real number, then

- The sequence $\sigma = (2^{-j\alpha})_{j\in\mathbf{N}}$ is an admissible sequence with $\underline{s}(\sigma) = \overline{s}(\sigma) = -\alpha$.
- Let $\varphi : [0, 1] \to (0, \infty)$ be a weakly varying function, i.e. a function satisfying

$$\lim_{x\to 0} \frac{\varphi(rx)}{\varphi(x)} = 1,$$

  for any $r > 0$. The sequence $\sigma = (2^{-\alpha j}\varphi(2^j))_{j\in\mathbf{N}}$ is an admissible sequence such that $\underline{s}(\sigma) = \overline{s}(\sigma) = -\alpha$.
- The sequences $\tau + \upsilon$ and $\tau\upsilon$ are admissible sequences (with $\underline{s}(\tau\upsilon) \geq \underline{s}(\tau) + \underline{s}(\upsilon)$ and $\overline{s}(\tau\upsilon) \leq \overline{s}(\tau) + \overline{s}(\upsilon)$)
- If $\alpha > 0$, $\alpha\tau$ is an admissible sequence (with $\underline{s}(\alpha\tau) = \underline{s}(\tau)$ and $\overline{s}(\alpha\tau) = \overline{s}(\tau)$).
- If $\alpha > 0$, τ^α is an admissible sequence such that $\underline{s}(\tau^\alpha) = \alpha\underline{s}(\tau)$ and $\overline{s}(\tau^\alpha) = \alpha\overline{s}(\tau)$.

– If $\alpha < 0$, τ^α is an admissible sequence such that $\underline{s}(\tau^\alpha) = \alpha\overline{s}(\tau)$ and $\overline{s}(\tau^\alpha) = \alpha\underline{s}(\tau)$.

For example, $\varphi = |\log|$ is a weakly varying function.

The following result gives informations about the convergence of the series associated to an admissible sequence.

Proposition 2. Let σ be an admissible sequence.

– If $\underline{s}(\sigma^{-1}) > 0$, there exists a positive constant C such that for any $J \in \mathbf{N}$,

$$\sum_{j \geq J} \sigma_j \leq C\sigma_J.$$

– If $\overline{s}(\sigma^{-1}) < n$ with $n \in \mathbf{N}$, there exists a positive constant C such that for any $J \in \mathbf{N}$,

$$\sum_{j=1}^{J} 2^{jn}\sigma_j \leq C2^{Jn}\sigma_J.$$

2.2 Generalized Hölder Spaces

We can now introduce a definition of generalized Hölder space.

As usual, $[\alpha]$ will denote the greatest integer lower than α,

$$[\alpha] = \sup\{p \in \mathbf{Z} : p \leq \alpha\}$$

and $\Delta_h^n f$ will stand for the finite difference of order n: given a function f defined on $\mathbf{R}^d$ and $x, h \in \mathbf{R}^d$, $\Delta_h^1 f(x) = f(x+h) - f(x)$ and

$$\Delta_h^{n+1} f(x) = \Delta_h^1 \Delta_h^n f(x),$$

for any $n \in \mathbf{N}$.

Definition 2. Let $\alpha > 0$ and σ be an admissible sequence; a function $f \in L^\infty(\mathbf{R}^d)$ belongs to the space $\Lambda^{\sigma,\alpha}(\mathbf{R}^d) = \Lambda^{\sigma,\alpha}$ if there exists $C > 0$ such that

$$\sup_{|h| \leq 2^{-j}} \|\Delta_h^{[\alpha]+1} f\|_\infty \leq C\sigma_j,$$

for any $j \in \mathbf{N}$.

One sets $\Lambda^\sigma = \Lambda^{\sigma,\overline{s}(\sigma^{-1})}$. The application

$$|f|_{\Lambda^{\sigma,\alpha}} = \sup_j (\sigma_j^{-1} \sup_{|h| \leq 2^{-j}} \|\Delta_h^{[\alpha]+1} f\|_{L^\infty})$$

defines a semi-norm on $\Lambda^{\sigma,\alpha}$ and therefore $\|f\|_{\Lambda^{\sigma,\alpha}} = \|f\|_{L^\infty} + |f|_{\Lambda^{\sigma,\alpha}}$ is a norm on this space. We have the following result.

Theorem 3. Let $\alpha > 0$ and σ be an admissible sequence; the space $\Lambda^{\sigma,\alpha}$ equipped with the norm $\|\cdot\|_{\Lambda^{\sigma,\alpha}}$ is a Banach space.

Many equivalent definitions of theses spaces can be given. In particular, they can be written using polynomials. Let us denote the set of polynomials of degree at most n by $\mathbf{P}_n$.

Theorem 4. Let σ be an admissible sequence, $\alpha > 0$ and $f \in L^\infty(\mathbf{R}^d)$; f belongs to $\Lambda^{\sigma,\alpha}$ if and only if there exists a constant $C > 0$ such that

$$\inf_{P\in\mathbf{P}_{[\alpha]}} \|f - P\|_{L^\infty(2^{-j}B+x_0)} \leq C\sigma_j, \tag{3}$$

for any $x_0 \in \mathbf{R}^d$ and any $j \in \mathbf{N}$.

For example a sample path of a Brownian motion belongs to Λ^σ with

$$\sigma = (2^{-j/2}\sqrt{\log j})_j \tag{4}$$

almost surely.

One also can show that these spaces are related to the regularity of their elements. For example, we have the following result:

Theorem 5. Let σ be an admissible sequence and N, M be two positive integers such that

$$N < \underline{s}(\sigma^{-1}) \leq \overline{s}(\sigma^{-1}) < M.$$

Any element of Λ^σ is equal almost everywhere to a function $f \in C^N(\mathbf{R}^d)$ satisfying $D^\beta f \in L^\infty(\mathbf{R}^d)$ for any multi-index β such that $|\beta| \leq N$ and

$$\sup_{|h|\leq 2^{-j}} \|\Delta_h^{M-|\beta|} D^\beta f\|_\infty \leq C2^{j|\beta|}\sigma_j, \tag{5}$$

for any $j \in \mathbf{N}$ and $|\beta| \leq N$. Conversely, if $f \in L^\infty(\mathbf{R}^d) \cap C^N(\mathbf{R}^d)$ satisfies inequality (5) for $|\beta| = N$, then $f \in \Lambda^\sigma$.

3 Wavelet Characterization

3.1 Definitions

Under some general assumptions (see e.g. [2, 9]), there exists a function φ and $2^d - 1$ functions $(\psi^{(i)})_{1\leq i<2^d}$ called wavelets such that

$$\{\varphi(x-k)\}_{k\in\mathbf{Z}^d} \cup \{\psi^{(i)}(2^j x-k) : 1 \le i < 2^d, k \in \mathbf{Z}^d, j \in \mathbf{N}_0\}$$

form an orthogonal basis of $L^2(\mathbf{R}^d)$. Any function $f \in L^2(\mathbf{R}^d)$ can be decomposed as follows,

$$f(x) = \sum_{k\in\mathbf{Z}^d} C_k \varphi(x-k) + \sum_{j=0}^{\infty} \sum_{k\in\mathbf{Z}^d} \sum_{1\le i<2^d} c_{j,k}^{(i)} \psi^{(i)}(2^j x-k),$$

where

$$c_{j,k}^{(i)} = 2^{dj} \int_{\mathbf{R}^d} f(x)\psi^{(i)}(2^j x-k)\,dx$$

and

$$C_k = \int_{\mathbf{R}^d} f(x)\varphi(x-k)\,dx.$$

The above formulas are still valid in more general settings; they have to be interpreted as a duality product between regular functions (the wavelets) and distributions [3, 9]. In what follows, we will suppose that the wavelets are the Daubechies wavelets [2] (which are compactly supported and can be chosen arbitrarily regular, let us say r-regular with $r > \bar{s}(\sigma^{-1})$).

3.2 The Characterization

We aim at showing the following result.

Theorem 6. Let σ be an admissible such that $\underline{s}(\sigma^{-1}) > 0$. If f belongs to Λ^σ, there exists $C > 0$ such that

$$|C_k| \le C \qquad \text{and} \qquad |c_{j,k}^{(i)}| \le C\sigma_j, \tag{6}$$

for any $j \in \mathbf{N}$, any $k \in \mathbf{Z}^d$ and any $i \in \{1, \dots, 2^d - 1\}$.

Conversely, if $f \in L^\infty_{\mathrm{loc}}(\mathbf{R}^d)$ and (6) holds, then $f \in \Lambda^\sigma$.

Other hypothesis are given in [7], but the admissible sequence is requested to be strong.

Proof. Let $f \in \Lambda^\sigma$ and let us prove that (6) holds. Let $M \in \mathbf{N}$ and $j_0 \in \mathbf{N}$ be such that $M > \bar{s}(\sigma^{-1})$ and $\mathrm{supp}\psi^{(i)} \subset 2^{j_0} B$ for any i. We have

$$|C_k| = |\int f(x)\varphi(x-k)\,dx| \le C\|f\|_{L^\infty}.$$

Using Theorem 4, for $k \in \mathbf{Z}^d$ and $j \ge j_0$ one can find a polynomial P of degree less or equal to $M-1$ such that

$$\|f(\cdot) - P(\cdot - k/2^j)\|_{L^\infty(B')} \leq C\sigma_{j-j_0},$$

where $B' = k2^{-j} + 2^{-(j-j_0)}B$. One gets

$$\begin{aligned}
|c_{j,k}^{(i)}| &= 2^{jd}\Big|\int f(x)\psi^{(i)}(2^j x - k)\,dx\Big| \\
&= 2^{jd}\Big|\int_{B'} \big(f(x) - P(x - k/2^j)\big)\,\psi^{(i)}(2^j x - k)\,dx\Big| \\
&\leq C\sigma_j \sup_{i\in\{1,\dots,2^d-1\}} \|\psi^{(i)}\|_{L^1(\mathbf{R}^d)}.
\end{aligned}$$

Let us now suppose that (6) holds. We need to show that the hypothesis of Proposition 5 are satisfied. Let $N, M \in \mathbf{N}_0$ be such that

$$N < \underline{s}(\sigma^{-1}) \leq \overline{s}(\sigma^{-1}) < M \leq r$$

and let us define

$$f_{-1}(x) = \sum_{k\in\mathbf{Z}^d} C_k\varphi(x-k) \quad \text{and} \quad f_j(x) = \sum_{i=1}^{2^d-1}\sum_{k\in\mathbf{Z}^d} c_{j,k}^{(i)}\psi^{(i)}(2^j x - k),$$

for any $j \in \mathbf{N}_0$. One easily checks that f_j converges uniformly on every compact set, so that f_j has the same regularity as $\psi^{(i)}$. For every j, we have $|f_j| \leq C\sigma_j$ so that $g = \sum_{j=-1}^{\infty} f_j$ uniformly converges to a function belonging to $L^\infty(\mathbf{R}^d)$. One thus gets $f = g$. Using the properties of φ and $\psi^{(i)}$, one gets

$$|D^\beta f_j| \leq C2^{|\beta|j}\sigma_j,$$

for any multi-index β such that $|\beta| \leq M$. Consequently, g is differentiable up to order N, which shows that we have $f \in C^N(\mathbf{R}^d)$ and $|D^\beta f| \leq C$ for any β such that $|\beta| \leq N$. Let α be a multi-index such that $|\alpha| = N$ and let $h \in \mathbf{R}^d$ and $j_0 \in \mathbf{N}$ be such that $|h| < 2^{-j_0}$. We have

$$\begin{aligned}
&\|\Delta_h^{M-N} D^\alpha f\|_{L^\infty} \\
&\leq \sum_{j\leq j_0} \|\Delta_h^{M-N} D^\alpha f_j\|_{L^\infty} + \sum_{j>j_0} 2^{M-N}\|D^\alpha f_j\|_{L^\infty} \\
&\leq C\sum_{j\leq j_0} |h|^{M-N} \sup_{|\beta|=M-N} \|D^{\alpha+\beta} f_j\|_{L^\infty} + C\sum_{j>j_0} 2^{Nj}\sigma_j \\
&\leq C|h|^{M-N}2^{Mj_0}\sigma_{j_0} + C2^{N(j_0+1)}\sigma_{j_0+1} \\
&\leq C2^{Nj_0}\sigma_{j_0},
\end{aligned}$$

hence the conclusion.

3.3 Usefulness in the Estimation of the Hölder Exponent

The usual version of Theorem 6 theoretically allows to estimate the uniform Hölder regularity of a function f by looking at the behavior of the wavelet coefficients versus the scales j. For the sake of simplicity, let us suppose that f belongs to $L^2([0,1])$ (this is by no mean a restriction, see e.g. [1]) and let Ψ_j denotes the set of wavelet coefficients at scale j. Let us suppose that $f \in \Lambda^H$; at each scale j, one computes the wavelet power spectrum of f (see e.g. [12]):

$$S_f(j) = \sqrt{\frac{1}{\#\Psi_j} \sum_{i,k} |c_{j,k}^{(i)}|^2}.$$

Following the standard wavelet characterization [9], one should have

$$S_f(j) \sim C\omega^{(H)}(2^{-j})$$

with $\omega^{(H)}(r) = r^H$. This implies

$$\log_2 S_f(j) \sim -Hj + C,$$

so that a log-log plot can be used to estimate the slope H. Another method consists in fitting a parametric curve $C\omega^{(H)}(2^{-j})$ to the function S_f. In this case, one can also apply Theorem 6 and modify $\omega^{(H)}$ to obtain a better fit. For the Brownian motion, having (2) or (4) in mind, one should choose

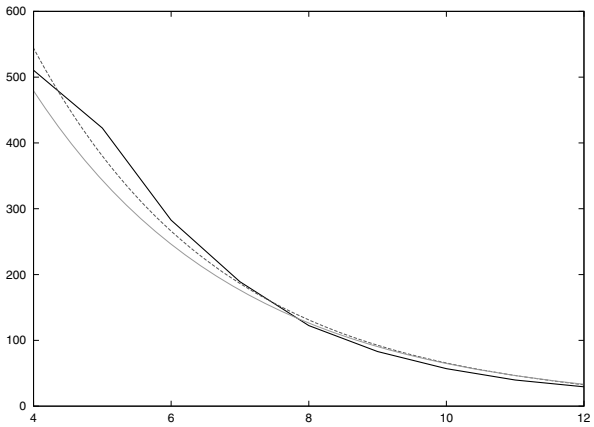


Fig. 2 The function S_W (thick black), with the curves $j \mapsto C2^{-jh}$ (grey) and $j \mapsto C\omega_W^{(H)}(2^{-j})$ (dashed lines) defined by equality (7). Both curves are obtained with the Levenberg-Marquardt algorithm.

$$\omega_W^{(H)}(r) = (r \log |\log r|)^H \tag{7}$$

in order to get a sharper estimation and help to discern between two models.

As an illustration, the wavelet power spectrum of a Brownian motion W (2^{15} points) is represented in Figure 2. When trying to fit the curve $C2^{-jh}$ to S_W using the Levenberg-Marquardt algorithm [8], one gets $h_0 = 0.48$ with (asymptotic) standard deviation $5\,10^{-2}$ (see Figure 2). The same computation with the curve $C\omega_W^{(H)}(2^{-j})$ gives $h_0 = 0.499$ with (asymptotic) standard deviation $3\,10^{-2}$, which is closer to the expected value $1/2$. Of course, additional work has to be done in order to suitably validate this method.

References

1. Dahmen, W., Prössdorf, S., Schneider, R.: Wavelet approximation methods for pseudodifferential equations: I Stability and convergence. Math. Z. **215**, 583–620 (1994)
2. Daubechies, I.: Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics (2006)
3. Jaffard, S.: Wavelet techniques in multifractal analysis. Proceedings of Symposia in Pure Mathematics **72**, 91–152 (2004)
4. Jaffard, S., Meyer, Y.: Wavelet methods for pointwise regularity and local oscillations of functions. Mem. Am. Math. Soc. **587**, 110 (1996)
5. Khintchine, A.: Über einen Satz der Wahrscheinlichkeitsrechnung. Fund. Math. **6**, 9–20 (1924)
6. Kreit, D., Nicolay, S.: Some characterizations of generalized Hölder spaces. Math. Nachr. **285**, 2157–2172 (2012)
7. Kreit, D., Nicolay, S.: Characterizations of the elements of generalized Hölder-Zygmund spaces by means of their representation. J. Approx. Theory **172**, 23–36 (2013)
8. Marquardt, D.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. Soc. Indust. Appl. Math. **11**(2), 431–441 (1963)
9. Meyer, Y.: Ondelettes et opérateurs : Ondelettes, vol. 1. Hermann (1990)
10. Moura, S.D.: On some characterizations of Besov spaces of generalized smoothness. Math. Nachr. **280**(9–10), 1190–1199 (2007)
11. Samoradnitsky, G., Taqqu, M.S.: Stable Non-Gaussian Random Processes. CRC Press (1994)
12. Torrence, C., Compo, G.P.: A Practical Guide to Wavelet Analysis. Bull. Amer. Meteor. Soc. **79**, 61–78 (1998)

A New Wavelet-Based Mode Decomposition for Oscillating Signals and Comparison with the Empirical Mode Decomposition

Adrien Deliège and Samuel Nicolay

Abstract We introduce a new method based on wavelets (EWMD) for decomposing a signal into quasi-periodic oscillating components with smooth time-varying amplitudes. This method is inspired by both the "classic" wavelet-based decomposition and the empirical mode decomposition (EMD). We compare the reconstruction skills and the period detection ability of the method with the well-established EMD on toys examples and the ENSO climate index. It appears that the EWMD accurately decomposes and reconstructs a given signal (with the same efficiency as the EMD), it is better at detecting prescribed periods and is less sensitive to noise. This work provides the first version of the EWMD. Even though there is still room for improvement, it turns out that preliminary results are highly promising.

Keywords Wavelets · Wavelet mode decomposition · Empirical mode decomposition · Continuous wavelet transform · ENSO index

1 Introduction

The aim of this paper is to provide and apply a new mode decomposition method using a wavelet-based approach. This decomposition has been introduced in [1]; we now improve it to develop a simple yet powerful mode decomposition method that we call empirical wavelet mode decomposition (EWMD). The basic idea is to extract successively "quasi-periodic" oscillating components from the continuous wavelet transform of the original signal. Each component is associated to a mean frequency but, unlike the Fourier transform, this decomposition does not lead to pure cosines: the amplitudes and frequencies slowly evolve through time. This allows to

A. Deliège(✉) · S. Nicolay
Department of Mathematics (B37), University of Liège,
Allée de la découverte 12, 4000 Liège, Belgium
e-mail: {adrien.deliege,s.nicolay}@ulg.ac.be

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_83

drastically decrease the number of terms needed to accurately rebuild the signal by taking into account only the terms carrying most of the information. By doing so, the reconstructed signal resolves the main variations of the original one without taking the noise into account. For a detailed description of the theory of continuous wavelet transforms and analogies with the Fourier transform, the reader is referred to e.g. [2, 3, 4]. Let us add that several wavelet-based mode decomposition methods have been developed in recent years (see e.g. [5] and references therein). We do not intend to write a review of these techniques nor to compare our results with these studies in the present paper.

In this work, we compare the skills of the EWMD with those of the famous empirical mode decomposition (EMD) (see e.g. [6, 7, 8]). This method has proven to be well-suited for the analysis of nonlinear and nonstationary signals, despite its lack of mathematical background. It allows to decompose accurately a signal into a finite (often small) number of modes, called intrinsic mode functions (IMFs) without leaving the time domain. The basic idea of the EMD is simple and consists in the following steps. First, compute $h = f - m$ where f is the signal and m is the mean of the upper and lower envelopes of f, then repeat the process with h instead of f until h is considered to be an IMF i_1. Compute $f_1 = f - i_1$ and repeat the whole process with f_1 instead of f to obtain the next IMF i_2, then compute $f_2 = f_1 - i_2$, etc. More details can be found in e.g. [6, 7, 8]. We found that the idea of extracting one component at a time, then subtracting it from the remaining data was appealing because it could reveal components that are overshadowed by the dominant modes of the signal. As explained in the next section, this idea combined with the wavelet decomposition provided in [1] led us to this first version of the EWMD. Usual criticism made about the EMD is its lack of solid theoretical background, the high sensitivity to noise, and the inability of separating modes with frequencies close to each other but with a large difference in amplitudes. Even though this mode-mixing problems and the robustness to noise have been improved (to the detriment of computational resources) in several revisions of the EMD, such as the Ensemble EMD (EEMD [9]) and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN [10]), the original EMD is not outdated and is still a good starting point for assessing the quality of a new mode decomposition method.

2 Method

In this section we describe step-by-step our empirical wavelet mode decomposition. The wavelet ψ used in this study is a Morlet-like wavelet with exactly one vanishing moment ([1]):

$$\psi(t) = \frac{e^{i\Omega t}}{2\sqrt{2\pi}} e^{-\frac{(2\Omega t+\pi)^2}{8\Omega^2}} \left(e^{\frac{\pi t}{\Omega}} + 1\right) \tag{1}$$

with $\Omega = \pi\sqrt{2/\ln 2}$. The Fourier transform of ψ is given by

$$\hat{\psi}(\omega) = \sin\left(\frac{\pi\omega}{2\Omega}\right) e^{\frac{-(\omega-\Omega)^2}{2}} . \tag{2}$$

The decomposition procedure explained below is largely inspired by the wavelet mode decomposition in [1] but has an added feature of the empirical mode decomposition: extracting only one component then subtracting it from the signal and iterating the process. The successive steps of the EWMD are the following.

a) Perform the continuous wavelet transform of the signal f:

$$Wf(a,t) = \int f(x)\bar{\psi}\left(\frac{x-t}{a}\right)\frac{dx}{a} \tag{3}$$

where ψ is the chosen wavelet, $\bar{\psi}$ is the complex conjugate of ψ, t stands for the time (position) parameter, and $a > 0$ is the scale parameter.

b) Compute the wavelet spectrum Λ associated to f:

$$\Lambda(a) = E\,|Wf(a,.)| \tag{4}$$

where E denotes the mean over time. Then look for the scale a^* at which Λ reaches its global maximum.

c) Extract the component related to a^*:

$$|Wf(a^*,t)|\cos(\arg Wf(a^*,t)) . \tag{5}$$

d) Subtract this component from f to get, say, f', and repeat steps (a) to (d) with f' instead of f.

e) For now, we stop the process when the extracted components are not relevant anymore. More precisely, in this paper, we decided to stop the process when $\Lambda(a^*) < 0.15$. A more adequate data-driven stopping criterion should still be found. The sum of the components successively extracted is an accurate reconstruction of f, though a slight vertical rescaling can be performed to minimize the root mean square error (RMSE) with f (this will also be done with the EMD).

If we denote by $c_1, c_2, ..., c_K$ the components successively extracted with the EWMD (i.e. these components are the counterparts of the IMFs of the EMD), then the signal $c_0 = f - \sum_{k=1}^{K} c_k$ is considered as the remaining noise and therefore the decomposition of f can be completed with this noisy component:

$$f = \sum_{k=0}^{K} c_k . \tag{6}$$

Similarly to the EMD, and contrary to the method presented in [1], where there is no iterative process, i.e the components are extracted from the local maxima of the spectrum, the procedure described above extracts one component at a time. This has

the advantage of unveiling components that are easily hidden by dominant modes. Also, the components successively extracted here are ordered following their energy level, while they are sorted according to their frequency with the EMD (which always gives a few noisy IMFs to start with). Consequently, more adequate data-driven stopping criteria could be based on the energy level of the component to extract.

Let us note that, in the following, the comparison between a reconstructed signal and the original one is made via two indicators: the RMSE between the signals and their Pearson product-moment correlation coefficient (PCC). Also, the reconstructed signals (with both the EMD and EWMD) will not be shown since they can barely be distinguished from the original ones with the naked eye.

3 Results

3.1 Accuracy of the Reconstruction

We first apply the EWMD on the two classic examples presented in [8] and for which the EMD gives accurate results. We show that our method is comparable to the EMD regarding the accuracy of the reconstruction of such signals. The first example consists of the sum of two frequency-modulated sinusoidal signals and a Gaussian wavepacket (see [8]). The signal is represented in Fig. 1 (left) with the successively extracted components with the EWMD. Though these components are

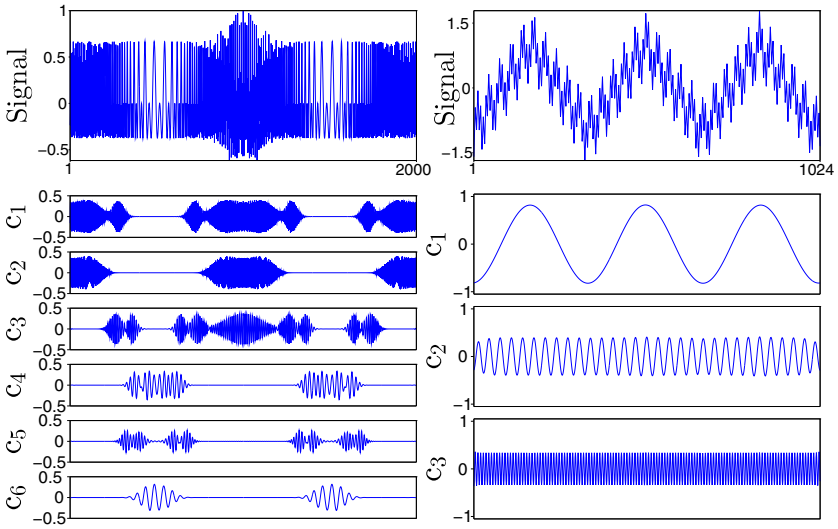


Fig. 1 Two signals analyzed with the EMD in [8] and the components extracted from the wavelet-based mode decomposition.

not the exact components of the original signal, they still manage to decompose it in an effective way. Indeed, when their sum is computed, the RMSE between the reconstructed signal and the original one is 0.086 and the PCC is 0.968. Such results are comparable to those obtained with the EMD, for which the RMSE is 0.07 and the PCC is 0.979. The second signal is the sum of a sinusoidal wave (constant frequency ω) and two triangular waveforms (one with a frequency larger than ω, the other smaller than ω). The signal and the components extracted are displayed in Fig. 1 (right). It can be seen that the three components are clearly recovered, though the triangular waveforms are somehow smoothed out. The RMSE and PCC related to the reconstructed signal compared to the original one are respectively 0.085 and 0.992, compared to 0.023 and 0.999 with the EMD. Although it seems that the EMD does a slightly better job at decomposing and reconstructing the signals, one can more than reasonably consider that the EWMD passed this test.

3.2 Period Detection Skills

One of the key features of both the EMD and wavelet decomposition is their ability to identify the periodicities hidden in a signal. In other words, they usually allow to recover the period of a sinusoidal signal as well as the mean period of an amplitude-modulated and frequency-modulated (AM-FM) signal (provided the frequency does not vary too much). Let us test the period detection skill of both methods on a toy example. We consider the signal $f(x) = \sum_{i=1}^{4} f_i(x)$ where

$$f_1(x) = \left(1 + 0.5\cos\left(\frac{2\pi}{200}x\right)\right)\cos\left(\frac{2\pi}{47}x\right) \tag{7}$$

$$f_2(x) = \frac{\ln(x)}{14}\cos\left(\frac{2\pi}{31}x\right) \tag{8}$$

$$f_3(x) = \frac{\sqrt{x}}{60}\cos\left(\frac{2\pi}{65}x\right) \tag{9}$$

$$f_4(x) = \frac{x}{2000}\cos\left(\frac{2\pi}{23 + \cos\left(\frac{2\pi}{1600}x\right)}x\right). \tag{10}$$

We thus have a signal made of three AM-components and one AM-FM component; the target periods to detect are $\approx 23, 31, 47, 65$ units. The signal and its components are plotted in Fig. 2. The successive steps (wavelet spectrum - period detection - component extraction - repeat the process) leading to the final decomposition are illustrated in Fig. 3. One can clearly notice that the extracted components match the original ones, confirming the fact that the EWMD is well-suited for that type of task. The RMSE between the reconstructed signal and f is 0.069 and the PCC is 0.996. Moreover, the periods detected by this method are 21.6, 30.6, 46.4 and 65.6 units, which are extremely close to the target periods ($\approx 23, 31, 47, 65$ units). On the other

hand, the IMFs obtained with the EMD are plotted in Fig. 4. One can clearly see that they do not recover the original ones as good as the components obtained from the EWMD; it can also be noticed that the first IMF on its own is extremely similar to the signal, having alone a PCC of 0.925. Also, even though the RMSE and PCC of the reconstructed signal are still excellent (resp. 0.068 and 0.998), the periods extracted from the Hilbert-Huang transform are of $\approx 41, 75, 165, 284$ units and are thus far from the expected ones.

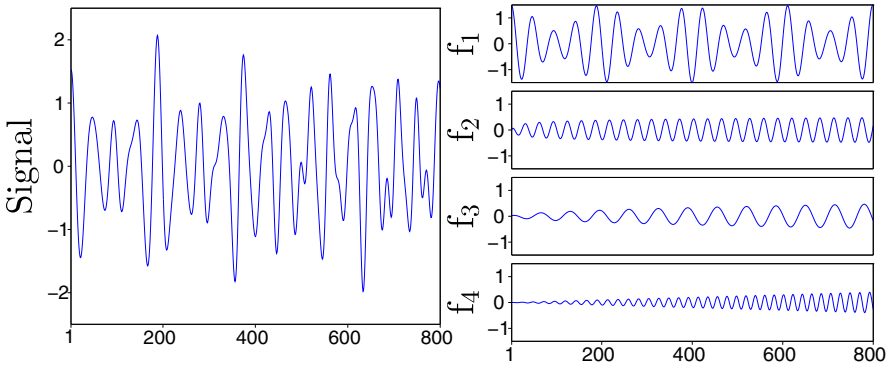


Fig. 2 The signal f (left) made of four amplitude-modulated and period-modulated components (right) used to compare the extraction and period detection skills of the EWMD and the EMD.

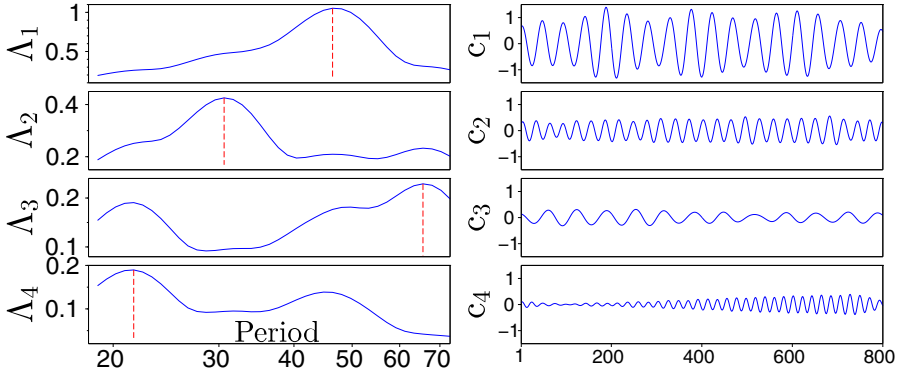


Fig. 3 Left: the wavelet spectra obtained from the successive wavelet transforms. The red dashed lines indicate where the global maximum is reached thus giving the period at which the corresponding component has to be extracted. Right: the components extracted from the wavelet transform, based on the associated wavelet spectra. The periods detected by the method are respectively (from top to bottom) 46.4, 30.6, 65.5 and 21.6 units (targets: 47, 31, 65, 23 units). The extracted components clearly match the original ones. The RMSE between the reconstructed signal and the original one is 0.069 and the PCC is 0.996.

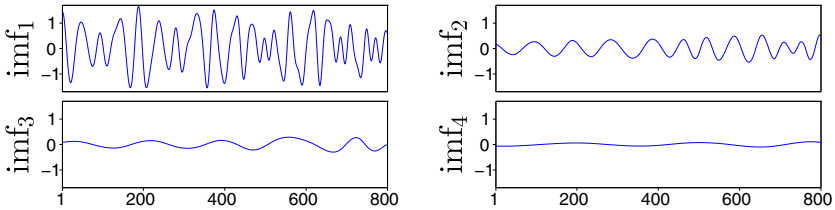


Fig. 4 The IMFs extracted from the EMD. One can see that they do not match accurately the original ones; the first IMF on its own is extremely similar to the whole signal. The periods extracted from the Hilbert-Huang transform (41, 75, 165, 284) are far from the target values ($\approx$ 23, 31, 47, 65 units).eps

3.3 Real-Life Data: ENSO Index

The EWMD is now applied to a real-life signal. For that purpose, the El Niño Southern Oscillation (ENSO) index is analyzed (ERSST.V3B SST Niño 3.4 time series provided by the Climate Prediction Center). It is a climate pattern consisting of monthly-sampled sea surface temperature anomalies (SSTA, in Celsius degrees) in the Eastern Pacific Ocean recorded from Jan. 1950 to Dec. 2014 (see Fig. 5). An anomalous warming in the SSTA is known as El Niño, while an anomalous cooling bears the name of La Niña. ENSO is well recognized as the dominant mode of interannual variability in the tropical Pacific Ocean. It affects the atmospheric general circulation which transmits the ENSO signal to the other parts of world; these remote effects are called "teleconnections" and induce changes in the occurrence of severe weather events, which dramatically affect human activities and ecosystems worldwide (see e.g.[11, 12, 13, 14]). Therefore, the ENSO index is of primary importance for climate scientists and appears as an interesting choice for testing the EWMD.

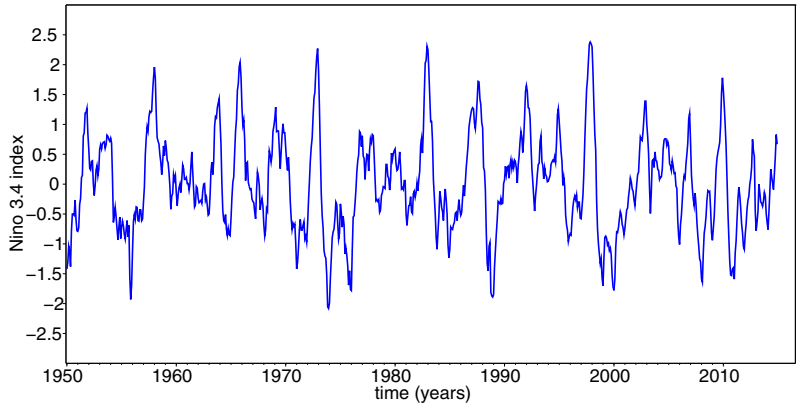


Fig. 5 The ENSO index, i.e. monthly sea surface temperature anomalies in the Equatorial Pacific Ocean.

As it can be seen in Fig. 6, the components extracted from the EWMD and the EMD are somehow comparable. The periods detected are respectively (from top to bottom in Fig. 6) 44.8, 28.6, 17, 65.6, 140.6 months for the EWMD and 9.8, 21, 38.6, 75.9, 138.4 for the EMD. It turns out that the periods detected with the EWMD seem more in agreement with previous studies than those obtained from the EMD (see e.g. [15, 16]). Let us note that both methods recover the famous $\approx$ 11-years period of the solar cycle. Regarding the reconstruction skills of the methods, the former has a RMSE of 0.277 and a PCC of 0.941, while the latter has a RMSE of 0.193 and a PCC of 0.973. Although these numbers seem to be better with the EMD, the 9.8-months component is likely a side effect due to the noise within the signal. Indeed, the EMD is known for having trouble to deal with noise, exhibiting IMFs very valuable for the reconstruction but with a rather poor physical interpretation. If that noisy component is not taken into account, the RMSE rises to 0.355 and the correlation drops to 0.903. Let us add that our first tests (not shown) indicate that the EWMD is not as sensitive to noise as the EMD, thus giving more reliable results especially for real-life data analysis. This has to be more carefully studied.

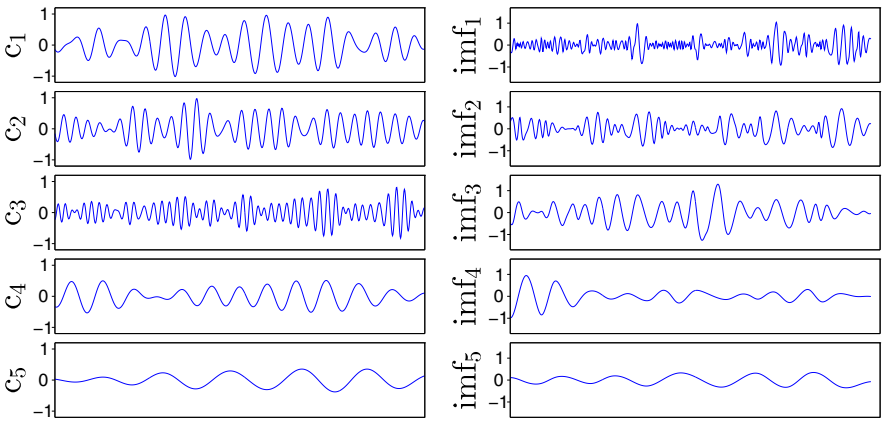


Fig. 6 Left: the components extracted from the ENSO signal with the EWMD. Right: the IMFs given by the EMD. The periods detected are 44.8, 28.6, 17, 65.6, 140.6 months (EWMD) and 9.8, 21, 38.6, 75.9, 138.4 (EMD). Those obtained with the EWMD seem to be more in agreement with some previous works. The reconstructions (not shown) are both satisfying. Let us note that the EMD displays a noisy component that artificially improves the reconstruction skills.

4 Conclusion

We presented a new wavelet-based method (EWMD) for decomposing an oscillating signal into several quasi-periodic components and compared the results with the famous empirical mode decomposition (EMD). It turns out that the decomposition-reconstruction skills of the EWMD are globally as good as those of the EMD. Moreover, the period detection abilities of the EWMD seem to outperform those of the

EMD. Regarding real-life data, we analyzed the ENSO index, i.e. temperatures of the Pacific Ocean. It appears that the periods detected by the EWMD seem more in agreement with previous studies, while both methods excel in decomposing-reconstructing the original signal. In general settings, the EWMD extracts components following somehow a decreasing level of energy while the EMD successively gives components with increasing mean frequency. Therefore, the EMD is highly sensitive to noise while the EWMD is barely affected; this should be investigated in detail in future works.

Let us note that the present paper is merely of an experimental nature and the skills of the promising EWMD have to be explored in more details. Also, we are well-aware of the fact that the EMD has been considerably improved since its first version (which is the one chosen here for comparing results). Therefore, it will be necessary to compare the EWMD with e.g. the EEMD and CEEMDAN as well as with other mode decomposition methods (such as e.g. [5] and methods mentioned therein). Let us finally note that this is the first version of the EWMD; improvements will be made in the near future, but the examples presented here show that the EWMD appears as a challenging candidate in the area of mode decomposition methods; we are optimistic about the possibilities it offers.

References

1. Nicolay, S.: A wavelet-based mode decomposition. European Physical Journal B **80**, 223–232 (2011)
2. Arneodo, A., Audit, B., Decoster, N., Muzy, J.F., Vaillant, C.: Wavelet based multifractal formalism: applications to DNA sequences, satellite images of the cloud structure, and stock market data. In: The Science of Disasters: Climate Disruptions. Heart Attacks, and Market Crashes, pp. 27–102. Springer, Berlin (2002)
3. Daubechies, I.: Ten lectures on Wavelets. SIAM (1992)
4. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press (1999)
5. Gilles, J.: Empirical wavelet transform. IEEE Transactions on Signal Processing **61**(16), 3999–4010 (2013)
6. Flandrin, P., Rilling, G., Goncalves, P.: Empirical mode decomposition as a filter bank. IEEE Signal Processing Letters **11**(2), 112–114 (2004)
7. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A **454**, 903–995 (1998)
8. Rilling, G., Flandrin, P., Goncalves, P.: On empirical mode decomposition and its algorithms. In: IEEE-EURASIP Workshop Nonlinear Signal Image Processing (NSIP) (2003)
9. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Advances in Adaptative Data Analysis **1**, 1–41 (2009)
10. Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P.: A complete ensemble empirical mode decomposition with adaptive noise. In: IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP) (2011)
11. Ashok, K., Behera, S., Rao, S., Weng, H., Yamagata, T.: El niño modoki and its possible teleconnection. Journal of Geophysical Research **112**(10.1029) (2007)
12. Glantz, M.: Currents of Change: Impacts of El Niño and La Niña on climate society. Cambridge University Press (2001)

13. Hsiang, S., Meng, K., Cane, M.: Civil conflicts are associated with the global climate. Nature **476**(7361), 438–441 (2011)
14. Yeh, S., Kug, J., Dewitte, B., Kwon, M., Kirtman, B., Jin, F.: El niño in a changing climate. Nature **461**(7263), 511–514 (2009)
15. Moron, V., Vautard, R., Ghil, M.: Trends, interdecadal and interannual oscillations in global sea-surface temperatures. Climate Dynamics **14**(7), 545–569 (1998)
16. Nicolay, S., Mabille, G., Fettweis, X., Erpicum, M.: 30 and 43 months period cycles found in air temperature time series using the morlet wavelet method. Climate dynamics **33**(7), 1117–1129 (2009)

Part X
Computer Vision, HCI and Image Processing/Analysis

Short Boundary Detection Using Spatial-Temporal Features

Muhammad Ali and Awais Adnan

Abstract Shot boundary detection is a major and challenging job in any video processing solutions. Different features are used for detecting shot boundaries in a video. Choice of features depends on the nature of processing. Mostly features are extracted from either spatial domain (visual features) or from temporal domain (motion information). To obtain more accuracy normally high level features are used but these features need complex computations. On the other hand low-level features like colour and brightness are easy to extract and compute but these features have less tolerance towards small changes in the frame. Similarly features in temporal domain have their limitations — they mostly ignore the spatial domain and hence give false results. In this proposed solution, features from both spatial and temporal domains are combined. In spatial domain brightness information is used that is easy to extract and fast to compute. From spatial domain motion information is used. Motion information is extracted from sub-blocks that add spatial information as well. Finally both the results are combined resulting in higher accuracy, more sensitivity and tolerance.

Keywords Segmentation · Image processing · Video processing · Motion vector

1 Introduction

The current era of computing is considered to be the era of imaging. Images and videos are gaining popularity and are used in all areas of life. One reason for the rapid popularity is the reduction in the prices of capturing devices and low price-capacity ratio of storage devices. Multi-terabyte videos are being captured and shared online. It is estimated that [1] at the end of 2015 more than three billion

M. Ali(✉) · A. Adnan
Institute of Management Sciences Peshawar, 1-A, E-5, Phase VII,
Hayatabad Peshawar 25100, Pakistan
e-mail: {muhammad.ali,awais.adnan}@imsciences.edu.pk

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_84

users will be using Internet, especially videos, which currently make up almost 50% of today's Internet traffic, will reach up to 69% of all global consumer Internet traffic in 2017 [2]. According to Forrester's research [3] "A minute of video is worth 1.8 million words". YouTube, one of the main video hosting services, is watched six billion hours per month which means around one hour for every person in this world. Managing, searching, indexing and filtering this huge visual data is a real challenge for the future. The dramatic growth in the use of videos requires efficient and effective video processing techniques.

In any video processing activity, the first and probably the most important step is detecting a shot-boundary. This process can be considered as a sequence of three incremental functions: First is the content representation in which visual contents are converted and represented in a way that forms a continuous function. Next is the evaluation of this continuity function to detect discontinuity. The third step is the classification where using discontinuity as change is used to identify shot boundary [4].

The cut is the simplest type of shot boundary that is easy to detect and compute. It usually appears when frames from two different shots are joined and hence can be detected as a sudden change in the two adjacent frames. In gradual transition shot translation is spread over multi frames. Different editing effects like fading, wiping and dissolving are example of this type. These types of shot-boundaries are hard to detect.

Sensitivity and invariance are two important parameters that contribute to the detection of shot boundaries and feature selection. The sensitivity shows how sensitive a feature is, which depends on the detail extracted from a video. Invariances on the other hand demonstrate the ability to tolerate the change in the feature space. At the extreme all the pixels present in an image can be taken as a feature. In this case sensitivity is high but at the same time it is less tolerant to changes. Simple geometric transformations such as rotation, scaling, or sheering etc. can easily effect the decision. For these types of features with more details more computations are needed. To reduce the computational complexity a set of features are extracted from the video. Such features are less sensitive but if selected intelligently show high tolerance towards changes and hence show high invariance.

Different types of features are used by researchers and can be grouped into four major categories: features based on text, features based on visual contents, audio based features, and last the combination of all three simultaneously. These features are usually extracted from one of two domains i.e. spatial domain or temporal domain. Visual features are extracted from the spatial domain whereas changes in these features along the temporal domain are used to build the community function. Mostly the spatial domain is used to extract a feature that forms a 1-D temporal signal where a discontinuity is identified as a shot boundary. Audio and other cinematic features are extracted in temporal domain where a change in these feature form the continuity signal.

2 Related Work

Due to its fundamental role in video processing, the shot-boundary detection remains a hot topic for researchers. A great deal of work has been done so far, especially after the launch of TRECVID (TREC Video Retrieval Evaluation) in 2001. This work can broadly be divided into four main categories: pixel based methods, histogram based methods, edge based methods, and temporal feature based methods.

Pixel based methods
Just like alphabets are units of text, pixels are the basic unit of image and hence the unit of digital video too. This is the reason why pixel based methods have been initially and are still one of the most important domains for feature extraction in image and video processing. Mostly in video processing, corresponding pixels from different frames are combined to construct a continuity function R which is analysed to detect discontinuity $\bar{R}$. A pixel based feature involves not only the intensity of the pixel but also its location. Such types of features are highly sensitive and less invariant. One of the early works that used pixel as a feature was written by Kikukawa and Kawafuchi [5]. In their work they used two adjacent frames to calculate the absolute difference of all pixels. The following equation (1) shows the mathematical representation of this method.

$$z(k) = \frac{1}{MN}\sum_{x=1}^{X}\sum_{y=1}^{Y}|I_k(x,y) - I_{k+l}(x,y)|_c \quad (1)$$

Where $I_k(x,y)$ is the intensity of the pixel at location (x, y) on the colour surface c and k and k and $k+l$ are two adjacent frames for all $l > 0$.

Otsuji et al. presented an improved concept in [6] by identifying only selected pixels for calculation. A threshold value for pixel intensity was used for pixel selection. Similar approaches with small modifications are used by different researchers such as Zhang et al. [7], and Nagasaka and Yeo [8, 9]. To reduce computation complexity and time, the authors of [10] presented the concept of visual rhythm in pixel based methods. M. Cooper [11] used the $\chi 2$ test to compare a histogram to adjacent frames to detect shot boundaries between them and yielded better accuracy then other histogram based methods [12].

2.1 Histogram Based Methods

Different statistical techniques are used to improve the sensitivity and complexity of simple pixel based methods --- one such method is histogram-based which is less sensitive than pixel based methods. However for some of the effects like filtering and zooming, the sensitivity of these methods is still high. Initially grey images were used, however the same methods can be applied on colour images

and videos by extending them in the colour dimension. The following equations (2 & 3) represent mathematically the approach used for grey and colour frames.

$$z(k) = \sum_{j=1}^{N_{Bins}} |H_k(j) - H_{k+1}(j)| \tag{2}$$

$$z(k) = \sum_{C} \sum_{j=1}^{N_{Bins}} \left| H_k^C(j) - H_{k+1}^C(j) \right| \tag{3}$$

Where $H_k(j)$ is histogram in intensity level j = 1. .255, for frame number k

$H_k^c(j)$ is the histogram for colour channel c whereas k is the number of frame.

M. Cooper [7, 11] and Gargi et al. [12] used the $\chi 2$ test for histogram comparison where the latter has shown that results from this technique are much better than simple histogram based techniques. It had been observed that in most of the histogram based techniques spatial information was ignored. This issue was resolved in [13] --- before the histogram calculation, they divided a frame into 16 blocks and then a separate histogram was calculated for each block and compared with the histogram of the corresponding block in the next frame.

2.2 Edge Based Methods

Histogram based methods are simple and relatively less sensitive than the pixel based methods, however they mostly work with global features extracted from the whole frame, irrespective of the nature and location of the object. Edges are, on the other hand, a type of feature that considers both the contents and the location of contents. In these techniques a mask is applied to extract the edges from the frame and then features extracted from these edges are compared to detect the shot boundary. Not only cut shots but other type of smooth transitions are detected using these methods [14, 15]. Edge-based contrast method, which is a modified version of simple edge based feature, is used by Rainer Lienhart in [16]. Priya and Domnic [17] used the strength of edges and an orthogonal vector for shot boundary detection that are applied on blocks of fames.

2.3 Temporal Techniques

Although visual features are a common choice for researchers in all types of image and video processing applications, but due to the close relationship between the shot, shot boundaries and the temporal feature, many researchers used the temporal domain in their research. Akutsu, et al. [18] used a motion vector for video indexing. With the motion vector they developed an approach to identify video segments by applying the smoothness on the vector that represents motion in the video. This approach not only identifies cut points but also different types of camera motions. Similarly authors of [19] used a support vector machine from the temporal domain to analyse videos in terms of boundary detection.

Osian and Van Gool, in their work presented in [20], used motion information by adaptive threshold for motion detection. The concept was improved and modified by Wang et al. in [21] where information about b-frames and p-frames were extracted directly from the MPEG compressed video. This modified approach however works with MPEG format only. Another limitation of this detection scheme is that it can identify only one type of shot i.e. the cut.

In [22], Whitehead et al. used feature tracking to find the dissimilarity in continuous function for shot boundary. According to the authors they used a complex technique because motion based shot detection methods are less reliable and effective than methods that are based on discontinuity in visual feature since reliable and accurate motion estimation is quite difficult.

Ferreira, Assuncao et al. [23] used different perceptual features to detect shot boundaries in 3D videos. They extracted the texture variation along the temporal dimension. K-mean clustering is used to classify boundaries and non-boundaries frames.

3 Proposed Solution

This proposed method can be broadly divided into two streams; one in the spatial domain and other in the temporal domain. The features are extracted separately from each of the two tracks and are then combined in the final step to extract shot boundaries in videos. In the first track the video is converted to RGB frames to extract visual features from the sorted histogram while in the second step the track motion vector is used to detect the shot boundary. These two tracks perform independently enabling the proposed system to exploit the concurrent processing (parallel processing). The proposed system is summarized in Figure 2.

3.1 Features Extraction in the Spatial Domain

In the first stream different visual features are extracted from two visual channels: Chroma channel that represents light and Luma channel that represents colour information of the frame. Novelty of this method is the reduction in the three mappings --- which are considered basic steps for any shot boundary identification process [4, 24, 25]. In this method feature extracted from the frame are used to construct the continuity function without any additional step of conversion of features into continuity function.

Chroma information is used from HSV colour model while Luma information frame is used in grey-level. Selection of these two channels separately makes this solution compatible with most of the component and composite video signals like S-video, YP_bP_r and RGB.

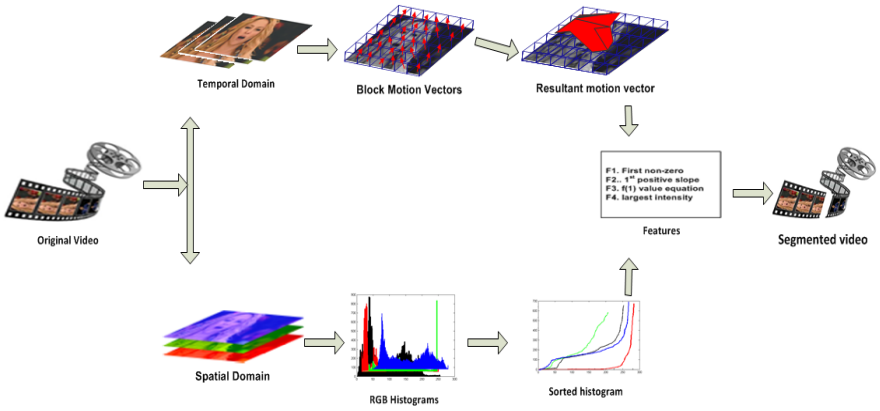


Fig. 1 Block Diagram of the Proposed Model

In the first step frame extracted from the video is converted into standard size of $240 \times 180 \times 3$ where 240 is the width of the frame and 180 is the height while 3 is the RGB colour channels. This size not only maintains 4×3 ratio but also contains 43200 pixels in each colour channel that is good enough to retain the important visual features and easy in computation using 16-bit integer.

In the next step, frame in RGB colour space is decomposed into Chroma and Luma channels. HSV colour model is used to extract feature in colour domain where grey-level representation is used to extract brightness information.

In the next step, after conversion of frame into four channels, histogram is calculated using 256 bits for each channel. In most of the systems 8-bits are used to represent intensity in any layer, so the suggested 256 bits include statistics for all the colour and brightness features. This histogram is then arranged in non-decreasing order such that Hn ≤ Hn+1 for all n = 1.. 256. Sorted histogram obtained in this step is used to construct polynomial curve of 3rd order to represent the histogram in simple mathematical form. Polynomial curve of order 3 is selected because it the nearest approximation of the sorted histogram. Least square method is used in this step to fit the curve on the sorted histogram. General form of polynomial curve of order m can be represented as following equation (4).

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + .. + a_mx^m \quad (4)$$

Where a0, a1, am are unknown coefficients

By solving above equations for m = 3, we obtained 3rd order polynomial curve as

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 \quad \text{where } x = 1, 2 \ldots 255 \quad (5)$$

Once the polynomial curve is obtained, the next step is to construct the feature vector F. It comprises the following four elements:

$$F = \{f_1, f_2, f_3, f_4\} \quad (6)$$

Details of the above elements are as

f_1 : This is the 1st non-zero value in the sorted histogram curve. $y = f(x)$fror $x = 1 \ldots 256$ then f1 = n such that $f(n) \neq 0$ and $f(n-1) = 0$.

f_2 : The second element in this feature vector is the coefficient of the polynomial curve which is obtained by using 1 in the variable as:

$$\begin{aligned} f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 \\ f(1) &= a_0 + a_1 + a_2 + a_3 \end{aligned} \tag{7}$$

We selected f (1) for this feature because the shapes of all the curves for other higher values of x are almost similar to x = 1 but the calculating value of the polynomial for x = 1 is simple, since it is the only addition to the four coordinates of the polynomial.

f_3: The third feature in this vector is the intensity level with the highest value in the histogram curve. If Hk,c is the value at k in histogram for channel c then f_3 for this channel is Hk,c if Hk,c ,cHm,c for all m = 1 .. 256.

f_4 : The last feature is the slope of the curve from f_1 to f_3. In our case $\Delta x = x_{n+1} - x_n = 1$ and $y_{n+1} > y_n$ for all n . the value of f_4 can be calculated as:

$$f_4 = \frac{dy_{nm}}{dx_{nm}} = \frac{dy_m - dy_n}{m-n} \tag{8}$$

To generate the continuity function we combine these four features according to the following equation:

$$F = [w_1 f_1 + w_2 f_2 + w_3 f_3]\, f_4^{w_4} \tag{9}$$

Where $w_1, w_2, w_3,$ and w_4 are the weights for the values that are calculated experimentally.

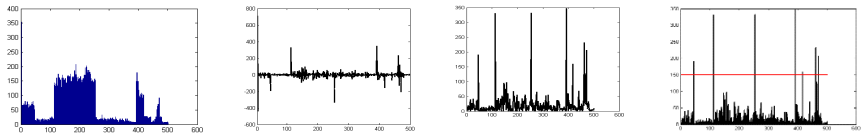
The shot boundary is identified where there is a sudden change in the continuation function F composed of four features. The boundary is identified at the point where the absolute difference in the continuation function is above threshold. Δf is the absolute difference as

$$\Delta f_n = |F_{n+1} - F_n| \tag{10}$$

and the point $\delta(f)$ where this difference is above the threshold is called the shot boundary.

$$\delta(f) = \begin{cases} \text{Yes}, & \Delta f \geq th \\ \text{No}, & \Delta f \geq th \end{cases} \tag{11}$$

Where th is the threshold value.



(a) Feature Point f b) Difference in f (c). Absolute difference (d) Thresholding

Fig. 2 Feature using thresholding on continuity function

3.2 Shot Detection Using a Motion Vector: Proposed Method

In the previous stream features were extracted from the visual domain while in this step the temporal domain is used. Once again, the video is extracted in RGB format that was resized and normalized into the 240 * 180 * 3 size. This normalized frame is then converted into 256 grey levels, 0 for the black (dark) and 255 for the white (bright) intensity level.

This two-dimensional image is then divided into 108 sub-blocks that are arranged into an array of 12 * 9. These 108 blocks when combined in the temporal domain construct different sub- channels.

Next is the identification of the motion vector within each channel (corresponding sub-blocks) for the neighbour frame with the distance l. Here the Three-Step Search (TSS) Algorithm [26] is used for the motion estimation. This process can be summarized as shown in Equation 12.

$$v(i,j)_n \leftarrow MV(B(i,j)_n, B(i,j)_{n+1}, s) \tag{12}$$

Where $v(i,j)_n$ is motion vector for block i, j in frames n and n+1, and s is the sensitivity level.

Then by combining all the motion in sub-blocks of a frame, a representative composite motion vector is formed by joining all the vectors.

$$V_n = f(X + Y) \tag{13}$$

where

$X = \sum_{i=1.j=1}^{12,9} x(i,j)$ and $Y = \sum_{i=1.j=1}^{12,9} y(i,j)$

Following figure show an example of shot boundaries identified fro, this method.

Fig. 3 Shot-boundaries detected

4 Results and Conclusion

To evaluate the performance of this proposed method, nine videos from the dataset obtained from the National Institute of Standards and Technology (NIST, http://www.nist.gov) are used. The total length of these segments is above

12817 seconds and contain 320437 frames. To overcome the computational complexity and evaluate the algorithm on normal machines available in the Institute, these segments are further divided into sub-segments such that no single segment has more than 10,000 frames --- thus a total of 35 video segments. The frame rates of all these segments are 25 whereas the frame size is 352 by 288. These are nature videos that contain all types of shot transition, however, in this work only cut shots are analysed.

Two common statistical tools, Precision and Recall are used to evaluate and compare the performance. Although accuracy rate is above 99% but this will not be a realistic indicator since out of 320437 frames only 1098 frames are from shot boundary and rest of the 319339 frames are non-boundary frames. Due to this unequal distribution of frames in the two classes (boundary and non-boundary), such a high accuracy rate does not reflect the performance of the algorithm. In such a situation precision and recall are more realistic indicators. All these results are summarized in figure 4 and figure 5

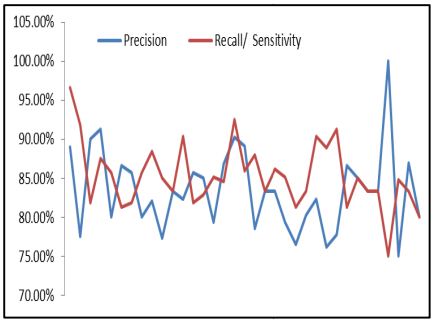


Fig. 4 Precision and recall

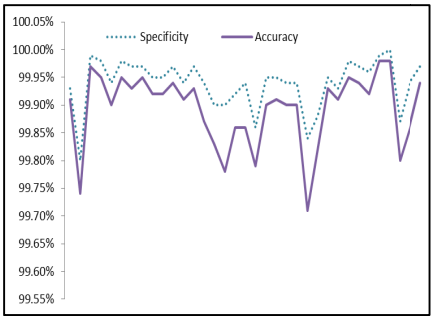


Fig. 5 Specificity and accuracy

Above figures shows that this system has high sensitivity (85.32%) as well as high invariance, whereas the overall accuracy is around 99.89%. All these results are tested on actual videos that contain all types of shot boundaries. Figure 4 also shows that on average more than 83% precision is obtained and even in some long videos it is about 100%. Recall is about 85%. In a related work presented in [27] where they used only temporal features, the value of recall was 75%. Similarly this proposed system show better precision (83.4) than the method [28] that used visual features only.

References

1. Kende, M.: Global Internet Report 2014, Geneva, Switzerland (2014)
2. Cisco: Cisco Visual Networking Index: Forecast and Methodology, 2013–2018, June 10, 2014
3. McQuivey, J.L., Lussanet, M.d., Wilkos, D.: How Video Will Take Over The World, June 17, 2008

4. Yuan, I., Wang, H., Xiao, L., Zheng, W., Li, J., Lin, F., et al.: A Formal Study of Shot Boundary Detection. IEEE Transactions on Circuits And Systems for Video Technology **17**, 168–186 (2007)
5. Kasturi, R., Jain, R.C.: Dynamic vision. In: Jain, R.K.R.C. (ed.) Computer Vision: Principles. IEEE Computer Society Press, Washington (1991)
6. Sethi, I., Patel, N.: A statistical approach to scene change detection. In: Storage and Retrieval for Image and Video Databases III, San Jose, CA, USA, pp. 329–338 (1995)
7. Zhang, H., Kankanhalli, A., Smoliar, S.W.: Automatic partitioning of full-motion video. Multimedia Systems **1**, 10–28 (1993)
8. Yeo, B.L., Liu, B.: Rapid scene analysis on compressed video. IEEE Transactions on Circuits and Systems for Video Technology **5**, 533–544 (1995)
9. Nagasaka, A., Tanaka, Y.: Automatic video indexing and full-video search for object appearances. In: Second Working Conference on Visual Database Systems II IFIP TC2/WG 2.6, Amsterdam, The Netherlands, pp. 113–127 (1992)
10. Chung, M.G., Kim, H., Song, S.M.-H.: A scene boundary detection method. Presented at the Proceedings of the 2000 International Conference on Image Processing (2000)
11. Cooper, M.: Video segmentation combining similarity analysis and classification. In: Proceedings of the 12th Annual ACM International Conference on Multimedia, MULTIMEDIA 2004, New York, pp. 252–255 (2004)
12. Choubeya, S.K., Raghavanb, V.V.: Generic and fully automatic content-based image retrieval using color. Pattern Recognition Letters **18**, 1233–1240 (1997)
13. Zhang, H., Low, C.Y., Smoliar, S.W.: Video parsing and browsing using compressed data. Multimedia Tools Appl. **1**, 89–111 (1995)
14. Ahmed, M., Karmouch, A., Abu-Hakimaˆ, S.: Key frame extraction and indexing for multimedia databases. Proc. Vis. Interface Conf. **1999**, 506–511 (1999)
15. Zabih, R., Miller, J., Mai, K.: A feature-based algorithm for detecting and classifying scene breaks. In: Proceedings of ACM Multimedia 1995, San Francisco (1995)
16. Lienhart, R.: Comparison of automatic shot boundary detection algorithms. In: Storage and Retrieval for Image and Video Databases, No. SPIE 3656, pp. 290–301 (1999)
17. Priya, G.G.L., Domnic, S.: Edge Strength Extraction using Orthogonal Vectors for Shot Boundary Detection. Presented at the Procedia Technology 2nd International Conference on Communication, Computing & Security [ICCCS-2012] (2012)
18. Akutsu, A., Tonomura, Y., Hashimoto, H., Ohba,Y.: Video indexing using motion vectors. In: Proceedings of SPIE Visual Communications and Image Processing (1992)
19. Yuan, J., Li, J., Lin, F., Zhang, B.: A unified shot boundary detection framework based on graph partition model. In: Proceedings of the 13th Annual ACM International Conference on Multimedia, MULTIMEDIA 2005, Singapore, pp. 539–542 (2005)
20. Osian, M., Gool, L.V.: Video Shot Characterization. Machine Vision and Applications **15**, 172–177 (2004)
21. Wang, X., Wang, S., Chen, H., Gabbouj, M.: A Shot Clustering Based Algorithm for Scene Segmentation. Presented at the International Conference on Computational Intelligence and Security Workshops (2007)
22. Whitehead, A., Bose, P., Laganiere, R.: Feature Based Cut Detection with Automatic Threshold Selection, *Lecture Notes in Computer Science, LNCS,* vol. 3115, pp. 2064–2065 (2004)
23. Ferreira, L., Assuncao, P., Cruz, L.A.d.S.: 3D video shot boundary detection based on clustering of depth-temporal features. In: 2013 11th International Workshop on Veszprem Presented at the Content-Based Multimedia Indexing (CBMI) (2013)

24. Hanjalic, A.: Shot boundary detection: unraveled and resolved? IEEE Transactions on Circuits and Systems for Video Technology **12**, 90–105 (2002)
25. Hanjalic, A.: Content-Based Analysis of Digital Video, 1st edn. Springer, Boston (2004)
26. Lu, J., Liou, M.L.: A Simple and Efficient Search Algorithm for Block-Matching Motion Estimation. IEEE Transactions on Circuits and Systems for Video Technology **7**, 429–433 (1997)
27. Adnan, A., Ali, M., Nawaz, M.: Shot Boundary Detection Using Motion Information in Macro Blocks. In: AWER Procedia Information Technology & Computer Science, pp. 1–9 (2013)
28. Adnan, A., Ali, M.: Shot Boundary Detection Using Sorted Color Histogram Polynomial Curve. Life Science Journal **10** (2013)

A Set of Usability Heuristics and Design Recommendations for u-Learning Applications

Fabiola Sanz, Raúl Galvez, Cristian Rusu, Silvana Roncagliolo, Virginica Rusu, César A. Collazos, Juan Pablo Cofré, Aníbal Campos and Daniela Quiñones

Abstract Usability is one of the most relevant attribute when evaluating a software product, application or a website, as it is important to analyze how the interaction design facilitates or hinders the user to achieve an objective in concrete. Moreover, the user experience expresses its positive or negative perception of a particular application through a set of factors and elements relating to user interaction. The paper presents a set of heuristics to detect usability problems in u-Learning applications. It also proposes a set of design recommendations focused on the user experience.

Keywords Usability · User experience · u-Learning · Usability heuristics · Design recommendations

1 Introduction

Because of the emergence of new technologies and the need for humans to acquire knowledge in other ways, it arises in the first instance the Electronic learning

F. Sanz · R. Galvez · C. Rusu · S. Roncagliolo · J.P. Cofré · A. Campos · D. Quiñones(✉)
Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: faby.sanz.b@gmail.com, raul.galvez.cabeza@gmail.com, jp.cofreuribe@gmail.com, anibal.campos.baez@gmail.com, danielacqo@gmail.com, {cristian.rusu,silvana}@ucv.cl

V. Rusu
Universidad de Playa Ancha de Ciencias de La Educación, Valparaíso, Chile
e-mail: virginica.rusu@upla.cl

C.A. Collazos
Universidad Del Cauca, Popayán, Colombia
e-mail: ccollazo@unicauca.edu.co

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_85

(e-Learning), based on three fundamental criteria: “the e-Learning is networked, is delivered to the end user through the use of computers using standard Internet technology and focuses on the broadest view of learning that go beyond the traditional para- digms of training” [1]. Due to the expansion of the concept e-Learning new learning, encompassing activities such as: Classroom Learning (c-Learning), Electronic Train- ing (e-Training), Mobile-Learning (m-Learning).

A new concept emerges, known as u-Learning. It is defined as "a model of learning that takes place in an environment of ubiquitous computing, which allows you to learn the right thing at the place and right time, the right way" [2]. It allows to "… move learning beyond the classroom to different environments of everyday life supported by a flexible, invisible and omnipresent technology that provides us the information we need at all times" [3].

Due to the specific characteristics of u-Learning applications and the lack of specific heuristics, it is difficult and expensive to assess their usability. We did not found user experience - oriented design recommendations for u-Learning applications.

The paper presents a set of usability heuristics and a set of user experience - oriented design recommendations for u-Learning applications. Section 2 presents concepts related to usability and user experience. Section 3 describes the proposed heuristics, including the methodology that we used. Section 4 presents specific design recommendations. Section 5 highlights conclusions and future work.

2 Usability and User Experience

According to the ISO 9241-11 standard the usability is "the degree to which a product can be used by specific users to achieve certain specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [4]. Nielsen defines usability in terms of: Learnability, Efficiency, Memorability, Errors, Satisfaction [5].

The ISO 9241-210 standard defines the user experience as a result of the percep- tions and responses of a person by use and anticipated use of a product, system or service [6]. Dillon proposes a model that defines the user experience as the sum of three levels: action that the user does; result, the user gets; and emotion that the user feels [7].

The user experience seeks to change user loyalty through an interaction that evokes positive emotions before, during and after using a software product. Usability is limited to the user’s goal achievement as easily and effectively as possible. User experience includes usability; it is a broader concept that includes the design, ergonomics, accessibility, emotional elements and usability, among other factors.

Helander et al. consider the usability evaluation as a process that measures the level of usability of a software product [8]. Assessing usability includes an application that is being evaluated, and a process by which one or more usability’s attributes are judged. Fernandez et al. affirm that a method of

evaluating usability is a procedure that is composed by a series of activities well defined for collecting data related to the interaction of an end user with a software product, and/or how the specific properties of this product software contribute to reach some degree of usability [9].

Usability inspections are based on the ability of evaluators to examine whether a particular interface fulfills a series of usability principles. These methods depend on the opinions, judgments and reports generated by evaluators [10]. Usability tests are methods where a user or a group of users is/are asked to use a prototype or a system. They provide direct information on how people use computers and their specific problems on an interface that is being tested [11].

The heuristic evaluation is a widely used usability inspection method. Jakob Nielsen proposes a set of 10 heuristics (principles) that allow identifying potential usability problems. However, general heuristics may neglect specific domain – related usability issues. Many authors propose set of specific usability heuristics.

Morville proposes the User Experience Honeycomb [12], which indicates that user experience is significant and valuable when information is:

- Useful: The content should be original and satisfy a need.
- Usable: The site must be easy to use.
- Desirable: Image, identity, brand and other design elements are used to evoke emotion and gratitude.
- Findable: The content has to be localizable and navigable internal and externally.
- Accessible: The content has to be accessible to people with disabilities.
- Credible: Users must trust and believe what they are told.

3 Developing Usability Heuristics for u-Learning

A methodology that facilitates the development process and/or particularization of usability heuristics was used [13]. It includes 6 stages:

- STAGE1: An exploratory phase, to collect bibliography related to u-Learning applications, their characteristics, general and/or related (if there are some) usability heuristics.
- STAGE2: A descriptive phase, to highlight the most important characteristics of the previously collected information, in order to formalize the main concepts associated with the research.
- STAGE3: A correlational phase, to identify the characteristics that the usability heuristics for u-Learning applications should have, based on traditional heuristics and case studies analysis.
- STAGE4: An explicative phase, to formally specify the set of the proposed heuristics, using a standard template.

- STAGE5: A validation phase (experimental) to check the new heuristics against traditional heuristic principles through experiments.
- STAGE6: A refinement phase, based on the feedback from the validation stage.

The set of usability heuristics for u-Learning was specified following the template proposed by Rusu et al. [13]:

- ID, Name and Definition: Heuristic´s identifier, name and definition.
- Description: The general concept that encompasses heuristics.
- Explanation: Explanation extended content and concepts covering heuristics, including tips and comparisons with other heuristics.
- Examples: cases are shown in which the heuristic has not been met, and in some cases positive examples of the application of heuristic principle.
- Benefits: Potential benefits of compliance with the principle.
- Anticipated problems: Potential difficulties foreseen in the implementation of heuristics.

3.1 A Set of Usability Heuristics for u-Learning

A set of 16 usability heuristics was built based on the distinctive features of U- Learning, the particularization of Nielsen's usability heuristics [14], and a set of preliminary heuristics proposed by Cofré [15]. The set of heuristics is presented below, in an abbreviated form (ID, Name and Definition).

- (UL01) Learning measurement: The application evaluates the user's learning progress, indicating the results.
- (UL02) Situated learning: The application solves doubts users in time when they arise
- (UL03) Collaborative learning: Users work together to maximize their own learning and that of others.
- (UL04) Continuity of learning resources: The application records all learning processes.
- (UL05) Connections and resources: The user has access to educational resources from their devices.
- (UL06) Synchronous and asynchronous interaction: The application provides synchronous and asynchronous communication.
- (UL07) Visibility state of the application: The application keeps users informed through feedback.
- (UL08) Coincidence between the application and the real world: The application speaks the language of the user.
- (UL09) User control and freedom: The application provides emergency exits, undo and redo.
- (UL10) Consistency and standards: The application follows the existing conventions, dependently on the device used.

- (UL11) Error prevention: The application warns of critical actions and irreversible.
- (UL12) Flexibility and efficiency of use: The application provides the ability to accommodate different work styles.
- (UL13) Minimize the user´s memory load: The user should not remember information from one part of the dialogue to another.
- (UL14) Aesthetic and minimalist design: The information displayed on the screen is relevant and is not overloaded with useless information or little used.
- (UL15) Help users recognize, diagnose and recover from errors: The application displays error messages simple, suggesting constructive solutions.
- (UL16) Help and documentation: The application provides the user documentation for the application.

Below is presented the full template of the heuristic "(UL01) Learning Measurement".

ID: UL01.

Name: Learning measurement.

Description: The application evaluates the user's learning progress, indicating the results.

Explanation: The application performs an initial evaluation, also has tools to measure the degree of achievement of learning, indicating its progress, approval or disapproval.

Examples:

- Fig. 1 shows the Level Test option, which allows users to select their English knowledge level or to make a test in order to identify it.
- Fig. 2 shows the evaluation test.

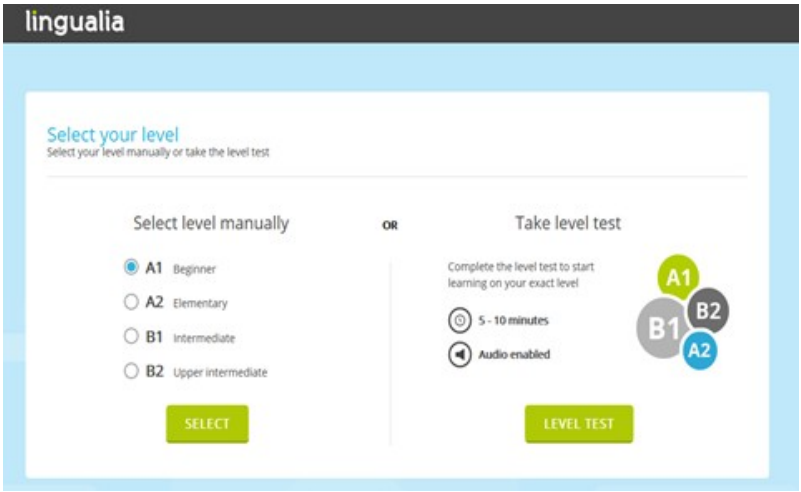


Fig. 1 Level Test

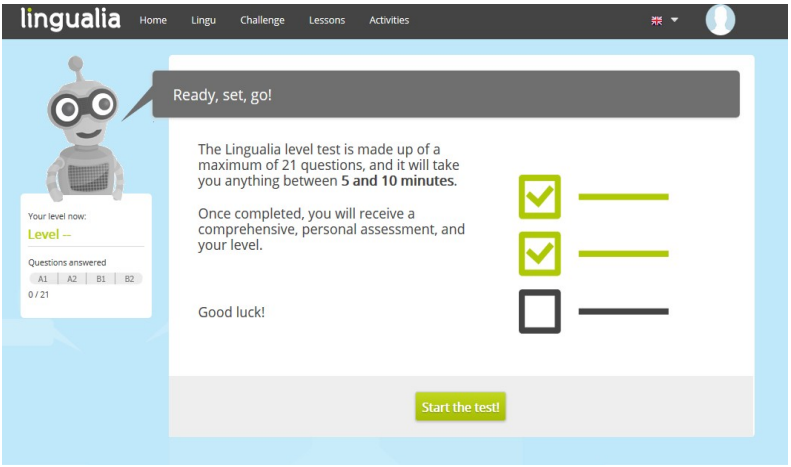


Fig. 2 Evaluation Test

Benefits:

- An initial evaluation of user knowledge allows particularizing the content according to his/her knowledge level.
- Provides quantitative data on the learning outcomes.

Anticipated Problem: The measurement of learning should be based on the results obtained after making any kind of evaluation, not indicate the number of daily, weekly and / or monthly hours performed.

3.2 Validating the Set of Heuristics

The proposed set of heuristics was refined in 2 iterations, through usability inspections, involving 4 expert evaluators. A semi-structured survey was also conducted, involving all evaluators. The experimental results allowed us to refine the set of heuristics.

Duolingo and *Lingualia* (desktop and mobile versions) were evaluated. Both applications meet u-Learning basic features: permanence, accessibility, immediacy, interactivity, adaptability.

A total of 34 problems were identified when evaluating *Duolingo*. 14 problems were associated with heuristics related to specific u-Learning characteristics (UL01 to UL06); the remaining problems were associated to general (Nielsen's) usability heuristics adapted to ubiquitous learning (UL07 to UL16). Fig. 3 shows the distribution of the problems encountered.

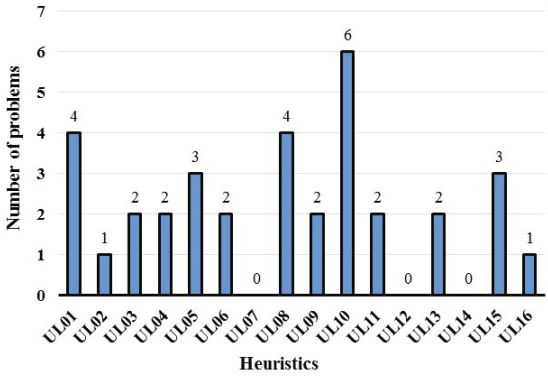


Fig. 3 Usability problems identified in *Duolingo*

Problems' criticality ranges from 3.25 to 7.0, in a 0 (minor) to 8 (sever) scale. Problems' severity varies between 1.50 and 3.37, in a 0 (minor) to 4 (sever) scale. Problems with higher criticality and severity were associated to heuristics:

- Measuring learning (UL01),
- Connection and resources (UL05),
- Synchronous and asynchronous interaction (UL06).

Problems with lower criticality were associated to heuristics:

- Consistency and standards (UL10),
- Minimize the burden of memory (UL13).

Problems with lower severity refer to:

- Coincidence between the application and the real world (UL08),
- Consistency and standards (UL10).

The heuristic evaluation of *Lingualia* allowed finding a total of 39 problems, of which 10 were associated with heuristics related to specific u-Learning characteristics (of UL01 to UL06); the remaining problems were associated general (Nielsen's) usability heuristics adapted to ubiquitous learning (UL07 to UL16). Fig. 4 shows the distribution of the problems encountered.

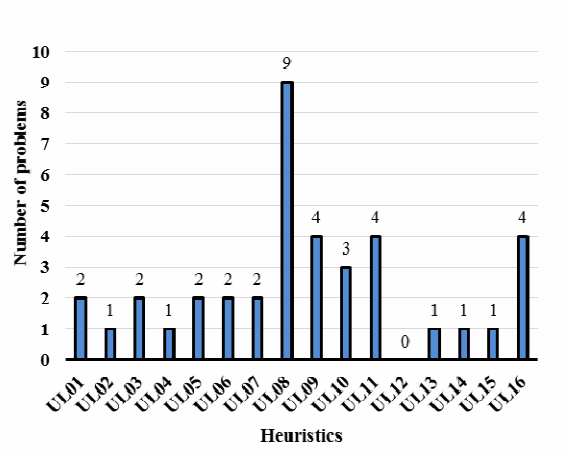


Fig. 4 Number of problems encountered in *Lingualia*

Problems' criticality ranges from 3.8 to 7.0. Problems with the highest criticality were associated to heuristics:

- Connection and resources (UL05),
- User control and freedom (UL09).

Problems with lower criticality were associated to heuristics:

- Coincidence between the application and the real world (UL08),
- Error prevention (UL11).

Problems' severity varies between 2.0 and 4.0. The most severe problems are related to:

- Connection and resources (UL05),
- Aesthetic and minimalist design (UL14).

Problems with lower severity refer to:

- Aesthetic and minimalist design (UL14),
- Error prevention (UL11),
- Coincidence between the application and the real world (UL08).

The experimental results prove that the set of u-Learning usability heuristics is a useful instrument. It allowed discovering additional problems, comparing to Nielsen's (particularized) heuristics. These problems are associated with the characteristics of u- Learning, directly related to the learning process, and they

have high levels of severity and criticality. Heuristic UL12 has no associated problems; both case studies adapt to different styles of work, so do not violate this heuristic.

4 Design Recommendations for u-Learning

A set of 10 design recommendations for u-Learning applications was developed, based and u-Learning characteristics, Peter Morville's user experience model, and usability issues identified in *Duolingo* and *Lingualia*. The set of recommendations is presented below.

(DR01) Learning Resources: The application should provide users educational activities through the course(s) that contain(s) units, lessons and support material for learning, considering an initial level test in order to suit your needs.

(DR02) Supporting the learning process: The application should provide users the opportunity to answer questions and conduct collaborative learning.

(DR03) Continuity and access to educational resources: The application must register all learning processes independently to the device used, since the user can access different devices, even without an internet connection.

(DR04) Interactivity: The application must allow communication with other users and / or experts, in real time, even if there is no coincidence in time.

(DR05) Ease of use and consistency: The ease and aesthetic consistency in the application are critical to providing independent of the device used satisfying experience.

(DR06) Error prevention and emergency exits: The application must warn of critical actions and irreversible. If the user reaches an undesired state, the application must provide emergency exits, undo and redo.

(DR07) User Help: When errors occur in the application must be given simple messages mistake, suggesting constructive solutions. Furthermore the application documentation is provided.

(DR08) Desirable application: images, icons, brand identity and other design elements are used to evoke emotion and gratitude. The application should keep users informed through feedback

(DR09) Findable resources: The content has to be localizable and navigable internal and externally, adapting to individual work styles that have users.

(DR10) Learning resources, educational and evaluative, should come from competent authorities, as users must trust and believe what they are told.

5 Conclusions

Because u-Learning has distinctive features, focusing on the process of learning that can take place anywhere, anytime, using various devices, it is vital to consider the characteristics of permanence, accessibility, immediacy, interactivity, activities education and adaptability located, when evaluating usability. Generic usability

heuristics may ignore domain – related usability issues. Specific u-Learning heuristics proved to work better.

Our research was based on a preliminary set of usability heuristics for u-Learning applications, proposed by Cofré and others [15]. The set of heuristics was experimentally validated through usability inspections, in two case studies: *Duolingo* and *Lingualia*. Experimental results and experts' opinion allow improving heuristics' definition.

Experiments proved the utility of the set of usability heuristics that we developed. It allows discovering problems associated with the characteristics of u-Learning not covered by the Nielsen's generic heuristics.

User experience aims to guide, control and improve the positive and negative feelings experienced by users. Based on the usability inspections' results, the u-Learning characteristics, and Peter Morville's user experience model, we proposed a set of user experience – oriented design recommendations for u-Learning applications.

As future work the set of heuristic usability for u-Learning applications will be validated in other case studies. The design recommendations will be validated through implementations of functional prototypes.

Acknowledgment The authors would like to thank all the participants involved into the experiments that the present study required, especially the members of the "UseCV" Research Group. The study was highly supported by the School of Informatics Engineering (Escuela de Ingeniería Informática) of the Pontifical Catholic University of Valparaíso (Pontificia Universidad Católica de Valparaíso) – Chile. This work has been part of the project CYTED 513RT0481 - "Red Iberoamericana de apoyo a los procesos de Enseñanza- Aprendizaje de competencias profesionales a través de entornos ubicuos y colaborati- vos". Daniela Quiñones has been granted the "INF-PUCV" Graduate Scholarship.

References

1. Rosenberg, M.: E-learning strategies for delivering knowledge in the digital age. Mc Graw Hill, United States (2001)
2. Ahmad, E., Yahya, S., Jalil, K.A.: The definition and characteristics of ubiquitous learning: A discussion. International Journal of Education and Development using Information and Communication Technology **6**(1), 117–127 (2010)
3. Morfi, M.: Learning Review (2012). http://www.learningreview.com/servplataformas-de-e-learning/2433-u-learning-aprendizaje-donde-quiera-que-estes
4. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDT's) – in Part 11: Guidance on Usability. International Organization for Standardization, Geneva (1998)
5. Nielsen, J., Norman, D.: Nielsen Norman Group (1195). http://www.nngroup.com/articles/usability-101-introduction-to-usability/
6. Usability Partners (2001). http://www.usabilitypartners.se/about-usability/
7. Dillon, A.: The University of Texas at Austin (2001). https://www.ischool.utexas.edu/~adillon/Journals/BeyondUsability.pdf

8. Helander, M., Landauer, T., V. Prabhu, P.: Handbook of Human Computer Interaction. Elsevier Sciencie B.V. (1997)
9. Fernandez, A., Insfran, E., Abrahão, S.: Usability evaluation methods for the web: A sys- tematic mapping study. Information and Software Technology **53**(8), 789–817 (2011)
10. Sánchez, W.: La usabilidad en Ingeniería de Software: definición y características. Ing-novación 3 **1**(2), 7–21 (2011)
11. Nielsen, J.: Usability Engineering. Academic Press Inc., United States (1993)
12. Morville., P.: Usability.gov. http://www.usability.gov/what-and-why/user-experience.html
13. Rusu, C., Roncagliolo, S., Rusu, V., Collazos, C.: A Methodology to Establish Usability Heuristics. In: The Fourth International Conference on Advances in Computer-Human Interactions (ACHI 2011), IARIA, pp. 59–62 (2011)
14. Nielsen, J.: Ten Usability Heuristics (2005). http://www.useit.com/papers/heuristic/heuristic_list.html. (Accessed September 17, 2013)
15. Cofré, J.P.: Usabilidad en u-Learning. Master Degree Thesis, Pontificia Universidad Católica de Valparaíso (2013)

Computer Input Just by Sight and Its Applications in Particular for Disable Persons

Kohei Arai

Abstract Computer input just by sight for disable persons is proposed together with its applications. It is confirmed that communication aids, phoning aids, service robot control, information collection aids, etc. are available by using the proposed computer input just by sight.

Keywords Human computer interaction · Computer input by human eyes only

1 Introduction

The number of disabled persons is getting large. There are 3.7 million of disabled persons in Japan. 0.3 million of disabled persons out of 3.7 million of disabled persons are blind or having problem on their eye(s). Therefore, 3.4 million of disabled persons have difficulties on their body and can use their eyes. Also, the number of elderly persons reached to 31 million persons in Japan and is increased rapidly. Such persons have difficulties on communications, having meals, information collections, watching TV, listening to radio, phoning, moving to somewhere with their Electric Wheel Chair: EWC, sightseeing, etc. The proposed Eye Based Human Computer Interaction: EBHCI is intended to provide such capabilities for disabled persons as well as elderly persons by using their eyes only.

Computer input by human eyes only has been proposed and developed [1]-[38]. It allows users to key-in operation and mouse operation by their eyes only allowing users' head movements. Conventional computer input by human eyes only has the following problems, (1) Key-in success rate is poor, (2) Influence due to illumination conditions, (3) It requires a time consumable calibrations, (4) Influence due to head movements, etc. The computer input by human eyes only uses visible to camera

K. Arai(✉)
Graduate School of Science and Engineering, Saga University, Saga, Japan
e-mail: arai@is.saga-u.ac.jp

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_86

to acquire human face. Eye is extracted from the acquired human face image. Then extract cornea and sclera. Meantime, feature points (two ends of eyes, two ends of mouth, two ends of eyebrows, for instance) which allows estimation of head pose are extracted. Therefore, geometric relations among feature points extracted from human face images are important. In order to determine the geometric relations, calibration is required. During the calibration process, users have to look at the four corners and the center of the computer screen. It takes a couple of minutes to acquire human face images when they are looking at the five points on the screen. Sometime, the calibration has to be tried again when the geometric relations are changed a lot. Using the location of cornea center in the image and the geometric relations, line of sight vector (at where users are looking) is estimated. Sometime, the geometric relations cannot be estimated due to changes of illumination conditions. There are influences due to shadow, shade, and feature points missing due to occlusion, blocking cornea by eye muscular, and so on. Near Infrared: NIR cameras with NIR light source (Light Emission Diode: LED) is effective to overcome the influences. On the other hand, cornea curvatures are different by human individually. Therefore, the calibration is required to adjust the geometric relation variations due to cornea curvature differences. Using double Purkinje images, cornea curvature can be estimated individually. Thus time consumable calibration is no longer needed when cornea curvature is estimated. More importantly, key-in success rate of the conventional computer input by human eyes only has to be improved. One of the reasons for poor key-in success rate of the conventional computer input by human eyes only is instability of gaze estimations. Due to computational errors, the estimated location at which users are looking at the computer screen is varied during the time when users are key-in. Therefore, it is hard to determine the designated keys when the distance between adjacent keys is shorter than variance of the gaze estimation results in poor key-in success rate. If the standard keyboard of key layout is used for computer input by human eyes only, the distance between adjacent keys is around 1.5 cm on the 16 inches computer screen. It is not so easy to stabilize the estimated gaze location within 1.5 cm. Therefore, some breakthrough is highly required. The proposed computer input by human eyes only introduced moving keyboard. Key layout can be simplified because keyboard can be moved to arbitrary locations. Therefore, all the keys on the keyboard can be selected and determined by using enlarged five keys, left/right/top/bottom/center. The keyboard which is displayed in the back of the five keys moves to the left direction when user is looking at left key. That is the same thing for the other four keys. Thus the entire keys on the keyboard can be selected and determined by using the five enlarged keys (transparent) for control of the keyboard movement. Five of enlarged keys can be selected and determined easily because these are displayed on the full computer screen with enough size. Thus key-in success rate becomes 100 %. Every 0.4 s, moving keyboard can be moved. Therefore, it takes less than 10 s even for 24 keys has to be travelled from the left to the right ends of the moving keyboard (keyboard consists of 24 by 6 keys). Computer input by human eyes only can be applied to a variety of fields, Mouse operations, Electric Wheel Chair control, Having meal aid, Communication aid, Information collection aid, Service robot control [39]-[45], e-learning system, etc. It is possible to manipulate mouse by human eyes only. Only thing users have to do is just looking at pointer dis-

played on to computer screen. All the mouse operations, click, drug, drop, etc. can be done with human eyes only. Electric Wheel Chair: EWC can be controlled by human eyes only. EWC turn right when user looks at right direction while EWC goes forward when user looks at the center. Meanwhile, EWC turn left when user looks at the left direction. Robot arm equipped beside the patient bed can retrieve a food from the meal tray through controlling robot arm motion (the tip of the robot arm) by patient's eyes. Using the proposed moving keyboard, sentences can be created by human eyes only. Then the sentences can be read by using software tool for type to talk. Thus the user can talk with other persons by human eyes only. Also, user can make phoning, search information through web sites, watching TV, listening to radio, search e-book, e-learning contents, etc. by human eyes only. Lecture can monitor the locations at which student is looking on the computer screen through lessons with e-learning contents by human eyes only. Students' gaze locations can be monitored by lecturers when the student uses computer input by human eyes only [46]-[49]. Evaluation of students' concentration to the e-learning contents at which lecturers would like students look can be done with computer input by human eyes only. Service robot can be controlled by human eyes only. Using image acquired with camera mounted on the service robot, patient lay down on bed can control service robot movement. Also, patient can create sentences with computer input by human eyes only. Therefore, patient can enjoy conversation with the other peoples through service robot and also can ask something to the other peoples. The paper proposes the computer input system by human eyes only together with the aforementioned applications of the system.

The proposed system is described in the following section followed by some applications together with some experimental results. Then concluding remarks with some discussions is described in the final section.

2 Proposed System

2.1 System Configuration

Fig. 1 shows the entire system configuration of the proposed computer input system just by sight for disabled and elderly person. Table 1(a) and (b) shows major specifications of single Head Mount Display: HMD and Near Infrared: NIR camera used for the proposed system.

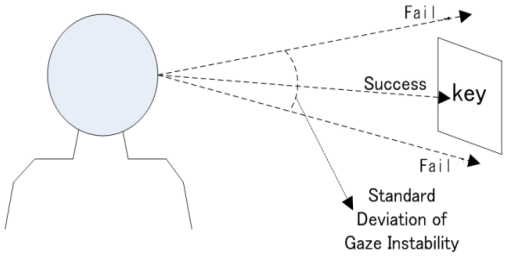


Fig. 1 System configuration of the proposed computer input system

Table 1 Major Specifications of HMD and NIR camera

(a) HMD Resolution	SVGA(800×600pixels)
Supposed distance	2m
Supposed size	60inch
Field of view	42degree
Input type	RGB
Operable temp。	0～40℃
Size	28mm(W)×35.24mm(H)×56mm(D)
Weight	20g
(b)NIR camera Resolution	1,300,000pixels
Minimum distance	20cm
Frame rate	30fps
Minimum illumination	30lux
Size	52mm(W)x70mm(H)x65mm(D)
Weight	105g

Five transparent enlarged keys are displayed onto HMD screen as shown in Fig. 2. The key displayed onto screen (in this case "G") is candidate for selection. If the user is looking at the right key, then key layout is moved right direction by 0.4s unit time. That is the same thing for the other four keys. If the centered key is desired key, then the candidate becomes selected key. And if the user looks at the selected key for 0.7s, the key is determined. Thus key-in can be done every 0.7s in minimum. This method is referred to "Moving keyboard" while traditional keyboard is called "Fixed keyboard" hereafter.

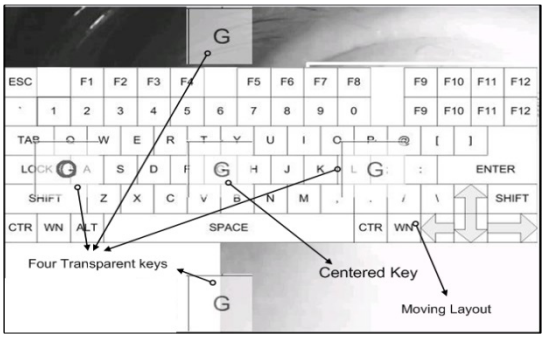


Fig. 2 Key layout of the proposed key-in system just by sight

2.2 Experiments

Key-in success rate and the required time for key-in is evaluated by three users. The results are shown in Table 2.

Table 2 Key-in Success Rate and the Time Required for Key-in

		Trial#1	Trial#2	Trial#3	Average	Time
User#1	Moving	100.00	100.00	100.00	100.00	101
	Fixed	78.57	92.86	85.71	85.71	197
User#2	Moving	100.00	100.00	100.00	100.00	159
	Fixed	92.86	71.43	64.29	76.19	219
User#3	Moving	100.00	100.00	100.00	100.00	109
	Fixed	71.43	64.29	64.29	66.67	227

As shown in Table 2, key-in success rate for the moving keyboard is 100% while that for the fixed keyboard is 76.19(%) in average. Meanwhile, the time required for key-in for the moving keyboard is 123 s while that for the fixed keyboard is 214 s. Therefore, key-in time of the moving keyboard is 74.25% much faster than that of the fixed keyboard.

3 Applications

3.1 Communication Aids

The proposed system can be used for communication aids. Namely, users may create sentences by using the proposed key-in system just by sight. Then created sentences are read by Type-to-Talk of software tools. Fig. 3 shows one of the examples of key layouts for Japanese key-in. Two layered Japanese "Hiragana" characters are layout in the key layout for this case. Therefore, users may select the desired key through two stages of selection of keys.

Fig. 3 Key layout of the proposed communication aids for Japanese language

3.2 *Information Collection Aids*

Main menu of the proposed information collection system just by sight is shown in Fig. 4. Users may select the desired functions in the menu just by sight. If users selected "TV", then the TV program is displayed onto the screen as shown in Fig. 5 (a). Channel selection, screenshot, information extraction from screenshot images and troubleshooting can be done with selection of functions from the submenu.

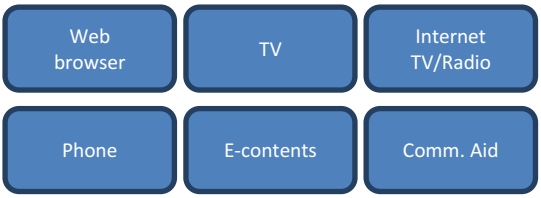


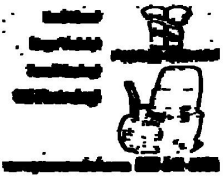
Fig. 4 Main menu for the information collection system

(a)Sub-Menu of TV monitoring

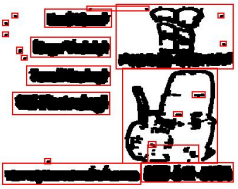
(b) Half-tone Image

(c) Inverted Image

(d) Binarization Image



(e) Morphologic Filtered Image



(f) Blob Detected Image

Fig. 5 Process flow of the information extraction from TV screenshot images

Screenshot images are converted to halftone image and then binary image through inverted image. After the isolated pixels are removed by using morphologic filter, blob images are extracted for keyword (such as sale providers' name, price, contact information, etc.) selections. This information required for purchase the product for sale are extracted from the screenshot images of TV program. Experiments are conducted with 50 of TV screenshot of images. Test results shows acceptable performance as shown in Table 3.

Table 3 Performance of information (required for purchase the desired sales products) extraction from TV screenshot of images

Number of images	50
Number of text lines	243
#correct detected	232
#false positives	40
Recall (%)	85.29%
Precision (%)	95.47%

Phoning is also available just by sight. Fig. 6 shows the menu for phoning. In the middle of row of the menu, there is the number from 0 to 9. These numbers can be moved. Also there are enlarged three transparent keys, LEFT, RIGHT and the selected key ("6" in this case). This is the same thing for moving keyboard. The candidate key is selected just by sight and determined keys are displayed between middle and bottom layers. Back-Space (BS) is also selected. If the telephone numbers users selected are correct, then users look at the top-center key of "CALL". Thus dialing can be done.

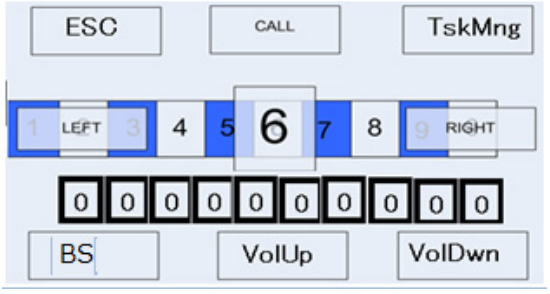


Fig. 6 Phoning system of the proposed information collection system

e-Book can be chosen and read by users using the proposed information collection system just by sight. Fig. 7 shows an example of screenshot images of e-Book reading system. Page forward and backward is available. Also content of the book can be referred and then the desired content item is selected and display onto the screen.

3.3 Service Robot Controlled Just by Sight

Another application of computer input just by sight is service robot control just by sight. Fig. 8 shows system configuration while Fig. 9 (a) and (b) shows outlook of the robot and components of the system. Major specification of the robot is shown in Table 4. The robot can be controlled just by sight. Users wear HMD and NIR camera. NIR camera is for acquisition of users' eye. Meanwhile, Menu shown in Fig. 10 (a) is displayed on HMD screen. There are two operating modes, Automatic

and Menu together with key-in mode. Usually, automatic mode is selected. In the middle, there are three arrows, left/right/up. If user looks at the “up arrow”, then the service robot is moved in the forward direction. Also if user looks at left or right arrows, then the robot turns left or right directions. At the tip of the robot, there is another camera which acquires forward images. Therefore, users can look at the image and control the robot just by sight. Other pictorial submenus are also available for users as shown in Fig. 10 (b).

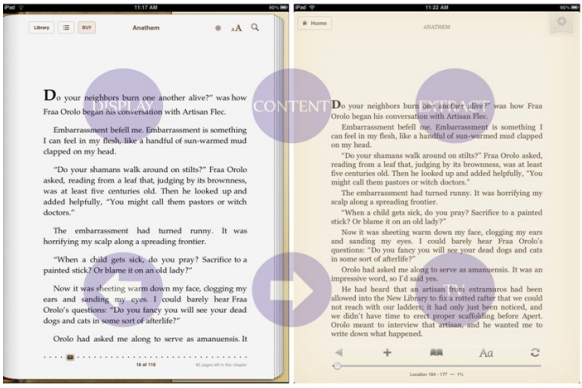


Fig. 7 e-Book reading system

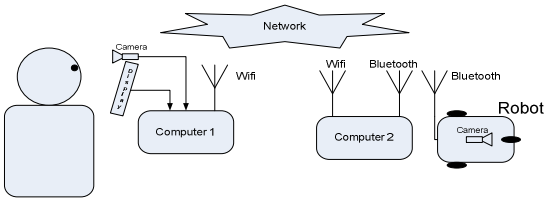
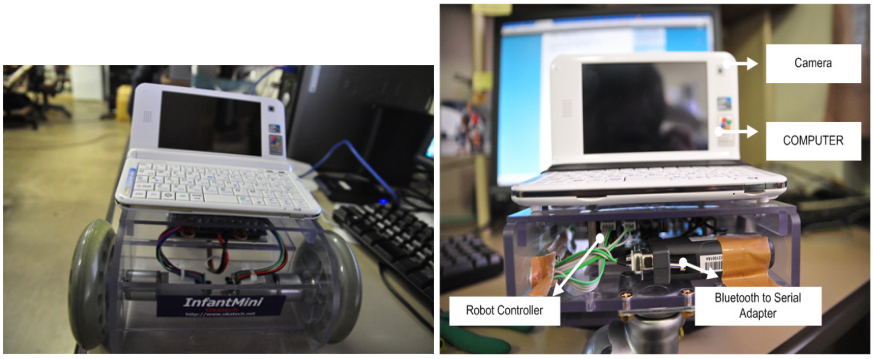


Fig. 8 System Configuration of the proposed service robot controlled just by sight



(a) Outlook (b) Components

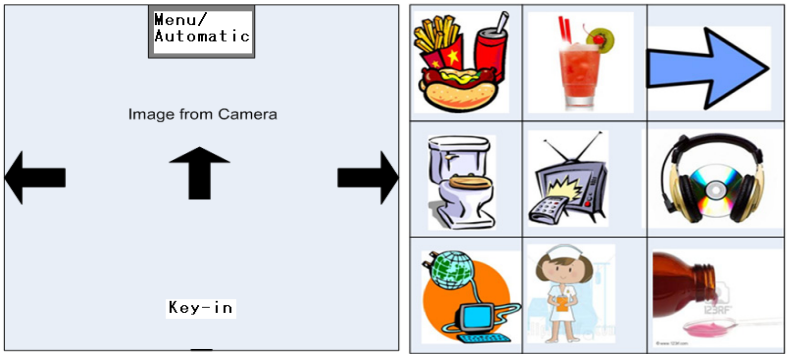
Fig. 9 Outlook and components of the proposed robot (InfantMini)

Table 4 Specification of OkaTech Infant Mini Robot Platform

Size	W213xD233xT105mm
Weight	1.7kg
Max.payload	5kg
Wheel_rotation	8000rpm
Speed	18.12cm/s(0.65km/h)
Motor	12V,1.5W(with_encoder)
Tire	100mm
Caster	60mm
Battery	Lithum_polimer(11.1V,2200mAh)

Experiments are conducted in the Arai laboratory of which the layout is illustrated in Fig. 11. One shot image of the robot during the experiments is shown in Fig. 12.

The test results shows marginal performance (it takes 177 s to 202 s, 190.2 s in average). Also, it has communication aid capability. Therefore, users can enjoy communication with the peoples in front of the robot.



(a) Automatic control (b) Menu control

Fig. 10 Menu for service robot control

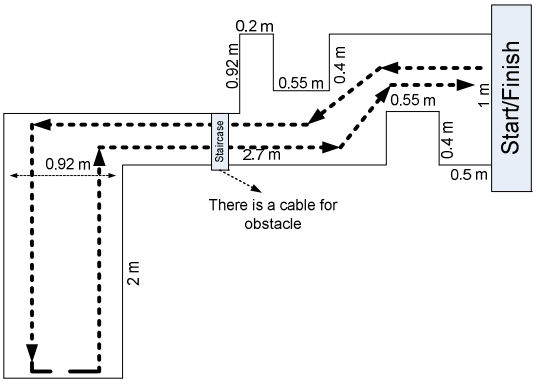


Fig. 11 Route for domestic robot in the experiments

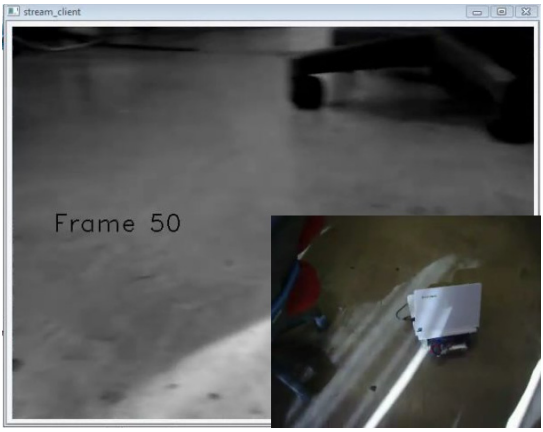


Fig. 12 Acquired image with the camera mounted at the tip of service robot and the image of service robot from the hand held camera

4 Conclusion

Computer input just by sight for disable persons is proposed together with its applications. It is confirmed that communication aids, phoning aids, service robot control, information collection aids, etc. are available by using the proposed computer input just by sight.

Smartphone and i-phone based computer input just by sight is also available. Applicable ranges of the smart-phone, i-phone based system are much wider than the PC based computer input just by sight. It also is applicable for wearable computing device. These applications will be demonstrated in the near future.

Acknowledgement The author would like to thank Dr. Ronny Mardiyanto of ITS and Herman Tolle of Brawejaya University, Indonesia for his effort for creation of systems and conduction of the experiments.

References

1. Arai, K., Uwataki, H.: Computer input based on line of sight estimation using detecting method for cornea center which allows users' movements. Journal of Electric Society of Japan, Tansaction C **127**(7), 1107–1114 (2007)
2. Arai, K., Yamaura, M.: Blink detection accuracy improvement by means of morphologic filter in designated key selection and determination for computer input by human eyes only. Journal of Image Electronics Society of Japan **37**(5), 601–609 (2008)
3. Purwanto, D., Mardiyanto, R., Arai, K.: Electric wheel chair control with gaze detection and eye blinking. In: Proceedings of the International Symposium on Artificial Life and Robotics, GS9-4 (2009)

4. Purwanto, D., Mardiyanto, R., Arai, K.: Electric wheel chair control with gaze detection and eye blinking. Artificial Life and Robotics, AROB Journal **14**(694), 397–400 (2009)
5. Arai, K., Mardiyanto, R.: Computer input by human eyes only with blink detection using Gabor filter. Journal of Visualization Society of Japan **29**(2), 87–90 (2009)
6. Arai, K., Mardiyanto, R.: Real time blinking detection based on Gabor filter. International Journal of Human Computer Interaction **1**(3), 33–45 (2010)
7. Arai, K., Yamaura, M.: Computer input with human eyes only using two Purkinje images which works in a real time basis without calibration. International Journal of Human Computer Interaction **1**(3), 71–82 (2010)
8. Arai, K., Mardiyanto, R.: Camera mouse and keyboard for handicap person with trouble shooting capability, recovery and complete mouse events. International Journal of Human Computer Interaction **1**(3), 46–56 (2010)
9. Arai, K., Uwataki, H.: Computer input system based on viewing vector estimation with iris center detection from face image acquired with web camera allowing users' movement. Electronics and Communication in Japan **92**(5), 31–40 (2009)
10. Arai, K., Yajima, K.: Communication Aid and Computer Input System with Human Eyes Only. Electronics and Communications in Japan **93**(12), 1–9 (2010)
11. Arai, K., Mardiyanto, R.: A prototype of electric wheel chair control by eye only for paralyzed user. Journal of Robotics and Mechatronics **23**(1), 66–75 (2010)
12. Arai, K., Yajima, K.: Robot arm utilized having meal support system based on computer input by human eyes only. International Journal of Human-Computer Interaction **2**(1), 120–128 (2011)
13. Arai, K., Mardiyanto, R.: A prototype of electric wheel chair controlled by eyes only for paralyzed users. Journal of Robotics and Mechatronics **23**(1), 66–75 (2011)
14. Arai, K., Herman, T.: Method for extraction product information from TV commercial. International Journal of Advanced Computer Science and Applications **2**(8), 125–131 (2011)
15. Arai, K., Herman, T.: Text extraction from TV commercial using blob extraction method. International Journal of Research and Review of Computer Science **2**(3), 895–899 (2011)
16. Arai, K., Mardiyanto, R.: Autonomous control of eye based electric wheel chair with obstacle avoidance and shortest path finding based on Dijkstra algorithm. International Journal of Advanced Computer Science and Applications **2**(12), 19–25 (2011)
17. Arai, K.: Mardiyanto, R:Eye-based human-computer interaction allowing phoning, reading e-book/e-comic/e-learning, Internet browsing and TV information extraction. International Journal of Advanced Computer Science and Applications **2**(12), 26–32 (2011)
18. Arai, K., Mardiyanto, R.: Eye based electric wheel chair control system-I(eye) can control EWC-. International Journal of Advanced Computer Science and Applications **2**(12), 98–105 (2011)
19. Arai, K., Mardiyanto, R.: Evaluation of users' impact for using the proposed eye based HCI with moving and fixed keyboard by using eeg signals. International Journal of Research and Reviews on Computer Science **2**(6), 1228–1234 (2011)
20. Arai, K., Mardiyanto, R.: Electric wheel chair controlled by human eyes only with obstacle avoidance. International Journal of Research and Reviews on Computer Science **2**(6), 1235–1242 (2011)

21. Arai, K.: Mardiyanto, R.: Evaluation of users' impact for using the proposed eye based HCI with moving and fixed keyboard by using eeg signals, International Journal of Research and review on Computer. Science **2**(6), 1228–1234 (2012)
22. Arai, K., Mardiyanto, R.: Electric wheel chair controlled by human eyes only with obstacle avoidance. International Journal of Research and review on Computer Science **2**(6), 1235–1242 (2011)
23. Arai, K., Mardiyanto, R.: Robot arm control with human eyes only and its application to help having meal for patients. Journal of Electrical Engineering Society of Japan, Transaction C, **C132**(3), 416–423 (2012)
24. Arai, K.: Human-Comuter Interaction with human eyes only and its applications. Journal of Image Electronics Society of Japan **41**(3), 296–301 (2012)
25. Mardiyanto, R., Arai, K.: Eye-based Human Computer Interaction (HCI) A new keyboard for improving accuracy and minimizing fatigue effect. Scientific Journal Kursor **6**(3), 1–4 (2012). ISSN 0216-0544
26. Arai, K., Mardiyanto, R.: Moving keyboard for eye-based Human Computer Interaction: HCI. Journal of Image and Electronics Society of Japan **41**(4), 398–405 (2012)
27. Arai, K., Mardiyanto, R.: Service robot which is controlled by human eyes only with voice communication capability. Journal of Image Electronics Society of Japan **41**(5), 535–542 (2012)
28. Arai, K.: Moving domestic robot control method based on creating and sharing maps with shortest path finding and obstacle avoidance. International Journal of Advanced Research in Artificial Intelligence **2**(2), 23–28 (2013)
29. Arai, K., Mardiyanto, R.: Eye-based domestic robot allowing patient to be self-services and communications remotely. International Journal of Advanced Research in Artificial Intelligence **2**(2), 29–33 (2013)
30. Arai, K., Mardiyanto, R.: Camera mouse including "Ctl-Alt-Del" key operation using gaze blink and mouth shape. International Journal of Advanced Computer Science and Applications **4**(3), 83–91 (2013)
31. Arai, K., Mardiaynto, R.: Method for psychological status estimation by gaze location monitoring using eye-based Human-Computer Interaction. International Journal of Advanced Computer Science and Applications **4**(3), 199–206 (2013)
32. Arai, K., Hasegawa, K.: Method for psychological status monitoring with line of sight vector changes (Human eyes movements) detected with wearing glass. International Journal of Advanced Research in Artificial Intelligence **2**(6), 65–70 (2013)
33. Arai, K.: Wearable computing system with input output devices based on eye-based Human Computer Interaction: HCI allowing location based web services. International Journal of Advanced Research in Artificial Intelligence **2**(8), 34–39 (2013)
34. Arai, K., Mardiyanto, R.: Speed and vibration performance as well as obstacle avoidance performance of electric wheel chair controlled by human eyes only. International Journal of Advanced Research in Artificial Intelligence **3**(1), 8–15 (2014)
35. Arai, K.: Service robot with comminucation aid together with routing controlled by human eyes. Journal of Image Laboratory **25**(6), 24–29 (2014)
36. Arai, K.: Information collection service system by human eyes for disable persons. Journal of Image Laboratory **25**(11), 1–7 (2014)
37. Arai, K.: Relations between psychological status and eye movements. International Journal of Advanced Research on Artificial Intelligence **4**(6), 16–22 (2015)
38. Arai, K.: Computer input by human eyes only and its applications. In: Intelligent Systems in Science and Information 2014. Studies in Computer Intelligence, vol. 591, pp. 1–22. Springer Publishing Co. Ltd. (2015)

39. Shaokun, W., Xiao, X. Hongwei, Z.: The wireless remote control car system based on arm9. In: International Conference on Instrumentation, Measurement, Computer, Communication and Control, pp. 887–890, October 2011
40. Chengjun, D., Bingsen, Y., Ping, D.: The remote control of mobile robot based on embedded technology. In: International Conference on Measuring Technology and Mechatronics Automation, pp. 907–910, January 2011
41. Zhigang, N., Yanbo, W.: Research on wireless remote control for coal mine detection robot. In: International Conference on Digital Manufacturing and Automation, pp. 315–318, December 2010
42. Goldstain, O.H., Ben-Gal, I., Bukchin, Y. Evaluation of telerobotic interface components for teaching robot operation. In: IEEE Transactions on Learning Technologies, pp. 365–376 (2011)
43. Zhigang, N., Yanbo, W.: Research on wireless remote control for coal mine detection robot. In: International Conference on Digital Manufacturing and Automation, pp. 315–318, December 2010
44. Artem, E., Bagayev, D.: System remote control of the robotized complex – Pegas. In: East-West Design & Test Symposium, pp. 358–361, September 2010
45. Lai, W-D.: ZigBee remote control interface for a robot dog design for older users. In: Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication, pp. 1–5, February 2011
46. Felzer, T., Nordmann, R.: Alternative text entry using different input methods. In: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 10–17 (2006)
47. Majaranta, P., Ahola, U., Špakov, O.: Fast gaze typing with an adjustable dwell time. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 357–360 (2009)
48. Ng, S.C., Raveendran, P.: EEG Peak Alpha Frequency as an Indicator for Physical Fatigue. In: Proceedings of 11th Mediterranean Conference on Medical and Biomedical Engineering and Computing 2007, pp. 517–520 (2007)
49. Klimesch, W.: EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Research Reviews, Rev **29**(2–3), 169–195 (1999)

User Impressions About Distinct Approaches to Layout Design of Personalized Content

Anelise Schunk, Francine Bergmann, Ricardo Piccoli, Angelina Ziesemer, Isabel Manssour, João Oliveira and Milene Silveira

Abstract Nowadays, a single person may produce large quantities of pictures, text and other digital content in a variety of devices, adding to it information collected from the Web. Usually, those information items are organized later to be presented/shared in more than one output device. Solutions for organizing information into personalized documents with a suitable design have been proposed for text and images, either in a fully automatic or interactive fashion. To analyze the use of two distinct approaches, we implemented an application using two layout design algorithms and performed user studies to collect user impressions. We present here the two algorithms as well as the results and the analysis of data collected in the user studies showing their preferences.

Keywords Digital documents · Authoring · Personalized content · Automatic · Physics-based

1 Introduction

The Internet has changed the way people perform many tasks and increased the importance of digital documents. Such documents may be produced by the users themselves as single information items or collections of items assembled into a larger document. Some simple examples of such documents may be calendars, news clippings, single page flyers, greeting cards, photo albums or travel guides. In the case of travel guides the text items, images, notes and links can be inserted according to user preferences and interests, either before traveling or afterwards. Before a trip, such a document may be produced by the user, printed and used as an inexpensive travel guide for the day or kept in digital form in a mobile device, whereas afterwards it may

A. Schunk · F. Bergmann · R. Piccoli · A. Ziesemer · I. Manssour(✉) · J. Oliveira · M. Silveira
PUCRS — Faculdade de Informática, Porto Alegre, RS, Brazil
e-mail: {anelise.schunk,francine.bergmann,ricardo.piccoli,angelina.ziesemer}@acad.pucrs.br,
{isabel.manssour,joao.souza,milene.silveira}@pucrs.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_87

be used as a travel journal or memento. However, for a user with little or no design experience, one of the problems of producing such documents (especially when they are to be shared with other users) is to easily and quickly distribute elements on a page. Moreover, additional visual quality is required when those documents are to be printed or presented to an audience. In particular, it is difficult for amateurs to obtain aesthetic quality similar to that achieved on a daily basis by professionals. Although there are tools to help a designer to produce layouts, those tools do not automatically generate a layout by themselves and are time-consuming and fairly hard to use by non-specialists. Taking into account that new approaches for non-specialized users have to be easy to use, we developed two algorithms to help the layout design of personalized content:

– a fully automatic algorithm able to place text and images on a page or split content across several pages; thus, items that are dragged into a page are automatically positioned and a new layout is generated;
– an interactive algorithm based on a physics engine able to organize items (e.g., images, text boxes, shapes and decorations) on a page; thus, items that are already in a page, position themselves to make room for new items that are dragged into the page due to a physics-based force model.

The main goal of this work is to analyze the adequacy of these two distinct approaches for layout design of personalized documents. The analysis was performed through a user study with a calendar application that we implemented using these two approaches. We performed a user study with 17 users to obtain their impressions about the proposed approaches. We concluded that each approach has its niches and fits different situations and applications.

2 Related Works

The page layout problem has been studied extensively in the literature [12]. Algorithms for generating designs are usually modeled after optimization approaches that attempt to maximize a set of aesthetic measures [21]. Depending on the approach, optimizing layouts can be a computationally challenging task [12, 14], and thus several computational techniques and layout models have been proposed in order to make the problem tractable. Concerning the different techniques used for generating layouts, works in the literature can be broadly classified into one of the four categories described below.

2.1 Template-Based

For most high-quality publications such as monthly magazines, graphic artists produce a collection of page templates to be used repeatedly in the publication, thereby

giving it a distinctive look. A library of previously generated templates is used to produce document instances: typically, templates are composed of fixed elements and variable regions where content is to be inserted. Templates may be entirely static [7, 18] or have adaptable parts to better suit the input content [5, 22].

Adaptive templates are normally specified in a higher level of abstraction (as opposed to specifying absolute positions for items), and employ linear constraint-solving techniques to produce instances. In both cases, a template-selection and content splitting mechanism are required in order to produce paginated output.

Lin [15] describes an approach to create templates adjusting content but with constraints to adapt text regions that will receive text items of different sizes using a linear relationship between width and height of the text boxes. The approach of Piccoli et al. [18] explores the division of work in two steps – pagination and rendering –, which results in a faster algorithm and a more flexible architecture.

2.2 Stochastic Optimization and Meta-Heuristics

In several works [8, 13, 21] document layouts are modeled as high-level geometric primitives, and the problem thus is to find positions, shapes and sizes for input content in an empty page. The proposed approaches usually describe only a small set of rules for constraining item placement, and general stochastic optimization algorithms such as Simulated Annealing or Genetic Algorithms are employed to explore parts of the problem search space. An objective function then attempts to quantify the aesthetic quality of candidate solutions in terms of placement, alignment, and others.

Goldenberg [8] proposed a genetic algorithm to solve the layout problem through successive cuts on a page but beyond the problems inherent to genetic algorithms (e.g., non-determinism, low performance), the layouts were aesthetically poor, since aesthetic criteria as alignment and homogeneity [9] were combined into a single objective function, hampering the convergence to a suitable solution as multiobjective optimization models are usually more challenging to solve efficiently [21].

2.3 Force or Energy-Driven

For applications of document layout requiring interactive work such as diagram editors or assisted design, force or energy-driven methods can be employed. In a force-driven layout, items are modeled as physical entities that produce forces toward each other, such as magnets or springs. The interplay between these forces cause items to move and accommodate themselves in a surface (page) until the physical system reaches equilibrium, e. g., when the total kinetic energy of the system reaches zero. From an optimization point of view, the solution may employ a gradient descent [11] method to find a local minimum of the energy function of the physical systems.

These methods were used in graph layout problems [3], and also in some document layout problems [1, 20]. In the first approach an interface for the interactive construction of calendars and brochures is based on magnetic repulsion to avoid the elements overlapping and produce a balanced distribution. The main advantage of a force-driven approach is in an interactive setting, where a layout can be computed incrementally and presented intuitively to the user in animation form.

2.4 Deterministic Algorithms

Works in this category [2, 6, 17, 19] define a reasonable set of constraints for item placement, in a way that is suited to the use of efficient techniques, such as dynamic programming. A type of layout that can be efficiently generated this way is a guillotine layout, where a document page is recursively split into smaller sections until there are enough sections to hold each different item.

Oliveira [17] proposed two deterministic algorithms able to produce document layouts, each one directed to a specific application: the first one receives the content as rectangles with fixed geometry (width and height) that must be arranged on a page while disallowing changes in area and aspect ratio. The other algorithm allows for arbitrary texts and images to be placed in rectangles on a page. Those rectangles are selected by the algorithm to ensure an optimal solution. Newspapers [10] and similar publications can usually be constructed using this method.

3 Algorithms Used

The problem of distributing items on a rectangular page can be solved in different ways. The first algorithm requires information about the areas and aspect ratios of the regions where items are to be placed and partitions the page accordingly, preserving the item input order and solving a minimization problem to produce a page.

The second algorithm distributes items on a page as they are inserted by the user: each new item forces others to move and spread to achieve a balanced distribution on the page. The user may also move an item around and other items will react, making room for the item to move. In this case, each item was modeled as a magnetic charge and we try to minimize the total energy of the system, reaching a state of equilibrium. The next two sections explain both algorithms used to solve the layout problem.

3.1 Automatic Page Layout Approach

In this approach (described in more detail in [17]) the layout algorithm receives a sequence of elements and places them on a page that has dimensions informed by

the user. It is fast for generating layouts with less than 30 items and may be adapted to many types of documents (e.g., newspaper pages, photo albums or travel guides). Results are generally good even though this layout approach does not require user intervention. However, a question that can be raised is whether users are prepared to use a tool that does not require any interaction except providing the content.[1]

The algorithm does not use templates: the content itself (texts, images and captions) is used to estimate the area required by each item, so that the page can be partitioned accordingly. The algorithm also requires an integer number of columns to be used in the document, as well as two real values that represent the acceptable minimum and maximum aspect ratios for items on the page, avoiding a bad layout generated by very wide or narrow regions.

With the elements to be placed, the algorithm makes consecutive page bisections, dividing the page into regions and attempting to place each item into a region. The bisection is made by a horizontal or vertical line, splitting a region and, similarly, the list of elements is also divided. When a sequence of items is available for placement, a large number of splitting possibilities are tested and the best decision is the one that reduces the maximum error of the items to be placed. The placement error is evaluated according to the following requirements: the area allocated for an item has to be similar to the required area to place that item [17]; the aspect ratio of the allocated area must be constrained to a specified range; for pictures, there may be optional constraints on its size in the page. The algorithm explores each possible splitting choice recursively and returns the best possible placement or an indication that placement was not possible.

A layout is composed by the best sequence of splitting decisions and the algorithm uses the sequence that has the lowest error measure. A function that evaluates the error of an item when placed on an area considering the item and area information as well as the number of columns on the page is given in Figure 1.

This method for generating layouts can be adapted to different types of documents as newspapers [17], photo books [4], personalized calendars, brochures or newsletters.

3.2 Physics-Based Document Layout Approach

In this approach as described in Piccoli et al. [20], the user starts with an empty page and the document is constructed interactively as content is inserted or moved by the user. Any number of pages can be used in a document, but each page is independent from the others and no attempts are made to perform automatic pagination or flow items across pages. Items already placed on the page continuously interact with each

[1] As even small changes to the content tend to generate completely different layouts, this might cause surprise to users attempting to make incremental changes to a document.

Algorithm 1. PlacementError

```
Input: Item item, region width w, region height h, number of columns c
Output: Item placement error for the given region parameters
aspectRatio ← w/h
if aspectRatio < item.minAR or aspectRatio > item.maxAR then
    return ∞
end
if item is a picture then
    if c < item.minCols and c > item.maxCols then
        return ∞
    end
    return max (h/item.height, item.height/h)
else
    proportion ← (w ∗ h)/(pageWidth ∗ pageHeight)
    return max (item.area/proportion, proportion/item.area)
end
```

Fig. 1 Placement error evaluation function

other and the user may insert, delete or resize new items, as well as move them around by dragging.

Items are either texts or pictures that may be placed into arbitrary polygons: a polygon is defined by a center (x_1,y_1) and its shape $(x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$. Each simulation step computes new positions for all polygons on a page, finding new positions so that they are evenly placed on the page. This can be modeled as an energy minimization approach: forces are computed for each time step and items move accordingly. From a usability perspective, items should move smoothly on the page so that the user is able to follow their movement when interacting with the system. A velocity parameter $(vx_i; vy_i)$ was also added for each item, allowing them to move with inertia and a state of equilibrium is the state that minimizes the total kinetic energy ε:

$$\varepsilon = \sum_i w_i\, h_i\, \left(vx_i^2 + vy_i^2\right). \tag{1}$$

Positions and velocities are computed using simple Euler integration [16]. This updates positions and velocities explicitly in discrete time steps Δ_t, until the system converges to an energy minimum. The basic layout force is the magnetic repulsion between items. Every item generates forces toward every other item, based on their distances and sizes. As a consequence, items will try to be as far apart as possible from each other, so overlapping seldom occurs. Thus, for all items u, v, the resultant force vector $\overrightarrow{F_u}$ is given by

$$\overrightarrow{F_u} = \sum_{v \neq u} \overrightarrow{f_{u,v}} \tag{2}$$

From the resultant forces, new velocities and positions can be computed for the next time step. After computing the force vectors $\overrightarrow{F_i}$ for every item, they are applied for one time step. Assuming that item i is placed on position (x, y) at time t, we compute positions for time $t + 1$ as shown in Equation (3), where Δ is a damping coefficient in the range $[0 \ldots 1]$:

$$(x_i, y_i)^{t+1} = (x_i, y_i)^t + \Delta_t \ (vx_i, vy_i) \, . \tag{3}$$

One problem with this approach is that repulsion forces will move items infinitely far. Therefore, additional forces are incorporated into our model so that items are confined within the document page. The physics-based algorithm can be used, for example, to generate personalized greeting cards and calendars.

4 User Study

To compare the two approaches we implemented two versions of a personalized calendar generator: on the physics-based version, forces make the elements interact and move on the page, whereas the automatic version produces the layout without user interaction. Both versions have the same user interface and functions: add external or pre-loaded content and manipulate it to produce a calendar. The main menu allows selection of images, adding text, defining backgrounds and selecting polygons. Furthermore, it is possible to save a file to continue editing later or export a PDF file when the document is ready. Figure 2 shows the calendar creation user interface.

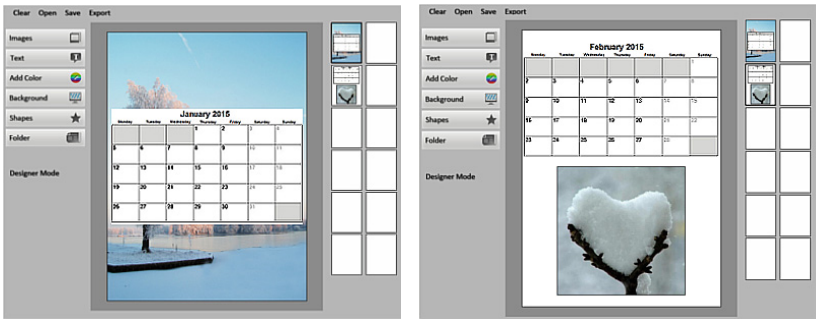


Fig. 2 Calendar user interface

In the physics-based version, user interaction is simple: items are dragged into the page and as the user moves them around other items make room for it due to the physics forces. Users may add monthly calendars as well as pictures, shapes and backgrounds.

The automatic version operates differently: as each item is dragged into the page the software adds it to the page, re-evaluates the layout and presents the new, finished layout. There is no animation to show the movement of items and the user is not able to drag and drop the items to absolute positions in the page.

4.1 Procedure

In a study with 17 users with varied professions, we asked them to fill a form with profile information and to use the two versions of the calendar generator, answering a questionnaire about each one. In order to not compromise the results, 9 users started with the automatic version and then used the physics-based version of the application; the other 8 did the opposite.

After trying out each prototype, participants had to answer a questionnaire specific to each prototype, and lastly, compare the two versions. We asked in what circumstances would they use each version, and found that the automatic version would be preferable to produce a simple calendar quickly, and the physics-based is better suited to produce a more creative or free-form calendar.

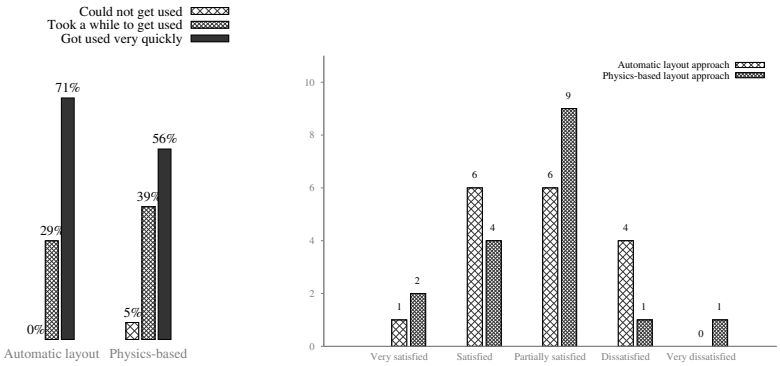
4.2 Participants

Among the 17 participants, 11 (64.7%) were female and 6 (35.3%) male. Seven participants were from 19 to 29 years old and 10 were 30 or more. With respect to their occupation/areas of expertise, 6 (35.3%) work with Information Technology, 3 worked at home (17.6%), 2 were teachers (11.8%) and the others (35.3%) have varied occupations (biologist, stylist, public officer, geologist, nutritionist and health care/nursing).

4.3 User Impressions

After trying out each prototype, participants had to answer a questionnaire specific to each prototype and finally compare the two versions. From the answers about each version we could detect the following impressions: All participants said that they could generate a calendar quickly with the automatic version, and 13 (76%) said that they could perform the same task with the physics-based version as well. Their general impressions about how they get used to each approach are presented in Figure 3(a). Considering the automatic approach, more than half of the participants (71%) got used to it very quickly, while 56% took a while to get used with physics-based approach.

Figure 3(b) summarizes the users' satisfaction with the calendars produced by each approach. Seven participants (41.1%) declared themselves very satisfied or satisfied with the calendar produced with the automatic version; 6 (35.3%) were partially satisfied; and 4 (23.6%) were unsatisfied or very unsatisfied. Considering the physics-based version, these numbers changed to 6 (35.3%), 9 (52.9%) and 2 (11.8%), respectively.



(a) Getting used with each approach (b) Users' satisfaction about the final calendar layouts

Fig. 3 User impressions from the study

Concerning the **automatic version** users liked the simple operation, fast results, and ease of use provided by an automatic generator of documents *("the possibility of automatically grouping photos and calendars reduces my work and effort in the calendars' construction"; "(. . .) I don't have patience to do much specialized editing"; "I can do a calendar very quickly (. . .), but I would like more editing freedom"), especially the auto-organization of the elements when a modification is made*. They complained about the quality of the final layouts *("(. . .) the photos do not go where I want them to go. I do not have (editing) freedom with this tool."; "the idea of automatic manipulation seems to be interesting when we are in a hurry, but I did not like the final result")* and also about the editing restrictions (they wanted to be able to position images in absolute positions, add borders and margins, and others).

In the **physics version** users liked the interaction with the page elements. We could observe that they like to be free to edit (as actions they cannot control are not well received). On the other hand, they complained about the auto-organization of the elements each time a new picture was added *("It sometimes rearranges what I'm doing, pushing the images towards each other. But it is not always the movement that I hope the tool would make"; "Sometimes the tool tries to "help me" changing the image locations, but not always in the way I expect."; "The tool tries to make you favors arranging the elements in the page, but in reality this hampers your work.")*, that they said disorganize the current page.

At the end of the questionnaire, participants had to choose a version and the preferred one was the physics-based approach with the requirement of providing more editing freedom. It shows that the ability to change the layout is very appreciated by the users, as they prefer an application that has a steeper learning curve and acts unexpectedly from time, but provides more editing freedom. After evaluating this scenario, we observe that users like ways to ease the generation of documents, but they are not prepared for editing documents without directly interacting with the result.

In summary, when asked about the possible uses of each version, the participants would use the automatic approach to produce a calendar quickly and the physics-based approach to produce a creative calendar. Despite this, 10 participants (58.8%) reported that they would print the final calendar generated with the automatic version and 11 (64.7%) with the physics-based version.

5 Final Considerations

This paper presented two algorithms for the layout design of personalized content (one fully automatic and the other using a physics-based interaction model) as well as the analysis of user impressions about the algorithms through the use of a simple prototype.

A calendar generator was implemented to provide a comparison between the two layout approaches. We noticed that for this application, the physics-based layout was more positively evaluated, because users felt they had more control about the final result, whereas that sensation seems to have been lost in a fully automated layout. Also, a negative experience for users of the calendar prototype was caused by the lack of editing options in the automatic version, such as control of image sizes or fixed positions, and the auto-organization of elements in the physics-based version. The automatic approach was less well received by users, but is still an effective solution for systems that have larger amounts of information to be organized in a layout providing the possibility of fast document generation by non-specialists.

As the participants highlighted, the approaches are better suited to different layout design situations: the automatic approach is better for rapid production and the physics-based approach is more suitable for creative exploration. We believe that when documents are built by the users themselves, they associate higher value to them and interactive methods may be recommended, as users feel they are more in control of the process. When documents have too much content or when the users are not required to take part in their construction, automatic methods that are able to handle large amounts of data with predictability are more easily accepted.

For future works, the physics-based approach has to be refined to provide finer control, increasing the perception that the amount of change in the page is small as new items are added. Another subject for further research the amount of time required by users to adapt to the new interaction model. Moreover, we would like to investigate new applications for the proposed automatic layout approaches.

Acknowledgments This paper was achieved in cooperation with Hewlett-Packard Brasil Ltda. using incentives of Brazilian Informatics Law (Law n. 8.2.48 of 1991).

References

1. Ali, K., Hartmann, K., Fuchs, G., Schumann, H.: Adaptive layout for interactive documents. In: Smart Graphics, pp. 247–254. Springer (2008)
2. Atkins, C.B.: Adaptive photo collection page layout. In: 2004 International Conference on Image Processing, ICIP 2004, vol. 5, pp. 2897–2900. IEEE (2004)
3. Battista, G.D., Eades, P., Tamassia, R., Tollis, I.G.: Graph drawing: algorithms for the visualization of graphs. Prentice Hall PTR (1998)
4. Bergmann, F.B., Manssour, I.H., Silveira, M.S., de Oliveira, J.B.S.: Automatic layout generation for digital photo albums: A user study. In: Human-Computer Interaction. Users and Contexts of Use, pp. 117–126. Springer (2013)
5. Damera-Venkata, N., Bento, J., O'Brien-Strain, E.: Probabilistic document model for automated document composition. In: Proceedings of the 11th ACM Symposium on Document Engineering, DocEng 2011, pp. 3–12. ACM, New York (2011)
6. Gange, G., Marriott, K., Stuckey, P.: Optimal guillotine layout. In: Proceedings of the 2012 ACM Symposium on Document Engineering, pp. 13–22. ACM (2012)
7. Giannetti, F.: An exploratory mapping strategy for web-driven magazines. In: Proceedings of the Eighth ACM Symposium on Document Engineering, DocEng 2008, pp. 223–229. ACM, New York (2008)
8. Goldenberg, E.: Automatic layout of variable-content print data. MCs Dissertation, School of Cognitive & Computing Sciences, Univeristy of Sussex, Brighton, UK, pp. 1–41 (2002)
9. Harrington, S.J., Naveda, J.F., Jones, R.P., Roetling, P., Thakkar, N.: Aesthetic measures for automated document layout. In: Proceedings of the 2004 ACM Symposium on Document Engineering. pp. 109–111. ACM (2004)
10. Harrower, T., Elman, J.M.: The newspaper designer's handbook. Brown & Benchmark Publishers, WCB (1995)
11. Hendrix, E.M., Boglárka, G., et al.: Introduction to nonlinear and global optimization. Springer, New York (2010)
12. Hurst, N., Li, W., Marriott, K.: Review of automatic document formatting. In: Proceedings of the 9th ACM Symposium on Document Engineering, DocEng 2009, pp. 99–108. ACM, New York (2009)
13. Johari, R., Marks, J., Partovi, A., Shieber, S.: Automatic yellow-pages pagination and layout. Journal of Heuristics **2**(4), 321–342 (1997)
14. Knuth, D.E., Plass, M.F.: Breaking paragraphs into lines. Software: Practice and Experience **11**(11), 1119–1184 (1981)
15. Lin, X.: Active layout engine: Algorithms and applications in variable data printing. Computer-Aided Design **38**(5), 444–456 (2006)
16. Müller, M., Stam, J., James, D., Thürey, N.: Real time physics: class notes. In: ACM SIGGRAPH 2008 Classes, p. 88. ACM (2008)
17. de Oliveira, J.B.S.: Two algorithms for automatic page layout and possible applications. Multimedia Tools and Applications **43**(3), 275–301 (2009)
18. Piccoli, R., Oliveira, J., Manssour, I.: Optimal pagination and content mapping for customized magazines. Journal of the Brazilian Computer Society **18**(4), 331–349 (2012)
19. Piccoli, R., Oliveira, J.B.: Balancing font sizes for flexibility in automated document layout. In: Proceedings of the 2013 ACM Symposium on Document Engineering, DocEng 2013, pp. 151–160. ACM, New York (2013)

20. Piccoli, R.F.B., Chamun, R., Cogo, N.C., de Oliveira, J.B.S., Manssour, I.H.: A novel physics-based interaction model for free document layout. In: Proceedings of the 11th ACM Symposium on Document Engineering, pp. 153–162. ACM (2011)
21. Purvis, L., Harrington, S., O'Sullivan, B., Freuder, E.C.: Creating personalized documents: an optimization approach. In: Proceedings of the 2003 ACM Symposium on Document Engineering, pp. 68–77. ACM (2003)
22. Schrier, E., Dontcheva, M., Acobs, C., Wade, G., Salesin, D.: Adaptive layout for dynamically aggregated documents. In: Proceedings of the 13th International Conference on Intelligent User Interfaces, pp. 99–108. ACM (2008)

Video Compression Using Variable Block Size Motion Compensation with Selective Subpixel Accuracy in Redundant Wavelet Transform

Ahmed Suliman and Robert Li

Abstract This paper proposed a high performance video coding system in redundant wavelet domain utilizing a multihypothesis variable size block motion compensation technique. The redundant wavelet transform (RDWT) provides several advantages over the traditional wavelet transform (DWT). The RDWT retains all the phase information of wavelet transform and provides multiple prediction possibilities in ME/MC in transform domain. The paper presents a new adaptive partitioning scheme and decision criteria that utilizes more effectively the motion content of a frame in terms of the size and shape of the blocks. The experimental results show that our proposed MH-VSBMC has a better performance and uses less number of partition blocks. In addition, the approach using selective strategy for subpixel accuracy achieves similar PSNR and other quality values while using less computation steps.

Keywords Video compression · Motion compensation · Block matching · Redundant Wavelet Transform · Subpixel accuracy

1 Introduction

The VSBMC technique has been used in the video coding standards, such as H.264 [1] and MPEG-4 [2], and a discrete cosine transform (DCT) is used to decorrelate the residual frame. However, for the last few years the RDWT has been an attractive subject for video coding and compression. The RDWT not only provides wavelet-based video coding with shift- invariant property, but also increases the precision of the motion vectors by increasing the number of hypothesis of the prediction [3].

A. Suliman · R. Li(✉)
Department of Electrical Engineering, North Carolina A&T State University, Greensboro, NC, USA
e-mail: eeli@ncat.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_88

The VSBMC technique for motion vector[4], is to partition a frame in such a way that regions with uniform translational motions are divided into larger blocks while those containing complicated motions into smaller blocks. The VSBMC technique generally relies on a binary tree or a quadtree decomposition structure. Such a scheme is efficient in representing the partitioning, but the resulting blocks are restricted to be rectangular and the sizes and locations of the blocks are also restricted by the tree structure. Kim and Lee [5] proposed a non-rectangular blocks, where each block is repeatedly partitioned into smaller blocks along a quadtree structure until a predefined maximum partition level is reached. The partitioning is indirectly represented by a full-quadtree with a one-bit code for each node or leaf, indicating whether or not a motion vector is transmitted for the corresponding block denoted by the leaf or the node of the tree. The drawback of this method is the inefficiency of its tree code scheme, which always requires a fixed number of bits regardless of the final partitioning result. A two phases partitioning technique [6] fixed the coding inefficiency by introducing a two-bit quadtree code scheme. This scheme is to be used here for its efficiency.

Also, subpixel motion estimation plays an important role in compression efficiency within modern video codecs such as H.264 and MPEG-4. Subpixel motion estimation is implemented within these standards using interpolated values at 1/2 or 1/4 subpixel accuracy. Such interpolation gives a good reduction in residual energy for each predicted macroblock and, therefore, improves compression. However, this leads to a significant increase in computational complexity at the encoder [15]. We try to improve this situation in our effort. In addition, multihypothesis motion compensation (MHMC) [7] forms a prediction of pixel s(x, y, t) in the current frame as a combination of multiple predictions in an effort to combat the uncertainty inherent in the ME process [8]. One approach to MHMC is to implement multihypothesis prediction in the wavelet domain. Other approaches include fractional-pixel MC [9] and overlapped block motion compensation(OBMC) [10].

This paper presents a novel approach to VSBMC in the redundant wavelet domain, where a new adaptive partitioning scheme has efficiently utilized the multi-hypothesis motion content of a frame to determine the shapes and sizes of the block. Moreover, the paper proposed a new method of selective accuracy algorithm that reduces unnecessary computation produced by the subpixel accuracy.

2 Redundant Wavelet Transform

The Redundant wavelet transform (RDWT) can be considered to be an approximation to the continuous wavelet transform that removes the downsampling operation from the traditional critically sampled DWT to produce an overcomplete representation. The shift-variance characteristic of the DWT arises from its use of downsampling, while the RDWT is shift invariant since the spatial sampling rate is fixed across scale. There are several ways to implement the RDWT, and several ways to represent the resulting overcomplete set of coefficients. The most obvious implementation is a direct implementation of the algorithm 'a trous. We obtain results in subbands that are exactly the same size as the original signal. The advantage of this “spatially coherent”

representation is that each RDWT coefficient is located within its subband in its spatially correct position [11].

In essence, the RDWT removes the downsampling operation from the traditional critically sampled discrete wavelet transform (DWT) to produce an overcomplete representation. The RDWT is easily created by employing filtering as in the usual critically sampled DWT; however, all phases of downsampled coefficients are retained rather than discarded. Thus, the RDWT consists of a number of distinct, critically sampled DWTs, each one being a separate phase of the RDWT. In a J -scale 2D RDWT, there exist 4^j distinct critically sampled DWTs; each one corresponds to a unique choice between even and odd phase for the horizontal subsampling as well as the vertical upsampling at each scale of decomposition (Fig. 1).

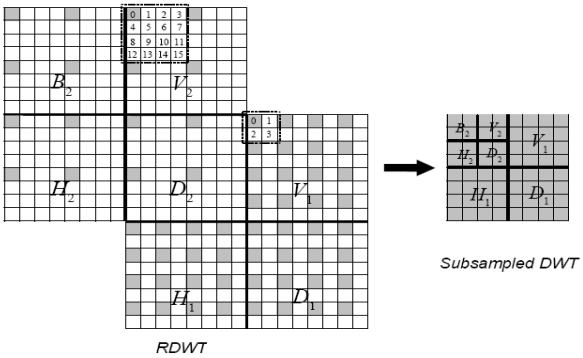


Fig. 1 2-scale RDWT subsampling recovering one of the 4^j critically sampled DWTs.

3 Overall Encoder Architecture

The encoder of our RDWT-MH video-coding system is depicted in Fig. 2.

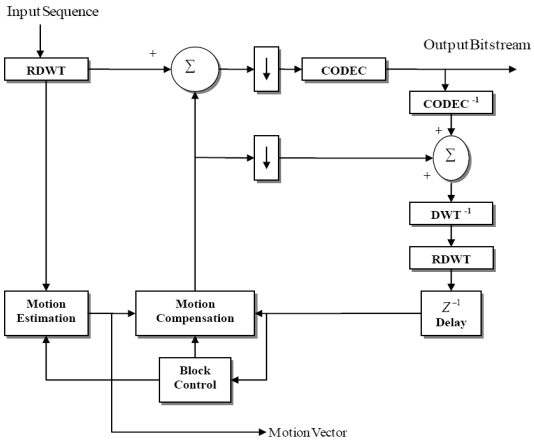


Fig. 2 Block diagram of the RW-ME/MC video-coding system.

The current and reference frames are transformed into RDWT coefficients. Both ME and MC take place in the redundant wavelet domain. In a J-scale RDWT decomposition, each block in the original spatial domain corresponds to 3J+1 blocks of the same size, one in each subband. The collection of these co-located blocks is called a set. In the ME procedure, block matching is used to determine the motion of each set as a whole. Specifically, a block-matching procedure uses a cross-subband distortion measure that sums absolute differences for each block of the set similar to the cross-subband ME procedure of [12]. To speed up the search, a 1-scale RDWT, rather than the full J -scale transform, is used for the block-matching ME procedure. Therefore, each frame will produce four subbands (LL, LH, HL and HH).

An adaptive variable size window is used for the block search. The all-phase correlation mask and approximation subband (LL) are used to construct a multihypothesis decision criteria for the block size. The motion of the variable block sizes from the reference frame to the current frame is estimated in the RDWT domain, and motion vectors are transmitted to the decoder. Multihpothesis MC is accomplished by using a multiple reference frames (subband) algorithm to generate bidirectional prediction. Residing in the RDWT domain, the motion-compensated residual is itself redundant; consequently, it is downsampled before coding. The final encoding step consists of a wavelet-domain still-image coder. In this case, we use the SPIHT method [13].

4 Video Frame Partition Process

The new partition process proposed a new decision criterion that partitions a given frame into variable size region according to the motion information of the frame. The partitioning information is efficiently represented by a two-bit quadtree coding scheme. The frame partitioning is accomplished by: first; splitting 16x16 block into 8x8 and 4x4 blocks, second; merging four neighbors of 16x16 blocks into 32x32 block.

First, decide which 16x16 macroblock is a candidate to be split. As it was mentioned before, each frame has four subbands in redundant wavelet domain. The direct multiplication of the RDWT coefficients at adjacent scales (all-phase correlation mask) distinguishes important features from the background due to the fact that wavelet coefficient magnitudes are correlated across scales. We will use all-phase correlation mask of the reference frame to determine which 16x16 macroblock is the candidate to be split by setting a number of thresholds.

To create the all-phase correlation mask for the current frame, we multiply the vertical (V), horizontal (H), and diagonal (D) bands together across scales and combine the products. Figure 3 shows an example of correlation mask for “Foreman” sequence. Two-bit code scheme (TBC) [4] is used to keep track of the different block sizes and its position at the decoder side. The first bit is used to distinguish between a leaf and a non-leaf node, while the second is used to indicate whether a motion vector is being transmitted. In the first-bit position of

the code, a 0 or 1 represents, respectively, a leaf or a non-leaf node; in the second-bit position, a 1 represents the transmission of a motion vector and a 0 the lack of it. For example, the code “10” represents a non-leaf node with no motion vector being transmitted, while the code “01” represents a leaf with its motion vector being transmitted. Since a block at the bottom level always corresponds to a leaf, only one bit regarding the transmission of the motion vector is needed for such a block. Figure 4 shows an example of MV's obtained after splitting and merging processes and its corresponding quardtree.

Fig. 3 The image shows an example of “Foreman” sequence correlation mask.

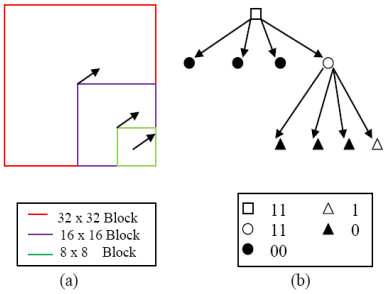


Fig. 4 (a) Example of MV's after splitting and merging process (b) Corresponding tree using the TBC system.

5 Selective Accuracy Algorithm

The subpixel accuracy is a powerful tool to achieve high accuracy, but it takes huge computation time. It uses a full search algorithm to find the accurate coordinates for each motion vector [15]. To perform subpixel motion estimation, the encoder interpolates pixel value at subpixel positions using pixel values at integer pixel position. Although the coding efficiency is highly increased by the subpixel motion estimation, the computational complexity of this repetitive subpixel motion search is very large in comparison with the case of fast integer-pixel motion search. Moreover, the subpixel motion estimation without considering the macroblock characteristics is not efficient in terms of the computational complexity. To reduce this additional complexity, a new method of selective refinement algorithm (Figure 5) is desired for the low complexity encoder and it works in two steps:

Step 1: Use the decision tree from the variable size block matching to decide the size of the block. Notice that we do not include 32x32 and 4x4 blocks in this procedure. We assume that 32x32 block does not have important detailed texture and most likely its motion vector will be zero. We also assume that the MV for 4x4 block will end up to 2x2 if 3-step or 4-step block matching algorithm were used, therefore no reason to go further.

Step 2: Calculate Sum of Absolute Difference (SAD) for each 16x16 and 8x8 macroblocks from approximation subband (LL) and test against thresholds for 16x16 and for 8x8 macroblocks. If the macroblock SAD is less than a given threshold, then keep the half pixel and the integer accuracy unchanged. Otherwise calculate quarter pixel accuracy for 16x16 and half pixel accuracy for 8x8. The SAD is a sum of absolute difference between co-located macroblock in previous and current frame macroblock.

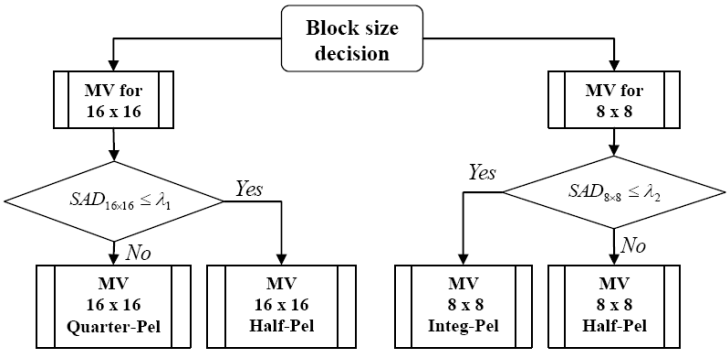


Fig. 5 Selective accuracy algorithm procedure.

6 Experimental Results

In the experimental results we use 60 frames, 256x256 pixels "News" sequence and 70 frames, 256x256 pixels "Foreman" sequence, Common Intermediate Format CIF (A standard video format used in video- conferencing). The sequence is in grayscale and has a temporal sampling of 25 frames/sec. The first frame is intra-encoded (I-frame) while all subsequent frames use ME/MC (P and B frames). All wavelet transforms (RDWT) use the Daubechies 9-7 filter with symmetric extension and a decomposition of J = 2 level. The parameters α, β and α_n are 0.4, 0.68 and 0.73 respectively. The core compression engine in all experiments is SPIHT. Since SPIHT produces an embedded coding, each frame of the sequence is coded at exactly the specified target rate with compression rate 0.5 bpp. In the experimental results, we use the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [14]. The SSIM is a method for measuring the similarity between two images, it can be viewed as a quality measure of one of the images being compared, provided the other image is regarded as of perfect quality. The dynamic range of the pixel values is 255 for 8-bit grayscale images.

The PSNR and SSIM were calculated for the coding system in Figure 2 using both our proposed MH-VSBMC and conventional FSBMC method using 16x16 block size. The experimental results in Figure 6 (with "Foreman" frame) and Table 1 show that our proposed MH-VSBMC method has a better performance in PSNR and SSIM, and uses less number of partition blocks. In addition, the

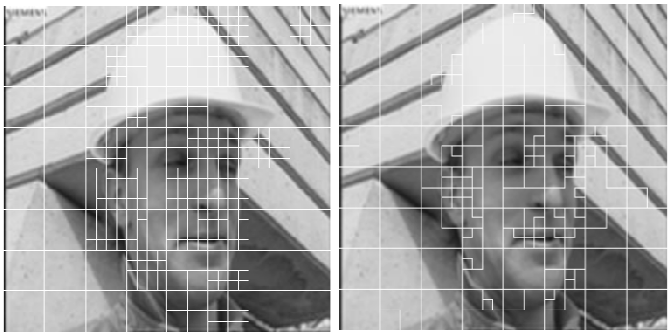


Fig. 6 The fourth frame "Foreman" (Left) The conventional VSBMC partition (Right) The proposed MH-VSBMC partition with less number of blocks used

Table 1 Averaged SSIM and PSNR for "News" and "Foreman" video frame

	News		Foreman	
	SSIM	PSNR	SSIM	PSNR
FSBMC+Subpixel	0.884	32.02	0.866	29.70
VSBMC+Subpixel	0.941	34.93	0.927	32.27
MB-VSBMC+Subpixel	0.987	35.71	0.954	33.06
MB-VSBMC+Selective	0.986	35.56	0.942	32.81

7 Conclusion

In this paper, we propose a MH-VSBMC system which incorporate the idea of MHMC and VSBMC in redundant wavelet domain. Recognizing the different phases in RDWT coefficients and viewing the motion from different perspectives, this method allows us to partition a frame more flexibly according to its motion content. Moreover, a new selective accuracy algorithm has been introduced to reduce unnecessary computation for the traditional 1/2 and 1/4 subpixel accuracy algorithm. Overall, our approach implemented here has a superior performance than the existing video compression techniques.

References

1. ITU-T, Advanced Video Coding for Generic Audio-visual Services, ITU-T Recommendation H.264, May 2003
2. ISO/IEC 14496-2, Information Technology-Coding of Audio-Visual Objects-Part 2: Visual, MPEG-4 Coding Standard (1999)
3. Cui, S., Wang, Y.: Redundant wavelet transform in video signal processing. In: International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV 2006) (2006)
4. Lee, J.: Optimal quadtree for variable block size motion estimation. ICIP 1996 (1996)

5. Kim, J.W., Lee, S.U.: Rate-distortion optimization between the hierarchical variable block size motion estimation and motion sequence coding. SPIE **2501**, 822–833 (1995)
6. Zhang, J., Omair, M., Swamy, M.: A new variable size block motion compensation. In: Proceedings of the IEEE (1997)
7. Sullivan, G.J.: Multi-hypothesis motion compensation for low bit-rate video coding. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Minneapolis, MN, April 1993, vol. 5, pp. 437–440 (1993)
8. Cui, S., Wang, Y., Fowler, J.E.: Multihypothesis motion compensation in the redundant wavelet domain. In: Proceedings of the International Conference on Image Processing, Barcelona, Spain, September 2003, vol. 2, pp. 53–56 (2003)
9. Girod, B.: Motion-compensating prediction with fractional-pel accuracy. IEEE Transactions on Communications **41**(4), 604–612 (1993)
10. Orchard, M.T., Sullivan, G.J.: Overlapped block motion compensation: An estimation-theoretic approach. IEEE Transactions on Image Processing **3**(5), 693–699 (1994)
11. Fowler, J.E.: Analysis of redundant wavelet multihypothesis for motion compensation. In: Proceedings of the IEEE Data Compression Conference, March 2006, pp. 352–361 (2006)
12. Park, H., Kim, H.: Motion estimation using lowband-shift method for wavelet-based moving picture coding. IEEE Transactions on Image Processing **9**(4), 577–587 (2000)
13. Said, A., Pearlman, W.A.: A new, fast, and efficient image codec based on set partitioning in hierarchical trees. IEEE Transactions on Circuits and Systems for Video Technology **6**(3), 243–250 (1996)
14. Wang, Z., Bovik, A.C., Sheikh, H.R., Simon-celli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing **13**, 600–612 (2004)
15. Jung, J., et al.: Fast subpel motion estimation using selective motion vector accuracy of inter-mode decision for H.264/AVC. In: IEEE International Symposium on Industrial Electronics (ISlE 2009), Seoul Olympic Park Hotel, Seoul, Korea, July 5–8, 2009

PPMark: An Architecture to Generate Privacy Labels Using TF-IDF Techniques and the Rabin Karp Algorithm

Diego Roberto Gonçalves de Pontes and Sergio Donizetti Zorzo

Abstract Layman and non-layman users often have difficulties to understand privacy policy texts. The amount of time spent on reading and comprehending a policy poses a challenge to the user, who rarely pays attention to what he or she is agreeing to. Given this scenario, this paper aims to facilitate privacy policy terms presentation regarding data collection and sharing by introducing a new format called Privacy Label. Using natural language processing techniques, a model able to extract information about data collection in privacy policies and present them in an automated and easy-to-understand way to the user was built. To validate this model we used a precision assessment method where the accuracy of the extracted information was measured. The precision of our model was 0.685 (69%) when recovering information regarding data handling, making it possible for the final user to understand which data is being collected without reading the whole policy. The PPMark architecture can facilitate the notice-and-choice by presenting privacy policy information in an alternative way for online users.

1 Introduction

There are a wide range of options of the online services to a enormous diversity of users. Users can access social networks, internet banking, e-mails and a number of other options. Online services have privacy policies to determine and inform what data they may collect and to what end. The user must read and agree with the privacy policy if he wants to use one of these services. However, according to McDonald and Cranor[11], users are bound not to read such policies. They simply accept the terms without knowing what data is being collected and why. Companies may use

D.R. Gonçalves de Pontes(✉) · S.D. Zorzo ·
Department of Computer Science, Federal University of São Carlos (UFSCar),
São Carlos, SP, Brazil
e-mail: {diego.pontes,zorzo}@dc.ufscar.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_89

these data to select what products they will advertise for a specific user, for example. Kelley et al. [8] developed an approach to show terms in privacy policies related to data collection and sharing in a graphical manner called Nutrition Label for Privacy. Such information was presented in a table and the authors used the Platform for Privacy Preferences[1] (P3P) to build it. However, the P3P was not adopted by many online companies. Given this scenario, this paper aims to process privacy policies from online services, which are written in natural language, extract information about data collection and fill the table proposed by Kelley et al. [8] with some modifications to better fit the scope of this work. This work used a number of well known natural language processing algorithms to extract information in privacy policy texts and to fill full the label privacy. The scope of this paper is not to go deep in artificial intelligence techniques, but rather to use what is already available to present privacy policies to users in an alternative automated format. In the following section we present related works. In section 3, we briefly describe concepts about the techniques used in this paper. In section 4 we describe in details the privacy policy information extraction model and how such information is presented to the user. In section 5 and 6 we explain how the Learning Module and and the Analysis Module was applied in this case study. Section 7 contains details about the assessment method used to validate our proposal. Results are presented in Section 8 and Section 9 concludes the paper.

2 Related Work

Costante et al. [5] developed an approach to analyze the integrity of privacy policies from online services. The user first chooses which information he wants to share, and given these choices the proposed method analyzes the privacy policy and check to see if the service collects any other data that the user did not specify or did not allow; If so, the user is notified. Such approach may prevent users from using a given service, since they can only use those that comply with all their privacy choices. However, users should rather be warned – not prevented – about what is being collected and then decide by themselves to use a service. Online stores, for instance, must collect user data and are widely used nowadays.

Lobato and Zorzo [10] did a research and developed an application that verifies privacy items in a digital service. This application, called PrivPerson, has as the main goal automate the analysis and the customize of the privacy level of the websites.

Conger et al. [4] did a study on the privacy of personal information and described it into four types of interested parties on user data. The description of the entities took place in four parts. The first part was the consumer or individual who has the ownership of the data and provides information in order to use an online service. The second part is the one who sells or provides services, which requires personal information to perform some operation. The third part involves several entities: a third party consulting the individual credit information, entities regulated by the government

[1] http://www.w3.org/P3P/

or even companies that offer personalized advertising based on data shared by the second part. The third part may be known, but does not participate in information sharing decisions of users at the transaction act with some service, that is, the third part is hidden from the user and the service you are using. Finally the fourth part, which is the one that offers more risks to user data and are illegal hackers, thieves or third-party employees who violate company policy.

Kelley et al. [8] developed an alternative model to present privacy policies from online services. Their proposal was based on food charts, which included nutritional values of food products. The idea was to extract the privacy policies information and provide this information in a private label format, where the rows and columns indicates what data will be collected and used. To extract information from privacy policies, the mechanism proposed by the authors used the P3P platform, which use has been discontinued by many e-commerce companies.

3 Background

In the development of the proposed architecture will be used the TF-IDF techniques and Karp algorithm.

The TF-IDF technique was selected to find keywords in a set of texts. In a text we can easily see that words such as "a", "the", "and", "or", "in" and others frequently appear. The TF-IDF technique tries not to take these words into account, focusing on keywords that are relevant in the text. To identify keywords, we not only need to check how many times a given word appears on one text, but also how frequently each word appears on other texts. The TF-IDF algorithm uses a semi-empirical approach and was first proposed by Jones [17]. Since then, a number of works were developed based on the original idea.

The Karp algorithm was selected because it is a technique implemented to find patterns in texts written in natural language. Chang and Wang [3] formally defines pattern matching as: given a text T of size n and a set of patterns $P = P_1, ..., P_r$ with an alphabet to size s, find occurrences of patterns P in T. Pattern matching techniques are methods to verify if there is patterns in a data set that using a given knowledge base.

There are many works regarding this field. Some techniques use character by character search, as described in [1, 2, 7, 12]. Other techniques use regular expressions and automata [18]. The algorithm *Rabin Karp*[2] [6] avoids doing a character by character comparison by using a search window. The computational cost of the *Karp Algorithm* algorithm is in the order of $O(n - m + 1)$[1]. The choice of using the *Karp Algorithm* algorithm was made based on the performance evaluation described in [1, 3] where it was observed that the Karp algorithm had good efficiency when dealing with moderate to big amount of data.

[2] Implementation available at: https://github.com/sarveshsaran/RabinKarp

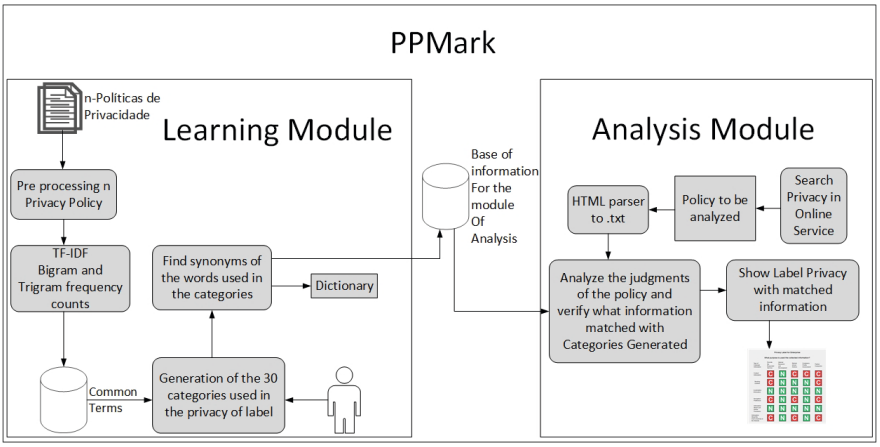


Fig. 1 Overview of the PPMark.

4 The PPMark Architeture

The architecture of the Privacy Policy Mark (PPMark) is divided in two modules: learning and analysis.

The Learning Module illustrated by Figure 1 is composed by four steps: **(i)** process *n* privacy policies; **(ii)** apply the TF-IDF technique on the *n* privacy policies by finding the most frequent bigrams and trigrams; **(iii)** generate data collection categories are based on bigrams and trigrams found; and **(iv)** to find synonyms of words from each category using a dictionary.

The steps of the Learning Module are detailed below.

1. ***Privacy policies*.** The first step consists of inputting privacy policies corpus with *n* policies that must be saved in ".txt" files.
2. ***Preprocessing of n privacy policies*.** This step is responsible for preparing the text to the learning process. In this preparation, special characters, accentuations e stopwords [3] are removed. After that, all words are capitalized and punctuations are removed – except for periods ".", that are kept to preserve complete sentences. Then the *stemming*[13, 16] is applied, leaving only the radicals of each word in the document. This technique is used to avoid taking into account words that have different endings but the same meaning (in Portuguese).
3. ***Find the most common terms using the TF-IDF technique*.** This step searches for the most common terms used in the policy, creating *bigrams*[4] and *trigrams*[5]. The technique adapted by Ramos [15] takes the processed documents as input and then does the search. The output are the most frequent *bigrams* and *trigrams*

[3] Words that are considered irrelevant

[4] Two-word set

[5] Three-word set

in the training corpus. They are used to generate rules to recognize similarities in a given privacy policy to the Analysis Module.

4. ***Create the categories used in the Privacy Label.*** The main purpose of this step is to generate the categories of collection and use of data. The categories are generated from bigrams and trigrams in the privacy policies. For example, two frequent *bigrams* in privacy policies are *inform-contat* and *prov-serv*. The expert must imply that these *bigrams* mean *"contact information to provide service"*, where the first *bigram* was related to the collected data and the second one to what end the data was collected. However, the *bigrams* do not need to necessairly appear in this order, hence we need an expert in this activity.
   Two sets of categories have been defined, the set of "type info" referred to in the column "Type of Collected Information", as shown in Figure 2, which contains terms related to all kinds of data to be collected. The group "used info" referred to in the column "What purpose the collected information is used?", shown in Figure 2, covers all terms relating to the use of the collected data purposes.
   These categories will be generated, selected and assigned to the privacy label. The set "type info" will be in the first column of the label and the set of "used info" will be the first row of the label, thus being able to map the type of data being collected and to what end, as shown in Figure 1.
   After the creation of the privacy categories, rules are generated for pattern matching, which takes place as follows: data sets of categories ("type info" and "used info") and a document are used as input (privacy policy). The technique will select each paragraph of the text and will cycle through each of two lines in two words and make comparisons with clusters of categories. For example, if a word is associated with the set "type info" and the other word is associated to the set of "used info", the corresponding cell is plotted on the label if the data is being collected.

The learning steps needs to be done only once for each language. The required inputs are a corpus from privacy policies, a dictionary and an expert.

After running the Learning Module, privacy categories will be stored in the Basic Terms Privacy Policy for further use.

The Analysis Module, shown in Figure 1, is composed of four steps: **(i)** search for privacy policies of online services; **(ii)** use the HTML-to-txt parser and the stemming technique; **(iii)** analyze sentences from the privacy policy and check Which information matches the categories; and **(iv)** create the Privacy Label with the matched information.

The steps of the Analysis Module are detailed below.

1. ***Search for privacy policies of online services.*** This step searches for privacy policies from online services. When the link to the text is found, it sends the privacy HTML code to the next step.
2. ***Use a HTML-to-txt parser and the stemming technique.*** This step takes as input the HTML code and converts it to a common ".txt" file without HTML

tags using a Java library called Jsoup[6]. Then the stemming technique is used to reduce every word to its radical, making it easier to compare the terms with the privacy policy terms database.

3. ***Analyze sentences from the privacy policy and check which information matches the categories.*** This step is in charge of analyzing each paragraph of the policy and comparing the text of tokens with the categories of privacy previously generated by the Learning Module. For each match the same category is displayed on the Privacy label.
4. ***Create the Privacy Label with the matched information.*** This activity creates the Privacy Label according to the occurrences of the categories created in the Learning Module.

5 Learning Module Execution

To run the Learning Module and then review the privacy policies with the acquired knowledge, a corpus of texts for training and testing for the language that is being used the proposal is necessary. To assess the purpose of this paper, two Brazilian Portuguese case studies were selected from two databases, one used for training and the other one for testing. The datasets came from the Alexa website[7], which is a website that ranks the most used online services in each country. The training and testing corpus, combined, had 60 privacy policies from the most accessed online services in Brazil, 50 of which were used for training and 10 were used for testing.

The steps 2 and 4 the Learning Module were applied in the training corpus to create the categories that will be shown next. Such categories were selected from the most frequent *bigrams* and *trigrams* in the privacy policies.

5.1 Categories

After finding the main *bigrams* and *trigrams*, such information was inspected by an expert and we narrowed it down to 30 categories, grouped as follows. The methodology used in this grouping was the same as the one proposed by Kelley et al.[8], where similar information was grouped under a name that would represent their intent. Below we briefly describe how the information was grouped: **(i)** contact information such as phones and e-mails were grouped under the same category, named "Contact information"; **(ii)** capture cookies, such as recovery sessions, shopping carts, reminder passwords were grouped under the same category, named "Reading Cookies"; **(iii)** geographical information such as city, address and place of birth under the same category, named "Geographical information"; **(iv)** browsing history,

[6] Available at http://jsoup.org/

[7] http://www.alexa.com/topsites/countries;0/BR

search history and other data related to user browsing were grouped under the same category, named "Navigation Preference"; **(v)** place of purchase, seen products and bought products were grouped under the same category, named "Online purchases"; and **(vi)** browser version, IP address, click tracking and login in the analyzed websites were grouped under the same category, named "Website activity".

The goals of the collected data were also grouped into 5 main categories. Such categories were chosen for being the most frequent in the analyzed privacy policies. The selected categories were: **(i)** provide the requested service ; **(ii)** internal research and development; **(iii)** marketing actions; **(iv)** consumer profiling; **(v)** partner companies.

The last category, Partner companies, is related to sharing the data with third-parties.

Finding Synonyms of N-grams and the Privacy Policy Terms Database. After generating the rules with the most frequente N-Grams, each N-Grams word synonym was found and stored in the Privacy Policy Terms Database.

5.2 Creating the Privacy Label with General Categories

After selecting the categories and defining the rules, the Privacy Label is created as illusted in Figure 2.

The Privacy Label shown in Figure 2 is fillfull using the Privacy Policy Terms Database, along with the most frequent terms and rules to analyze a new privacy policy. From this point on, the training is done and made available to analysis [8]. Once again, the training step needs to be done only once for a given language.

6 Analysis Module Execution

A policy text corpus, with ten ".txt" files, was given as input to the Analysis Module. For each document, we applied the pattern matching algorithm to sentences of each paragraph. On each paragraph, the algorithm tries to find the current category (a privacy label cell, as shown in Figure 2). For example, to match "Reading cookies" to "internal research and development" the algorithm first looks through the paragraph searching for the word "cookie". If found, then it starts searching for the words "internal", "research" and "development". If the algorithm finds, for instance, "research" or "development", then it assumes that the current paragraph is talking about reading cookies for internal research and/or development. A snippet of the code using the Rabin Karp algorithm for pattern matching is shown in Listing 1.1.

[8] To access the training database, please contact the authors

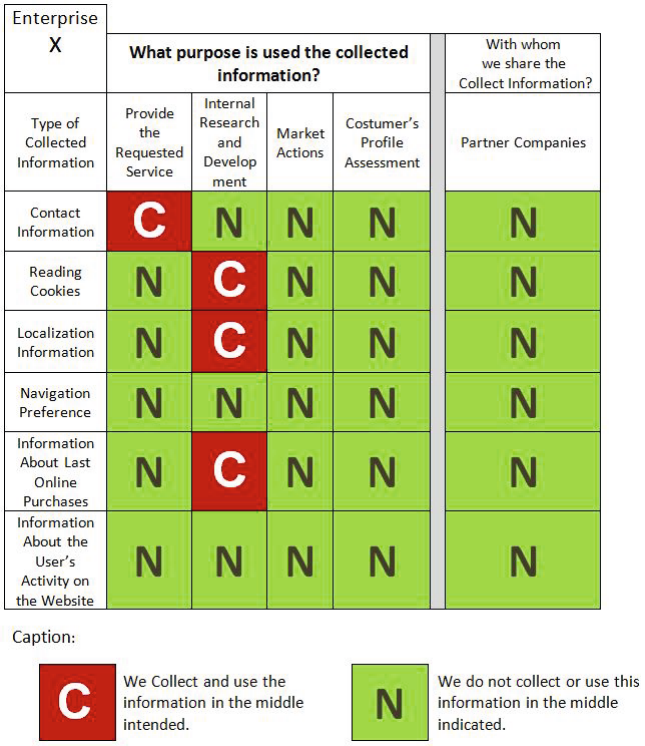


Fig. 2 The Privacy Label.

Listing 1.1 Excerpt code to match the patterns

```
...

public static int analysisTypeInfo(String tableCell, String paragraph) {
 int numMatchsTypeInfo = 0;
 String[][] patternTotal = table.get(tableCell);
 String[] patternTypeInfoInterested = patternTotal[typeInfo];
 for (int i = 0; i < patternTypeInfoInterested.length; i++) {
  String pattern = patternTypeInfoInterested[i];
  if(RabinKarp.findPattern(paragraph, pattern) >= 0 ){
   numMatchsTypeInfo++;
  }
 }

 return numMatchsTypeInfo;
}
...
public static int analysisUsedInfo(String tableCell, String paragraph) {
 int numMatchsUsedInfo = 0;
 String[][] patternTotal = table.get(tableCell);
 String[] patternUsedInfoInterested = patternTotal[usedInfo];
 for (int i = 0; i < patternUsedInfoInterested.length; i++) {
  String pattern = patternUsedInfoInterested[i];
```

```
23      if(RabinKarp.findPattern(paragraph, pattern) >= 0 ){
24       numMatchsUsedInfo++;
25      }
26     }
27     return numMatchsUsedInfo;
28    }
29
30    ...
```

According to Listing 1.1, we can see that the functions in line 5 and 19 take the label cell value and a paragraph as input. In lines 8 and 22, the algorithm takes the pattern to search for and, along with the paragraph, sends it to the Rabin Karp algorithm, which will return true if it finds a match (lines 10 and 27). At the end of each function, a number of matches is returned (for both type info and used info). If both values are greater than 1, it means that it was found some combination for some category. In this scenario, a character "C" (standing for "collects") is associated with that combination of type info and used info. On the other hand, a character "N" (standiing for "No collection") is used instead. These functions are applied to every category and every paragraph. In the end of the whole process, the privacy label matrix must be completely filled up, indicating which data is being collected and to what end.

7 Assessment Method

To assess the results of this proposal, we adapted the method used by Pérez Castillo et al. [14], in which the authors calculated the recovering precision of business process. The adaptation to this proposal is shown in Equation 1.

$$P = \frac{TP}{GT + FP} \tag{1}$$

where P is the precision result from extracting the terms, GT are Ground True found by the expert, TP are True Positives found by our model and FP are False Positives of our model. Péres Castillo et al. [14] defined the metrics used to assess the precision result, as shown in Table 1.

Table 1 Metrics to assess the performance of our technique, adapted from Pérez-Castillio et al. [14]

if	P	=	0.47			=	very low precision
if	0.47	<	P	=	0.56	=	low precision
if	0.56	<	P	=	0.63	=	moderate precision
if	0.63	<	P	=	0.72	=	high precision
if	P	>	0.72			=	very high precision

Table 2 Categories found by the expert in the privacy policies test cases

Policy	GT	TP	FP	Policy	GT	TP	FP
Text 1	27	19	1	**Text 6**	5	4	0
Text 2	17	14	3	**Text 7**	6	6	2
Text 3	14	12	0	**Text 8**	17	12	0
Text 4	4	3	3	**Text 9**	9	8	1
Text 5	6	6	3	**Text 10**	14	14	11
Total (Text 1 to Text 10)					**GT**	**TP**	**FP**
					119	98	24

To assess the precision of our model, we used 10 privacy policies text as test cases and evaluated the selection of categories.

The 10 test cases went through two analysis steps. The first one was done by an expert in which he identified which of the previously described categories were present in the policies texts. The second analysis was done by our model. The expert results and results of our model are shown in Table 2, with Ground True (GT) by expert, true positive (TP) and false positive (FP) metrics.

8 Results

The analysis done by the expert found 119 occurrences of data collection. For each privacy policy, a careful reading was done to verify if the categories were present in the text.

The analyis done by our model found 98 out of the 119 selected categories, and also got 24 false positives, as shown 2. The precision value of our model is given by Equation 1, where GT is the number of occurrences found by the expert, TP are true occurrences found by our model and FP are the false occurrences.

$$P = \frac{TP}{GT + FP} = \frac{98}{143} = 0.6853$$

Our model got a precision value of $P = 0.6853$ (69%), which according to Table 1 is considered high precision.

9 Conclusion and Future Work

According to the results presented in Section 8, we can say that our model has high precision in recovering information regarding data collection in privacy policies. However, the ambiguity in written text poses a challenge in the analysis of the texts. Kelley et al. [9] confirmed that the use of Privacy Labels can help users better

understand and pay more attention to privacy policies. The approach proposed by Kelley et al. [8] used files similar to XML with tags, which improved the analysis of policies. This work's proposal aims to improve the way privacy policies are shown to the user, so he or she can be aware of how their data is being handled by online services. Our technique does not use the P3P platform; rather, it process natural language written texts, extracts information regarding data handling and uses labels to make it easier for users to comprehend what they are agreeing to. With this work's results, a plugin is being developed for Google Chrome and Mozilla Firefox to warn users about privacy policies of the online services they use. This plugin will be first made available to be used with policies written in Portuguese, but we plan to train our model with English-written policies and make it available as well.

References

1. Alfred, V.: Algorithms for finding patterns in strings. Algorithms and Complexity **1**, 255 (2014)
2. Apostolico, A., Galil, Z.: Combinatorial algorithms on words, vol. 12. Springer Science & Business Media (2013)
3. Chang, C., Wang, H.: Comparison of two-dimensional string matching algorithms. In: 2012 International Conference on Computer Science and Electronics Engineering (ICCSEE), vol. 3, pp. 608–611, March 2012
4. Conger, S., Pratt, J.H., Loch, K.D.: Personal information privacy and emerging technologies. Information Systems Journal **23**(5), 401–417 (2013)
5. Costante, E., Sun, Y., Petković, M., den Hartog, J.: A machine learning solution to assess privacy policy completeness: (short paper). In: Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, WPES 2012, pp. 91–96. ACM, New York (2012). http://doi.acm.org/10.1145/2381966.2381979
6. Karp, R.M., Rabin, M.O.: Efficient randomized pattern-matching algorithms. IBM Journal of Research and Development **31**(2), 249–260 (1987)
7. Kearns, M., Pitt, L.: A polynomial-time algorithm for learning k-variable pattern languages from examples. In: Proceedings of the Second Annual ACM Workshop on Computational Learning Theory, pp. 57–71 (2014)
8. Kelley, P.G., Bresee, J., Cranor, L.F., Reeder, R.W.: A nutrition label for privacy. In: Proceedings of the 5th Symposium on Usable Privacy and Security, p. 4. ACM (2009)
9. Kelley, P.G., Cesca, L., Bresee, J., Cranor, L.F.: Standardizing privacy notices: an online study of the nutrition label approach. In: Proceedings of the SIGCHI Conference on Human factors in Computing Systems, pp. 1573–1582. ACM (2010)
10. Lobato, L.L., Zorzo, S.D.: Avaliação dos mecanismos de privacidade e personalização na web. Universidade Federal de São Carlos, São Paulo (2007)
11. McDonald, A.M., Cranor, L.F.: The cost of reading privacy policies. ISJLP **4**, 543 (2008)
12. Mooney, C.H., Roddick, J.F.: Sequential pattern mining - approaches and algorithms. ACM Comput. Surv. **45**(2), 19:1–19:39 (2013). http://doi.acm.org/10.1145/2431211.2431218
13. Orengo, V., Huyck, C.: A stemming algorithm for the portuguese language. In: Proceedings of the Eighth International Symposium on String Processing and Information Retrieval, SPIRE 2001, pp. 186–193, November 2001
14. Pérez-Castillo, R., García-Rodríguez de Guzmán, I., Piattini, M., Places, A.S.: A case study on business process recovery using an e-government system. Software: Practice and Experience **42**(2), 159–189 (2012)
15. Ramos, J.: Using TF-IDF to determine word relevance in document queries. In: First International Conference on Machine Learning (2003)

16. Savoy, J.: Light stemming approaches for the french, portuguese, german and hungarian languages. In: Proceedings of the 2006 ACM Symposium on Applied Computing, SAC 2006, pp. 1031–1035. ACM, New York (2006). http://doi.acm.org/10.1145/1141277.1141523
17. Sparck Jones, K.: A statistical interpretation of term specificity and its application in retrieval. Journal of documentation **28**(1), 11–21 (1972)
18. Watson, B.W.: A new regular grammar pattern matching algorithm. Theoretical Computer Science **299**(1), 509–521 (2003)

RGB and Hue Color in Pornography Detection

Awais Adnan and Muhammad Nawaz

Abstract Pornography and other such obscene material are now easily available on the internet, social media and other video and document sharing sites. Tons of videos are uploaded and shared daily on the Internets that have made it impossible to tag and filter obscene matters manually. Only feasible solution is automatic system that can detect pornography other similar materials in multimedia formats. Unfortunately most of such solutions are complicated and computationally expansive. Most of these solutions work with high-level features either special or in time space. On the other hand, low-level features are easy to detect and are fast in the computations, but contain less contextual information. In this article, a set of low-level features are analyzed to find the possibility of their use in pornography detection. Experimental results show that most of these low-level features are not suitable in this domain. However if different ratio of the same features (mix of these features) are used intelligently, they can be successfully used for the high-level content recognition as nudity and pornography detection.

Keywords Pornography · Brightness · Low-level features · Hue · Color models · Brightness

1 Introduction

Pornography can simply be defined as "any nudity or activity that stimulates erotic as opposed to aesthetic feelings in a community". According to available evidence, modern pornography started in Rome with the Italian renaissance of the print culture in fifteen century, with the publication of "I Modi", "the ways". This was composed of sixteen figures from Greco-Roman mythology designed by Giulio Romano and was published by Marcantonio Raimondi in 1524 [1] . In next two centuries pornographic art spread to other parts of the world but the pace was

A. Adnan(✉) · M. Nawaz
Institute of Management Sciences Peshawar,
A1-E5, Phase VII, Hayatabad Peshawar 25000, Pakistan
e-mail: awaisadnan@gmail.com, m.nawaz@imsciences.edu.pk

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_90

considerably slow. During the era between 1774 to 1788 a new form of political pornography arise in France with French Revolution that reached to its peak in 1789 when only in one year total pornographic novel published were one hundred and twelve [2]. Since then pornography spread all over the world where political pornography became commercial and was shifted from upper class to the all parts of the society.

Starting from the invention of "Daguerreotype" in 1839 to the latest Internet and mobile revolution, pornography have used these such advancements in technology for its creation and distribution very effectively. Invention of Video Cassette Recorder (VCR) in 1957 and Digital Versatile Disc (DVD) in 1995 spreads not only videos but also nudity and pornography at the level of streets and houses. With the addition of graphics and multimedia, Internet become more user friendly but at the same time it expands the boundaries of pornography in the entire world irrespective of any political or geographical boundaries [3]. Uncontrolled sharing of pictures and videos on social media and on other free media sharing sites are transforming this "information highway" into "nude-way".

According to a research conducted in 2010 [4], about 34% of cell phone users do video sharing and as a result on this, video traffic has been reached to about 40% of consumer internet traffic. In coming few years phones will be another main source of spreading such material in community. According to an estimate, around 70% of all prime time television shows (American and European) have some sexual contents and in the teen-shows percentage of such contents are above 83%. According to another research, on average there are 4.4 sexual scenes per hour while in teen shows it's about 6.7 per hours [5]. In an online discussion [6] about 56% were agreed that pornography is bad for society and should be banned because it is degrading human specially women and lowering human values.

Because of advancement in technology, nudity and pornography is no more a social problem, it is also a technological challenge to control this illegal use of such a material especially in underage groups. Although some controlling mechanisms are implemented, both in broadcast media and on the Internet, however most of these are based on manually tagging. Though some researchers have presented automatic solutions to detect nudity and pornography most of these solutions are using high-level features like shape, object, and gestures etc. Detection and extraction of these features is itself a challenging task. Not only these methods are computational expansive but the accuracy level obtained by these solutions is also too low to be used effectively.

In this paper different low-level features are examined to be used for pornography detection. For analysis two hundred clips are analyzed out of which 100 are nude or porno while the other set of hundred are normal visual clips. Experimental results reveal some interesting relationships between the presence of nudity in a scene and different low-level features like brightness and color distribution. Though these feature alone have no significant relationship with presence of nudity however their combination in different ratios give much better results.

2 Literature Review

Feature selection and extraction are two important decisions that play vital role in any processing activities especially in identification, classification, and retrieval problems. Different types of features are used in processing of images and videos where selection of these features mainly depends on the nature of the applications and the domain. All such features can be divided into two broad categories: low-level features and high-level features. Color, brightness, intensity of different colors channels, contrast and textures are examples of low-level features while features like shape, object, nature of contents, mood, type of scenes and shots come under high-level features. Most of the authors have used basic color information to extract different features like histogram, correlograms, moment etc. to retrieve image and videos. In an early work done by Wang et al. [7] an image passes through a series of tests to find objectionable contents. They uses a mix of features like color n edges and shape information using wavelets and moments. This is a slow solution that cannot be used efficiently on large videos where a huge numbers of frames are needed to process. Amir et al. [8] used simple color histogram in RGB space to retrieve videos. The main issue in this solution was the missing of locality information . This issues is solved by Yan and Hauptmann in [9] by splitting the image into 25 blocks that are arranged in 5 by 5 matrix to extract information containing color information with location reference.

Beside color, texture is another low-level feature that is used by many authors for image retrieval and understanding [10, 11]. Some authors like Foley et al. [12] have used shape feature with location information where they divide image into 4 by 4 non-overlapping blocks. Edges are extracted from these 16 blocks, arranged into 5 groups as 0, 45, 90, 135 and non-directional edges which form 16 x 5 = 80 bins histogram. In high-level features, shape and object are more common choices specially in content base retrieval and understanding [13]. Some special features like scene and shot information are used in video processing to get help in understanding higher level and advanced features like feeling , mood and emotions [14]. Many authors have used different audio features to make estimation about the content of scene. Williams and Stevens [15] used frequency of speech spectrum to determine five primary emotions—anger, fear, sadness happiness and disgust. Effects of these emotions are explained by three dimensions as Valence, Arousal and Control (V, A, C). Valence is the level of pleasure, in positive or negative, ranging from highly pleasant to extremely unpleasant [16]. Arousal, the other dimension is the experience intensity. Control is the dimension that differentiates between states that are similar in the first two dimensions. This Valence , Arousal paradigm is explained by A. Hanjalic in [14].

Skin detection is a common approach that is mostly used in nudity and pornographic detection. Forsyth et al. [17] are one of the pioneers in this work where they used smooth texture analysis and tuned skin filter to skin detection and then geometric shape analysis is used to identify nude parts of bodies. In another work conducted by Duan et al. [18] , skin color and SVM are used for nudity detection.

Their study is based purely on skin color detection and SVM. The images are first filtered by skin model and outputs are classified. Skin color and face are used by Rowley et al. in [19] to identify adult image. First, faces are detected using face detection system to eliminate the exposed face skin that can mislead the results. In a similar work, Lin et al. [20] use skin to classify porn image and non porn image. For classification they have used SVM. This solution starts with skin detection using skin-weight-map in RGB and HSV color space. Distance between these skin regions are calculated and grouped together using k-mean algorithm and then they found correlation in regions of skin and non-skin.

MPEG , a popular video format have some building features that are used by some researchers for processing. C. Jonson at el. [21] present a frame work to detect pornography in video. In their proposed method they estimated periodic motion using motion vector on MPEG-4 video. H. Sevimli et al. [22] present classification method that classify image into different level of adult content using MPEG-7 visual descriptor. This method, like many other early methods, works on skin regions but the novelty in this proposed method is the multi-level classification. This classification is based on the level of nudity i.e. sexual content, Nude, Topless, and swimming suit. In this work skin and face are used to extract four MPEG-7 feature that are Color Structure Descriptor (CSD), Color Layout Descriptor (CLD), Edge Histogram Descriptor (EHD) and Homogeneous Texture Descriptor (HTD). Using radial basis function (RBF) in multi-category SVM they classify images into the above mentioned four categories.

Sengamedu [23] present a method of pornography detection that detect body part in digital image and after identifying the body part it then classify the image accordingly. This system try to detect specific body part in image which itself is a difficult method.

Ulges and Stahl [24] a different techniques to detect special pornography i.e. Child pornography which is different and a bit difficult as compared to normal porn. To speed up the process and in a simple way, some authors have used proxy indicators like file name [25] and user activities [26]. These are easy to extract and fast in processing, but these methods have limited usability and do not analyze the content of the media.

3 Low-level Feature Analysis in Pornography Detection

In nudity and pornography detection most of the authors have used high-level features that are difficult to extract and process. Level of accuracy in such case are also low. Brightness and color intensity are some of the low-level features that are simple in nature, easy in extraction and involve less computation. Normally it is assumed that in content based applications, low-level features do not contain enough information to be used effectively. However in some video processing applications, low-level features can be helpful if used intelligently. One such application is pornography detection. Because of some interesting factors like cinematic principals and other production conventions, information extracted from different low-level features can be used to detect or identify such type of obscene in video clips.

3.1 Dataset

Here in this work, some of the basic features are tested to identify their relationship with nudity and pornography. Total of 200 video clips are used in this work that are selected randomly from dataset available at https://sites.google.com /site/pornographydatabase developed by S. Avila [27]. Out of these, 100 clips are pornographic while the other 100 are non-pornographic (normal clips). Total length this set is above 1400 minutes (88,546.6 seconds) that contain 2,538,550 frames. Summary of this dataset is given in Table 1.

Two main channels: Luma and Chroma are used to extract different features that represent brightness of intensity of different color channels. Such information are then analyzed to find their relation with pornography and nudity.

Table 1 Summary of Clips Used in Analysis.

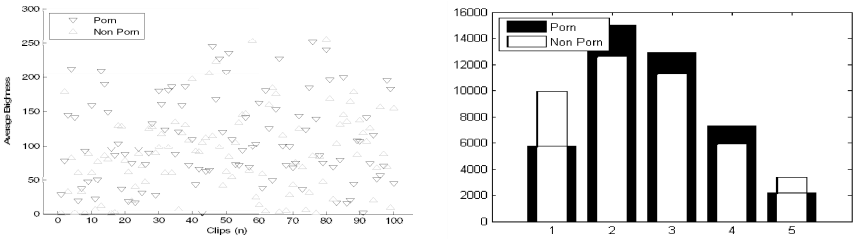
Category	Number of clips	Duration (Seconds)	Number of Frames
Porno	100	66,598.2	1,923,554
Non Porno	100	21,948.4	614,996
Total	200	88,546.6	2,538,550

3.2 Features Extraction

First feature in this analysis is the brightness in porno and non-porno scenes. RGB frames extracted from the clips are converted into gray-scale image using integer version of standard conversion equation shown in Equation (1).

$$\mathrm{Gray}_{(x,y)} = \frac{[\mathrm{R}_{(x,y)} * 299 + \mathrm{G}_{(x,y)} * 587 + \mathrm{B}_{(x,y)} * 114] + 500}{1000} \quad (1)$$

First feature is the average brightness that is simply the arithmetic mean of all the pixels in gray-scale frame and second feature is the distribution of brightness in five bins: 1 for the darkest and 5 for the brightest set of pixels. These features are shown in fig. 1. below.



(a). Average brightness (b). Brightness Distribution

Fig. 1 Brightness information

As shown in this figure brightness information do not provide a clear distinction between the two set of clips. Part b of the figure 1 shows that most of the pornographic videos have moderate brightness that lies at the mid while normal videos contain a border range of intensities. This is not significant to identify nudity in videos.

Colors are the next low-level information used in this analysis. Two color models RGB and HSV are used here. Average color intensity for red, green and blue channels, from RGB space and average intensity for Hue from HSV color space are extracted. Figure 3 below summarized the average intensity in RGB channels while average value of Hue is shown in figure 4. It can be seen from these figures that value in red for porno clips are slightly high than the average values for non-porno clips. Similarly some line can draw in hue channel for porn and non-porn clips. But like brightness, basic color information could not give significant information that can be used to distinguish pornographic and non-pornographic clips.

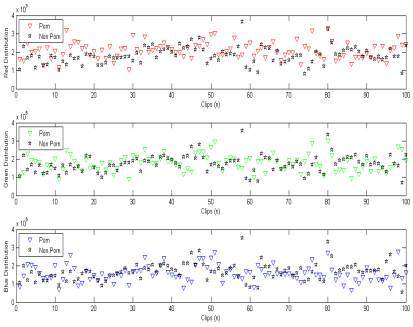


Fig. 2 Average Intensity of RGB Color

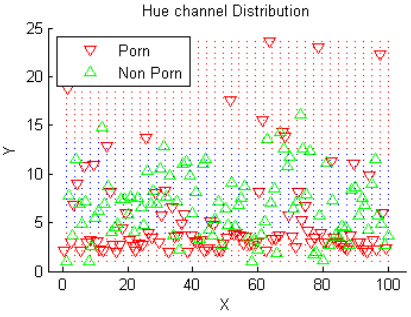
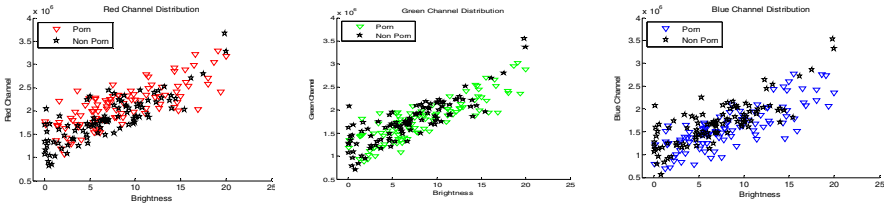


Fig. 3 Average Intensity in Hue channel

3.3 Results and Discussions

Main motivation of this analysis is to find a set of low-level features that are easy to extract and can be used to classify porno and non-porno video clips. But unfortunately initial results are not good enough to be used for classification. One possible reason is the variation in brightness which is almost similar in the two categories. Red, green, blue and hue are the four channels used in this analysis. Except hue, other three channels (RGB) are affected directly by the brightness. To overcome this issue features obtained from two type of channels i.e. color and brightness are combined and the results are surprisingly improved that can be seen in figure 5 below. In these figures, average value of red, green and blue and compared with the brightness level which gives much better patterns. In red and green channels, most of non-pornographic clips lie at lower-left corner while for pornographic video clips most of the values are at upper-right corner. Clear division in the combined effect indicates better classification could be done by using proper rations.



a: Red color channel b. Green color Channel c. Blue color channel

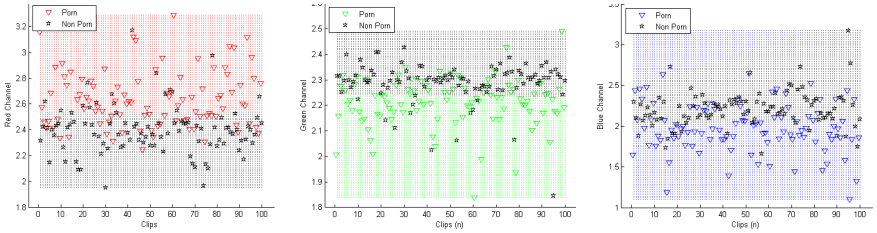
Fig. 4 RGB color channel vs. Brightness

In next step ratio between color and brightness for each frame is calculated using following equation 2.

$$R_c = \sum_{i=1,j=1}^{MN} C(i,j) \Big/ \sum_{i=1,j=1}^{MN} I(i,j) \quad (2)$$

Where Rc is the ratio in color channel c, $C(i,j)$ is the color intensity and $I(i,j)$ is the brightness of pixel at i,j, and $M \times N$ is size of image.

First ratio using above equation is for red channel that is shown in Figure 6. In this figure red triangle show the average values of all frame of a clip for pornographic while black star show the average of the ratios for non-pornographic clips. As it was expected, ratio of red vs. brightness is higher for porn clips than that of non-porn clips. One possible reason is the presence of skin and other low color tone in such videos. This is also confirmed by the figure where this is above 2.5 while for non-pornographic clips this value is less than 2.5. This division can easily be used for classification.



(a) Red / Brightness (b). Green / Brightness (c) Blue / Brightness

Fig. 5 Ratios of color channel with brightness

Similarly this ratio for green channel provides considerably good results. As clear from Figure 7, most of the pornographic clips have values less than 2.2 while non-pornographic clips this value lies between 2.3 and 2.5. Almost similar results are obtained from blue channel however division here is not as clear as the other two channels. Result of the blue channel is shown in Figure 8.

As in pornographic clips, some colors are more dominant in frame because of special effects and exposed skin. Hue channel in HSV color model in this situation give much better results. Following Figure 8 show the classification of porno and

non-porno clips based on Hue channels. As one can see that for pornographic clips (purple triangle) these values are less than 3.5 or above 13. For non-pornographic clips (black triangles) these values are between 3.5 and 13.

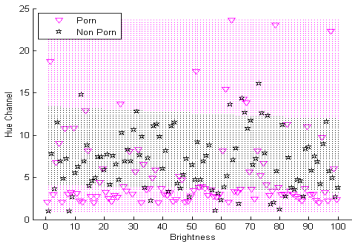


Fig. 6 Hue vs. Brightness

To test the significance of these features, Support Vector Machine (SVM) is used for testing. Gaussian Radial Basis Function with sigma value of 1 is used for testing. 50% (around) frames are sued for training and other 50% are used for testing. Results are summarized in table 2. Red and hue channel have higher accuracy (73% & 75%) and combination of all these channels have correction rate around 76%.

Table 2 Summary of Clips Used in Analysis.

Channel(s)	Positive productive value	Negative productive value	Sensitivity	Specificity	Correct Rate
Red	78.57%	70.24%	65.73%	81.85%	73.75%
Green	71.65%	73.11%	74.50%	70.16%	72.34%
Blue	69.80%	68.50%	68.13%	70.16%	69.14%
Hue	83.04%	59.17%	37.05%	92.34%	64.53%
RGB	76.33%	74.80%	74.50%	76.61%	75.55%
RGB + Hue	73.67%	79.82%	82.47%	70.16%	76.35%

4 Conclusion

It's a common understanding that low level features do not have sufficient information to be used for high-level decision in image and video analysis. However results in this analysis show that if properly used and compiled, these features can provide sufficient information can identify and classify pornography. These low-level features are easy to extract and process. Fast identification and classification system can be developed using such result that can work online for pornography detection or filtering system.

References

1. Hunt, L.: The Invention of pornography: obscenity and the origins of modernity, 1500-1800. Zone Books, New York (1993)
2. Wagner, P.: Erotica and the Enlightenment. Peter Lang Publishing Group, Frankfurt am Main (1991)
3. Döring, N.M.: The Internet's impact on sexuality: A critical review of 15 years of research. Computers in Human Behavior **25**, 1089–1101 (2009)
4. Cisco: Cisco Visual Networking Index: Forecast and Methodology, 2011–2016. Cisco Systems Inc. San Jose CA, USA, May 30, 2012
5. Strasburger, V.C.: Children, adolescents, and the media. Current Problems in Pediatric and Adolescent Health Care **34**, 54–113 (2004)
6. Is pornography bad for society and should it's sale and distribution be banned? http://www.debate.org
7. Wang, J., Li, J., Wiederhold, G., Firschein, O.: System for screening objectionable images. Computer Communications Journal **21**, 1355–1360 (1998)
8. Amir, A., Berg, M., Chang, S.-F., Hsu, W., Iyengar, G., Lin1, C.-Y., et al.: IBM Research TRECVID-2003 video retrieval system. In: Presented at the TREC Video Retrieval Eval, Gaithersburg, MD (2003)
9. Yan, R., Hauptmann, A.G.: A review of text and image retrieval approaches for broadcast news video. Information Retrieval **10**, 445–484 (2007)
10. Hauptmann, A., Baron, R.V., Chen, M.-Y., Christel, M., Duygulu, P., Huang, C., et al.: Informedia at TRECVID 2003: analyzing and searching broadcast news video. In: TREC Video Retrieval Eval, Gaithersburg, MD (2003)
11. Hauptmann, A., Chen, M.-Y., Christel, M., Huang, C., Lin, W.-H., Ng, T., et al.: Confounded expectations: Informedia at TRECVID 2004. In: TREC Video Retrieval Eval, Gaithersburg, MD (2004)
12. Foley, C., Gurrin, C., Jones, G., Lee, H., McGivney, S., O'Connor, N.E., et al.: TRECVID 2005 experiments at Dublin city university. In: TREC Video Retrieval Eval, ed. Gaithersburg, MD (2005)
13. Visser, R., Sebe, N., Bakker, E.: Object recognition for video retrieval. In: Presented at the Int. Conf. Image Video Retrieval, London, U.K (2002)
14. Hanjalic, A.: Extracting moods from pictures and sounds: towards truly personalized tv. IEEE Signal Processing Magazine **23**, 90–100 (2006)
15. Williams, C.E., Stevens, K.N.: Emotions and speech: Some acoustical correlates. Journal of the Acoustical Society of America **52**, 1238–1250 (1972)
16. Bradley, M.: Emotional memory: a dimensional analysis. In: Van Goozen, S.H., Van de Poll, N.E., Sergeant, J.A. (eds.) Emotions: Essays on Emotion Theory. Lawrence Erlbaum, Hillsdale, pp. 97–134 (1994)
17. Fleck, M.M., Forsyth, D.A., Bregler, C.: Finding naked people. In: 4th European Conference on Computer Vision-ECCV 1996, vol. II, pp. 593–602 (1996)
18. Duan, L., Cui, G., Gao, W., Zhang, H.: Adult image detection method base-on skin color model and support vector machine. In: Presented at the 5th Asian Conference on Computer Vision., Melbourne, Australia (2002)
19. Rowley, H.A., Jing, Y., Baluja, S.: Large scale image-based adult content filtering. In: International Conference on Computer Vision Theory and Applications (VISAPP), pp. 290–296 (2006)

20. Lin, Y.-C., Tseng, H.-W., Fuh, C.-S.: Pornography detection using support vector machine. In: Presented at the 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP 2003), Kinmen, ROC (2003)
21. Jansohn, C., Ulges, A., Breuel, T.M.: Detecting pornographic video content by combining image features with motion information. In: Presented at the 17th ACM International Conference on Multimedia MM 2009, Beijing, China (2009)
22. Sevimli, H., Esen, E., Ateş, T.K., Ozan, E.C., Tekin, M., Loğoğlu, K.B., et al.: Adult image content classification using global features and skin region detection. Computer and Information Sciences. Lecture Notes in Electrical Engineering, vol. 62, pp. 253–258 (2010)
23. Sengamedu, S.H.: Part-based pornography detection (2008)
24. Ulges, A., Stahl, A.: Automatic detection of child pornography using color visual words. In: 2011 IEEE International Conference on Presented at the In Multimedia and Expo (ICME) (2011)
25. Latapy, M., Magnien, C., Valadon, G.: First Report on Database Specification and Access including Content Rating and Fake Detection System (2008)
26. Latapy, M., Magnien, C., Tarissan, F., Valadon, G.: Content Rating and Fake Detection System (2009)
27. Avila, S., Valle, E., Araújo, A.d.A.: Pornography Database. C. CAPES, FAPEMIG, FAPESP (eds.) (2013)

Part XI
Potpourri

Sociology Study Using Email Data and Social Network Analysis

Wajid Rafiq, Shoab Ahmed Khan and Muhammad Sohail

Abstract Nowadays data mining and social network analysis techniques are broadly being used to study social structure of the underlying community. Internet has crept into our lives because of more and more dependency on online resources. We can precisely identify how society behaves from their presence in the cyber space. People are leaving so much foot prints on the cyber space that their online social structure can be extracted. In this research we have gathered email data of 23 graduate level students and applied social network analysis techniques. Social perspective of all the students has been extracted and threshold is applied on upper and lower bounds of data. We have extracted roles among people in the social setting and identified that how people behave while fulfilling those roles. We have also gathered survey data for those 23 students and extracted the social perspective in the same way as done for email data. We have performed the survey to extract what role a person is performing and related the social roles with the extracted social perspectives. At the end we have performed the validation process by comparing our results with the survey data. Our results showed strong correlation between emails data and survey data. Social structure of the students in the cyber space and in the real life was nearly same. To the best of our knowledge this is the first work in which both emails data and actual data in the form of survey for validation purposes is being used.

Keywords Emails social network · Social network analysis · Sociology · Social roles

1 Introduction

There exists a vast social life over the internet and the research in this field has become easy due to the availability of the interaction data available on social networking websites. Sociology study is the field in which we examine the

W. Rafiq(✉) · S.A. Khan · M. Sohail
Department of Computer Engineering, College of EME,
National University of Science and Technology (NUST), Islamabad, Pakistan
e-mail: rafiqwajid@gmail.com, msf321@gmail.com, shoabak@ceme.nust.edu.pk

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_91

association between people interacting on duty or in sport, in war or in love and in trade or worship. It is the social relationship that peoples' interest can find expressions and their desires become realized [1]. Due to the recent developments in the internet technology, communication technology and network theory, online networks are widely being used to study the social structure among people. Online social media websites are also playing a vital role in the online social interaction among users. Online social networking websites has a significant impact in our lives. After so much intervention of social networking websites it is interesting to study human behaviors during interaction with each other using these websites. It is more likely that people will interact and socialize with familiar people and objects [2]. Every individual owns a personal social network, people in this network and their connection mode portrays some characteristics about the network and has effect on thinking and behavior of people [3]. Sociology of cyber world is relatively new field, but it has a great potential to be explored. Social network analysis makes use of graph theory, statistics and probability theory, mathematical foundation and algebraic models to quantify the interaction taking place. Sociology study of cyber world makes use of structural variables to measure the graph relational ties such as friendship, business and trade.

This paper presents a focused sociology study to prove that emails social networks are actually correlated with the sociology of humans by using social network analysis approaches that incorporates refined statistical methods for exploring relational information. Our goal is to examine social perspective of students using emails data and by using survey data. We will extract social structure of students using email and survey data. After getting the social perspective of the students we will identify students that are playing some key roles, by applying threshold technique. Our results will prove that email social network of individuals is precisely the same as their actual social network.

2 Related Work

In the last decade the work on community structure analysis has gathered a lot of attention. A lot of work has been performed in the field of social structure analysis using social networks. In this research [4] authors demonstrated the cultural differences on Facebook by using association among exemplary participants of peculiar and collectivistic cultures, authors have also performed multilevel comparison on the basis of national level indices. In [5] authors validated that neighborhood based relations are dependent on personal and family characteristics for example income, education ethnicity, work status and time living on a particular address. In another work [6] authors highlighted the relevance of the science of groups its background generation and applications of the longitudinal and cross-sectional network statistics which are related to group researchers. Authors gave a background and general introduction to the new researchers in the field of group science. In [7] a model using call detail record has been proposed for human behavior and social network analysis. Authors have used statistical methods and probability to evaluate the relationships,

social groups and patterns of communication to detect change in human behavior. In [8] ninety two students took part in a survey in which they filled a questionnaire every day for a week recording daily time usage of the social network website Facebook and answering checklist of activities. Facebook was used by the individuals with whom they had offline relationship. Effects of Facebook usage for identity development and improving age-groups relationships were discussed. In [9] authors discussed the groups, role of persons in the groups evolution of groups and groups structure, they have also discussed the form of the linkage that motivate other persons to join the group. In [10] authors demonstrated that individuals are likely to share interests that talk each other by using applications of instant messaging, the web searches of those individuals were also same their relationship gets stronger as they spend more time in chatting., it is more likely that they share personal characteristics like age, place and are more possibly to be opposite in gender. Such results were also extracted for the persons who don't chat with each other but have a friend in common. In [11] a study that involved 190 employees comprising 38 work groups relating to 5 different organizations given indication that social networks affects group and individual's performance. Positive and negative effects of centrality on the job performance were discussed. Latent social network for eLearning system data was established in [12]. The social network discussed in this research paper consisted of students having similar contacts and similar contact circle. Data clustering was applied to the groups, the results showed that latent ties existed between group members.

3 Methodology

The purpose of this work is to analyze the social structure and characteristics of individuals in the social setting using email data and to study email social networks. Research was divided into two phases on the basis of data extraction process. First of all we examined and analyzed the community under consideration and selected the people for data extraction process. In phase-1 we gathered and analyzed email data and in the second phase survey data was collected and analyzed.

3.1 Phase-I

In the first phase we extracted the cyber space data from the participants specifically using email data for social interaction because email is a communication medium which is mostly used for well-defined purposes as oppose to other social networking websites that are mainly used for entertainment purposes. For the social network analysis NodeXL a template of Microsoft Excel was used. The simplest method to convert email lists into network interactions is the use of NodeXL 'import from Email Network' feature. When the windows search utility indexed files, emails can be imported directly by using NodeXL. In our example we have filtered emails from a specific date we included emails from September 2014 to February 2015.

NodeXL imports data in three columns, one column for sender, second column for receiver and third column for edge weight. Edge weight is the frequency of interaction between a specific sender and receiver. It was a huge amount of data so data was passed through preparation steps, first of all unwanted nodes were removed, links that were other than the population under consideration were eliminated. Sometimes people use different email addresses to communicate so their link came from two separate nodes, to eliminate this we merged two different emails addresses form same person into one. We have also eliminated self-loops in order to prepare the data.

Now we were only left with the sender and receiver nodes that were specific to the concerned population. We only considered interactions which had a specific weight. Sometimes people send emails for any information purpose so this is only one way interaction. We considered only those interactions that were reciprocal and occurred more than or equal to five times. For example we included a connection from 'A' to 'B' if node 'A' contacted node 'B' having edge weight of minimum five and node 'B' contacted node 'A' more than four times. Data was analyzed by calculating graph metrics in which graph measures were calculated for every node. Graph measures like degree, betweenness centrality, and closeness centrality were calculated, and a circular layout was used to plot the graph.

3.2 Phase-II

In the next phase we gathered actual data from the participants by preparing a questionnaire for the individuals. Individuals were asked to complete an adjacency matrix for all the other persons in the network. We asked the individuals to rate every other person's social perspective ranging from 0 to 5. We asked them to rate every other person in the class as to how close a friend he or she was. The question was to rate in the space that best describes your relationship with every other person on the list. Names of every person were included in the list, 0 being "don't have a connection" 5 being a well-defined connection. The values between 0 and 5 were in the increasing order of the sociability. In Table 1 social rating of five students is shown. The questionnaire was filled up by all the students, the data was then plotted and analyzed. After filling of the adjacency matrix another questionnaire was asked from the students which comprises of the following questions.

Table 1 Adjacency matrix for computing social perspective

	P1	**P2**	**P3**	**P4**	**P5**
P1	5	3	5	4	1
P2	5	5	4	5	2
P3	5	4	5	4	1
P4	5	3	5	5	1
P5	4	3	5	2	5

- To whom did they most often turn for advice and information?
- With whom did they most often ask for the study related help?
- Whom among their class mates they found most social?

With response to those questions names of four persons was requested.

4 Results

4.1 Email Data Analysis Results

Dataset after all the preparation steps was introduced to the social network analysis tool. Degree centrality measure was used for emails data interaction ties. We have considered only those interactions which took place at least five times and were reciprocal. All other ties which had edge weight less than five were ignored. We have calculated social perspective of every student based on the email interaction with other class members. We have used the following formula to weight social perspective of all the students, the result are shown in Table 2.

$$\frac{1}{n}\sum_{1}^{n} xi \tag{1}$$

Here:

x=degree of a node

Two thresholds were set that are as follows:

UTH: social perspective> =75%

LTH social perspective<=25%

Table 2 Social perspective calculated using email data

Person	Social Perspective	Person	Social Perspective
P1	57	P13	39
P2	30	P14	23
P3	22	P15	39
P4	57	P16	48
P5	96	P17	52
P6	65	P18	43
P7	91	P19	43
P8	61	P20	96
P9	57	P21	39
P10	23	P22	17
P11	57	P23	46
P12	91		

The number of students that were above the UTH range were four and the number of students laying below the LTH range were also four. The students that were below the LTH range were reserve students because of their less contact with other students.

4.2 Results of Survey Data

We used the following mathematical factor to calculate the social perspective of all the individuals.

$$\frac{1}{n}\sum_{1}^{n} xi \tag{2}$$

Social perspective of all the students was analyzed using bar graph. The thresholds that were assigned for cyber data analysis were also selected here and we got the same results.

UTH: social perspective> =75%
LTH social perspective<=25%
Here:
x=social rating of a specific individual
UTH is the upper threshold
LTH is the lower threshold

There were four students that were laying above 75% range and four students that were laying below 25% range as shown in Table 3. The second questionnaire was asked about four persons that were above UTH range. The second questionnaire resulted in determining of two students, P20, P7, P22 and P5. P20 was rated by the class as the most helping student and P5 was rated as the most sociable persons in the class. The other two persons, P7 and P22 were the class representatives.

5 Analysis of Results

Survey results show that two persons over UTH range were class representative (CR) of female students (P7) and class representative of male (P22) students.

P20 was most sociable person he had good contacts with other people in the class, he was also working as Research Assistant in the university so he used to help other students regarding administrative matters. The last student was most helping student in the class, he used to help all the students regarding their studies and he was also active in arranging parties and trips. Table 3 shows the social perspective of students using survey data technique.

Table 3 Social perspective calaulated using survey data

Person	Social Perspective	Person	Social Perspective
P1	52	P13	43
P2	26	P14	17
P3	25	P15	35
P4	52	P16	52
P5	87	P17	48
P6	61	P18	39
P7	96	P19	35
P8	70	P20	91
P9	61	P21	57
P10	18	P22	22
P11	61	P23	61
P12	78		

Individuals laying beneath the 25% threshold were P2, P3, P14 and P22. They were not sociable and were not interacting with other students. That's why there social perspectives were laying in the lower ranges. Table 4 contains the accuracy of the results after comparing both techniques.

Table 4 Accuracy of the results

Person	Accuracy %	Person	Accuracy %
P1	92	P13	91
P2	85	P14	74
P3	87	P15	89
P4	92	P16	92
P5	91	P17	92
P6	94	P18	90
P7	95	P19	81
P8	87	P20	95
P9	93	P21	69
P10	78	P22	77
P11	93	P23	75
P12	85		

6 Validation of Results

Survey data was used as the bench mark and cyber data was compared with this data. Social perspective of every student was extracted and plotted on the graph. We got overall accuracy of the social perspective of 86% with an error rate of 8.8%. We have got the social roles accuracy of 87% with an error rate of 8.5%.

7 Conclusion

People interconnecting in the cyberspace show the same behavior as they show offline. We have used email data to prove this phenomenon. In this research social interactions are modeled in a measureable form. We have gathered the cyber world and real data of a specific community and extracted their social structure. We have assumed that the survey data is the actual real world data that shows the actual relationship among that community. Degree centrality measure was used to get the social perspective of every person in the email social network. Average social perspective of every person is calculated and then it is compared with the real data. Our results show that the interactions in the cyber world and in real world are correlated. Specifically email social world and actual social world are actually the same. If we have email interaction data we can tell about the social structure about the people. This work also shows that there are people with specific roles in the society, in our case class representatives of girl students and boy students, most sociable persons and less interactive persons.

8 Future Work

Email data can be used to classify individual's behavior, interacting online. We can collect behavioral data and categorize it using classification algorithms. A future work can study that how there is difference in interaction behavior between people that are working in different locations, their specialization level, their profession and gender. The results of the underlying study can be used to design a learning management system or digital library having features of social network. This work can be extended to classify terrorists' networks and focusing on the most interactive persons we can easily get to the leaders in the network. This work can also be extended to enhance the marketing of the products, people with the same interaction pattern will behave in the same way so people with the same behaviors, likes and dislikes can be used to predict the likings of other people.

References

1. Blau, P.M.: Exchange and power in social life. Transaction Publishers (1964)
2. Pang, Y., Du, S.: Research on the selecting mechanism of individual interactive object for online social network. In: 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 820–827. IEEE, August 2014
3. McCarty, C., Molina, J.L.: Personal networks: Research and Applications, pp. 1–165. University of Florida, Florida
4. Na, J., Kosinski, M., Stillwell, D.J.: When a New Tool Is Introduced in Different Cultural Contexts Individualism-Collectivism and Social Network on Facebook. Journal of Cross-Cultural Psychology **46**(3), 355–370 (2015)

5. Van den Berg, P., Timmermans, H.: A multilevel path analysis of social networks and social interaction in the neighbourhood. REGION **2**(1), 55–66 (2015)
6. Wölfer, R., Faber, N.S., Hewstone, M.: Social network analysis in the science of groups: Cross-sectional and longitudinal applications for studying intra-and intergroup behavior. Group Dynamics: Theory, Research, and Practice **19**(1), 45 (2015)
7. Zhang, H., Dantu, R., Cangussu, J.W.: Socioscope: Human relationship and behavior analysis in social networks. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans **41**(6), 1122–1143 (2011)
8. Pempek, T.A., Yermolayeva, Y.A., Calvert, S.L.: College students' social networking experiences on Facebook. Journal of Applied Developmental Psychology **30**(3), 227–238 (2009)
9. Laine, M.S.S., Ercal, G., Luo, B.: User groups in social networks: an experimental study on Youtube. In: 2011 44th Hawaii International Conference on System Sciences (HICSS), pp. 1–10. IEEE, January 2011
10. Singla, P., Richardson, M.: Yes, there is a correlation:-from social networks to personal behavior on the web. In: Proceedings of the 17th International Conference on World Wide Web, pp. 655–664. ACM, April 2008
11. Sparrowe, R.T., Liden, R.C., Wayne, S.J., Kraimer, M.L.: Social networks and the performance of individuals and groups. Academy of Management Journal **44**(2), 316–325 (2001)
12. Dráždilová, P., Slaninová, K., Martinovič, J., Obadi, G., Snášel, V.: Creation of students' activities from learning management system and their analysis. In: International Conference on Computational Aspects of Social Networks, CASON 2009, pp. 155–160. IEEE, June 2009

Evaluation of Usability Heuristics for Transactional Web Sites: A Comparative Study

Freddy Paz, Freddy A. Paz and José Antonio Pow-Sang

Abstract Nielsen's usability heuristics are the most commonly used assessment tool to perform a heuristic evaluation. However, when they are employed to evaluate the usability of transactional Web sites, they fail to cover all aspects of usability that are currently present in this kind of software. For this reason, we have developed a new proposal that is capable of providing more accurate results in this context. Our approach includes fifteen usability heuristics for the design of usable graphical user interfaces. This paper presents a comparative study between the classical Nielsen's proposal and the new set of heuristics for transactional Web sites. The experimental case study was conducted following the Method Adoption Model which establishes the analysis of three dimensions: perceived ease of use, perceived usefulness and intention to use. For this purpose, forty-six undergraduate students were asked to perform a heuristic evaluation, in which both proposals were employed. The results showed that the new heuristics are easier to use than the traditional approach. This study provides the validation of an effective tool that can be used easily to perform heuristic evaluations in the context of transactional Web applications.

Keywords Heuristic evaluation · Usability heuristics · Comparative study · Experimental evaluation · Perception model

1 Introduction

The continuous growth of Electronic Commerce has led to the development of more sophisticated systems. Nowadays, Web sites are developed with highly complex

F. Paz(✉) · J.A. Pow-Sang
Pontificia Universidad Católica del Perú, San Miguel, Lima 32, Lima, Peru
e-mail: fpaz@pucp.pe, japowsang@pucp.edu.pe

F.A. Paz
Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Peru
e-mail: freddypazsifuentes@yahoo.es

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_92

components to support a large number of functional requirements. However, although the software products have changed their nature, specialists still continue using traditional tools to evaluate the usability. Software products are increasingly incorporating new features that are not considered by conventional assessment tools [14]. The literature states that these instruments have become unsuitable to evaluate the usability of new categories of software applications that are currently emerging [9].

Nielsen's usability heuristics are the most commonly used approach in heuristic evaluations. Nevertheless, they provide inaccurate results when they are used to evaluate nontraditional types of software such as: transactional Web sites, mobile-based applications, videogames, virtual worlds, applications for interactive TV and grid computing applications [12]. These new generation of systems demands more accurate inspections, and assessment tools that allow evaluators to identify, besides generic usability problems, issues from the application domain.

In a previous work [10], we developed a new set of usability heuristics for transactional Web applications. This paper presents a comparative study between the traditional Nielsen's approach and our new proposal. The results of both are compared and discussed.

2 Background

2.1 Usability and Heuristic Evaluation

In Software Engineering, the term usability is related to the ease of use of a software product. The ISO 9126-1 standard provides a specialized definition where usability is considered as "the capability of a software product to be understood, liked, used and attractive to the user, when it is used under specified conditions" [3]. This definition emphasizes not only the relevance of an aesthetic and intuitive graphical interface, but also the potential of a specific software to meet user's expectations.

The importance of usability in the context of software development has led to the emergence of several evaluation methods. The purpose of these techniques is to determine systematically the level of usability of specific software interfaces [2]. Usually, these methods are employed during all phases of the software development process to ensure the design of a usable product that can meet high-quality standards.

Usability is one of the most important quality attributes of a software product. In Web domain, where there are several alternatives to a Web site, the adoption of a software can be determined by its usability. If a Web site is difficult to use or fails to clearly state what it is offering, then users leave [7]. For this reason, usability is an aspect that should always be considered in the software design.

Heuristic evaluation is a well-known inspection technique that is widely used by usability specialists. It was developed by Nielsen [5], as an alternative assessment to the time-consuming user testing. In contrast to methods which involve the participation of extensive amounts of end users, in a heuristic evaluation, only from 3

to 5 specialists are required. Heuristic evaluation involves having usability specialists judge whether each dialogue element follows established usability principles, called "heuristics" [6]. In the context of software applications, the evaluation process includes a complete review of all the graphical user interfaces. The assessment is conducted by professionals in HCI in order to determine if all principles are followed. In case of infringement, the issue is classified as a usability problem. There are many protocols to conduct a heuristic evaluation. However, an own proposal was employed for this study [11].

3 A Comparative Study

The purpose of this study is to compare the perceptions about two different approaches of usability heuristics. In this section, both proposals are presented.

The principles proposed by Nielsen are the most commonly used tool in heuristic evaluations. This proposal is considered as a classical assessment instrument [12]. However, when these heuristics are used to evaluate new categories of software that are available nowadays, the results are inaccurate [8]. The main reason is that existing software products are embedded of sophisticated designs, excessive functionality and real-time processing [9]. Software applications, especially Web sites, are in continuous development. They incorporate increasingly new features that were not considered when the classical usability evaluation tools were established.

In a previous work [13], we identified some important aspects of usability related to culture, design, transaction, and functionality that, besides being relevant in the context of transactional Web applications, are not considered by Nielsen's heuristics. These principles fail to deal with the new usability features that the current software products contain. Therefore, we have developed a new set of heuristics in order to provide a tool that is capable of evaluating the level of usability of a transactional Web site, considering all features that are currently embedded in this specific kind of software.

3.1 Nielsen's Usability Heuristics

The Nielsen's proposal is considered one of the best-known sources to perform a heuristic evaluation. These principles describe ten general rules to evaluate the usability of a software product. Nielsen's usability heuristics are [10]: (N1) visibility of system status, (N2) match between system and the real world, (N3) user control and freedom, (N4) consistency and standards, (N5) error prevention, (N6) recognition rather than recall, (N7) flexibility and efficiency of use, (N8) aesthetic and minimalist design, (N9) help users recognize, diagnose, and recover from errors, and (N10) help and documentation.

3.2 New Usability Heuristics for Transactional Web Sites

This new set of usability heuristics was developed because of the need for a more accurate tool in the field of transactional Web applications. Despite Nielsen's heuristics are still valid in this domain, they do not succeed in covering all aspects of usability that are present in this kind of software [13]. For this reason, we employed the methodology proposed by Rusu et al. [14] to develop specific heuristics for this particular domain. A preliminary version was presented in a previous study [10]. In this work, we detail the final version of our approach after some experimental case studies [9]. The usability heuristics for transactional Web sites are:

F1. Visibility and clarity of the system elements: According to the nature of the system, some components will be more important than others. Therefore, the most relevant elements of the interface must be clearly visible and understandable.

F2. Visibility of the system status: The system must always keep users informed about the current state of the software. When a complex transaction is required, users must be notified via messages within reasonable wait times.

F3. Match between system and user's cultural aspects: Users should not feel forced to use a software whose mechanisms substantially differs from their own customs, beliefs or ways of thinking. The interface design must be related to the cultural profile of end users.

F4. Feedback of transaction: The system must keep users informed about the final status of an operation. Users must be notified about the success or failure of all transactions that are performed when the application is used.

F5. Alignment to Web design standards: The system must follow Web platform conventions. The graphical interface should be aligned to standardized guidelines, patterns and design structures that are widely known by the users, because of their extended use over the time.

F6. Consistency of design: All sections of the system must follow the same style and design pattern. The whole structure must be consistent and well organized. Information should appear in a natural and logical order.

F7. Standard iconography: The interface design must be implemented by the use of standardized icons whose meaning is already part of the user's conceptual model. Icons must succeed in communicating their intended purpose.

F8. Aesthetic and minimalist design: The user interface must not only be attractive but also it must contain units of information which are certainly relevant to users. Each extra unit of information will be competing in importance with other elements that are indeed essential.

F9. Prevention, recognition and error prevention: The system must avoid the occurrence of errors. The software must implement a careful design that prevents users

from performing actions that lead to errors. However, once an error occurs, the system must help users to recognize and quickly overcome this scenario by displaying clear messages with suitable instructions to solve the problem.

F10. Appropriate flexibility and efficiency of use: The software interface must provide expert users with accelerators to expedite the use of the system, without affecting the normal workflow of novice users. The interface must allow both, experienced and inexperienced users, to effectively achieve their goals.

F11. Help and documentation: The system must provide support to help users to perform concrete actions in the application that are required for the achievement of the user's goals. The help and documentation of the system must contain a list of explicit instructions to be carried out, focused on the user's tasks.

F12. Reliability and quickness of transactions: The system must guarantee that all kind of transaction will be successfully completed within the expected time under specific operational conditions. In case of error, the system must be capable of solving the issue and reverting all operations to a safe state.

F13. Correct and expected functionality: The elements of the system must be correctly implemented, and provide the functionality that is expected from them. The result of an action should be related to the purpose that is established in the interface design.

Table 1 Mapping between Nielsen's Proposal and the New Approach for Transactional Web Sites

Nielsen's Usability Heuristics		**Usability Heuristics for Transactional Web Sites**	
Id	***Name***	***Id***	***Name***
N1	Visibility of system status	F2	Visibility of the system status
N2	Match between system and the real world	F1	Visibility and clarity of the system elements
		F3	Match between system and user's cultural aspects
N3	User control and freedom	F15	User control and freedom
N4	Consistency and standards	F7	Standardized symbology
		F5	Alignment to Web design standards
		F6	Consistency of design
N5	Error prevention	F9	Prevention, recognition and error prevention
N9	Help users recognize, diagnose, and recover from errors		
N6	Recognition rather than recall	F14	Recognition rather than recall
N7	Flexibility and efficiency of use	F10	Appropriate flexibility and efficiency of use
N8	Aesthetic and minimalist design	F8	Aesthetic and minimalist design
N10	Help and documentation	F11	Help and documentation
		F4	Feedback of transaction
		F12	Reliability and quickness of transactions
		F13	Correct and expected functionality

F14. Recognition rather than recall: Users should not be forced to remember information from a previous state to the current part of the dialogue. The system must minimize the user's memory load by providing a highly intuitive graphical interface.

F15. User control and freedom: Users can execute certain functions of the system by mistake. Therefore, the system must provide mechanisms to allow users to exit from undesired states, reverting their actions. A mapping between the classical Nielsen's proposal and the new set of heuristics is presented in Table 1. This mapping shows the gaps that are covered by our new approach.

4 An Evaluation Model and Variables

The Method Adoption Model (MAM) provides a suitable basis to predict the success of a software design tool. According to this model, a successful proposal is one that is adopted in practice by the specialists. However, this acceptance, given by the community, can only be possible if the approach is perceived as easy to use and useful.

The MAM proposes the study of three psychological aspects whose relations are presented in Figure 1. It is based on the Technology Acceptance Model (TAM) [1], a well-known and thoroughly validated approach for evaluating information technologies. Moreover, it is the core of the Method Evaluation Model (MEM) [4], a more complex evaluation framework for software development tools that are currently in use.

Although usability heuristics are used as an evaluation tool, they can be considered as design instruments. These guidelines are used as important components for the development and design of usable graphical interfaces [15]. Therefore, the MAM is a suitable evaluation model to predict the success of these principles. The intention to use the heuristics can be considered as success, that will become visible through the adoption of the heuristics in practice. In this study, we have focused on the

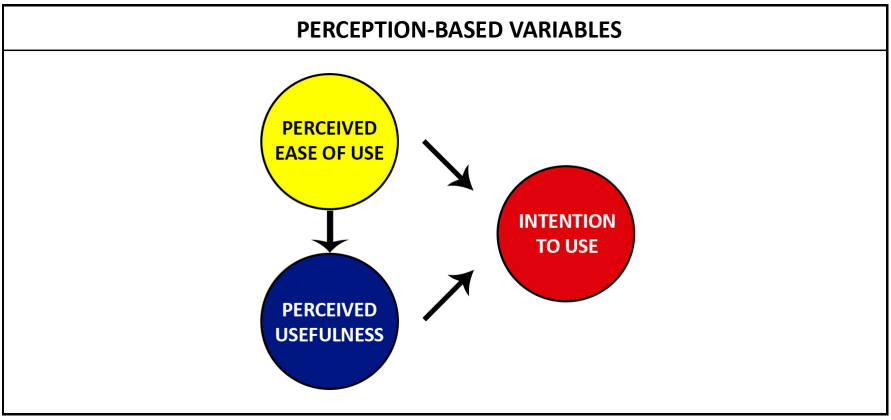


Fig. 1 Method Adoption Model (MAM)

perception-based variables of the MAM. However, the definition of these constructs was adapted to the context of usability principles. The variables that were considered are:

- *Perceived ease of use* (PEU): The extent to which an evaluator believes that using a particular set of usability heuristics would be free of effort.
- *Perceived usefulness* (PU): The extent to which an evaluator believes that a particular set of usability heuristics will achieve its intended objectives.
- *Intention to use* (IU): The extent to which a reviewer intends to use a particular set of usability heuristics in the future. This construct is an intention-based variable for predicting the adoption in practice of the heuristics.

5 Research Design

5.1 Participants

The participants in the study were forty-six undergraduate students in the final year of Engineering in Computer Science at the National University "Pedro Ruiz Gallo" (UNPRG). They were randomly chosen from two different sections of a same course called *Software Quality*. The students had similar backgrounds given that all had previously attended the same courses of the program.

All students voluntarily agreed to participate in our study without expecting any kind of compensation for their participation. Before conducting the experiment, students were informed that their answers would not affect their grades in the course. The experiment was conducted in October 2015.

As part of the course activities, students had to assess the quality attributes of a software product. This fact allows the instruction of heuristic evaluations as a technique to measure the level of usability of software graphical interfaces. Students have little or no previous experience in this topic. In order to perform the study case; they had to be previously trained on the concepts, procedures and tools that are required.

Although students were from two different sections, they were not mixed. In order to perform a comparative study, the heuristic evaluation was performed with a different evaluation tool in each section. The traditional Nielsen's proposal was used in Section 1 by twenty-one students. In the same way, the new set of heuristics for transactional Web sites was employed in Section 2 by twenty-five students.

5.2 Study Design

Our empirical study was focused on the analysis of the students' perceptions about the use of two different sets of usability heuristics. For this purpose, we have considered

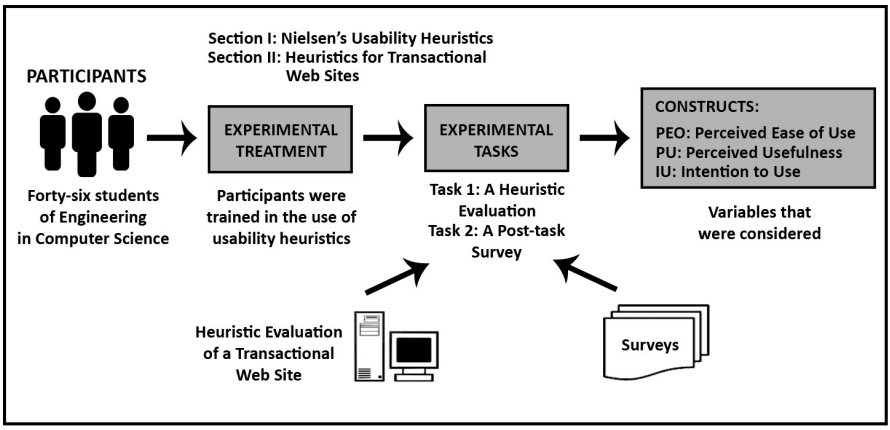


Fig. 2 Experimental Design

the experimental design that is illustrated in Figure 2. First, all participants were trained in the main concepts of usability and heuristic evaluations. Subsequently, they had to examine a transaction Web site using the set of heuristics that was assigned to their section. For this activity, *Booking.com*, an online Web site to make hotel reservations, was selected. The students inspected the graphical user interface of the system for about two hours. As a result of the evaluation process, each participant reported a list of usability problems related to the heuristics that was infringed by the interface design.

Finally, a post-task survey was employed to identify the students' perceptions about the use of heuristics with regard to PEO (*perceived ease of use*), PU (*perceived usefulness*) and IU (*intention to use*). The items of the survey were formulated using a 5-point Likert scale, where 1 was referred to a negative perception of the construct, and 5 to a positive opinion.

6 Data Analysis and Results

The survey results were analyzed to determine the PEO and PU of both proposals. The results were averaged in order to obtain a global score by heuristic. In Table 2, the results of the Nielsen's usability heuristics are presented. Similarly, the results of our new approach for transactional Web sites is presented in Table 3. The results show that Nielsen's heuristics are considered as useful. All the usability heuristics obtained a score above 3.0 (the neutral value) in *perceived usefulness* (PU). However, in the construct related to the *perceived ease of use* (PEO), some heuristics did not get a good score, and are classified as difficult to use. This result can be explained by the fact that the participants are novice reviewers. Additionally, it is the first time they are using an assessment tool as the usability heuristics.

Table 2 PEO and PU of Nielsen's Usability Heuristics

Nielsen's Usability Heuristics	PEO	PU
N1. Visibility of system status	3.05	4.05
N2. Match between system and the real world	3.00	3.81
N3. User control and freedom	3.19	3.86
N4. Consistency and standards	2.81	3.95
N5. Error prevention	3.05	3.71
N6. Recognition rather than recall	2.62	3.67
N7. Flexibility and efficiency of use	2.88	3.24
N8. Aesthetic and minimalist design	3.57	3.95
N9. Help users recognize, diagnose and recover from errors	2.79	3.67
N10. Help and documentation	3.19	3.52

Table 3 PEO and PU of the New Heuristics for Transactional Web Sites

Heuristics for Transactional Web Sites	PEO	PU
F1. Visibility and clarity of the system elements	4.04	4.08
F2. Visibility of the system status	3.68	3.88
F3. Match between system and user's cultural aspects	3.20	4.52
F4. Feedback of transaction	2.68	4.08
F5. Alignment to Web design standards	3.12	3.52
F6. Consistency of design	3.32	3.92
F7. Standard iconography	3.48	3.72
F8. Aesthetic and minimalist design	4.32	4.00
F9. Prevention, recognition and error recovery	3.20	3.56
F10. Appropriate flexibility and efficiency of use	2.80	3.08
F11. Help and documentation	4.28	3.28
F12. Reliability and quickness of transactions	2.88	3.76
F13. Correct and expected functionality	3.56	4.12
F14. Recognition rather than recall	2.72	2.72
F15. User control and freedom	3.32	3.40

The results of our approach for transactional Web sites are variable. There are both high and low scores. High scores can be considered as a positive aspect of our proposal. Low scores can be interpreted as an indicator that enhancements are required to increase the ease of use and usefulness.

A comparison of both proposals is presented in Table 4. Although there is an improvement in all aspects regarding the Nielsen's traditional heuristics, the differences were not highly remarkable. However, further studies are needed to generalize the results.

Table 4 Comparison between the Traditional Nielsen's Proposal and the New Set of Usability Heuristics

Approach	PEO	PU	IU
Nielsen's Usability Heuristics	3.02	3.74	3.62
Heuristics for Transactional Web Sites	3.37	3.71	3.64

7 Conclusions and Future Works

Heuristic evaluation is a widely used method to determine the usability of software products. This technique involves a group of specialists judging if all elements of a GUI follow specific guidelines called "heuristics". On the context of transactional Web applications, Nielsen' heuristics are the most common assessment tool to perform a heuristic evaluation. However, these principles fail to cover all aspects of usability that are currently present in this kind of software. Therefore, a new proposal was developed.

A new set of fifteen heuristics for transactional Web sites was proposed. In this study, we have performed a comparison between both approaches. The purpose was to determine how the heuristics are perceived by specialists that have little or none experience, in order to avoid any kind of preference. The experimental case study was conducted following the Method Adoption Model (MAM) which establishes the study of three psychological dimensions: *perceived ease of use* (PEU), *perceived usefulness* (PU) and *intention to use* (IU). Forty-six students from two different sections were asked to perform a heuristic evaluation. Both proposals were each assigned to a different section. After the evaluation, students were requested to fill a survey in which the three dimensions were evaluated.

The results showed that our new approach is easy to use and usefulness. Additionally, there are intentions to use these heuristics in the future. From these results, it is possible to identify an improvement regarding the traditional proposal. However, the differences are not highly remarkable. Therefore, it is still necessary to refine some heuristics that individually scored low. Given that these results are only valid in this environment; it would be also useful to repeat the experiment in other contexts. Furthermore, a checklist could be important to enhance the ease of use of these heuristics.

References

1. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly **13**(3), 319–340 (1989)
2. Fernandez, A., Insfran, E., Abrahão, S.: Usability evaluation methods for the web: A systematic mapping study. Information and Software Technology **53**(8), 789–817 (2011)
3. ISO: Software engineering – product quality – part 1: Quality model. ISO 9126-1:2001, International Organization for Standardization, Geneva, Switzerland (2001)
4. Moody, D.L.: Dealing with Complexity: A Practical Method for Representing Large Entity Relationship Models, 1st edn. Department of Information Systems, University of Melbourne, Melbourne, Australia (2001)
5. Nielsen, J.: Usability Engineering, 1st edn. Academic Press, San Diego (1993)
6. Nielsen, J.: Usability inspection methods. In: Conference Companion on Human Factors in Computing Systems, CHI 1994, pp. 413–414. ACM, New York (1994)
7. Nielsen, J.: Usability 101: Introduction to Usability (2012). https://www.nngroup.com/articles/usability-101-introduction-to-usability/ (accessed: November 24, 2015)

8. Otaiza, R., Rusu, C., Roncagliolo, S.: Evaluating the usability of transactional web sites. In: Third International Conference on Advances in Computer-Human Interactions, ACHI 2010, pp. 32–37, February 2010
9. Paz, F., Paz, F.A., Pow-Sang, J.A.: Experimental case study of new usability heuristics. In: Marcus, A. (ed.) Design, User Experience, and Usability: Design Discourse. Lecture Notes in Computer Science, vol. 9186, pp. 212–223. Springer International Publishing (2015)
10. Paz, F., Paz, F.A., Pow-Sang, J.A., Collantes, L.: Usability heuristics for transactional web sites. In: 2014 11th International Conference on Information Technology: New Generations (ITNG), pp. 627–628, April 2014
11. Paz, F., Paz, F., Villanueva, D., Pow-Sang, J.: Heuristic evaluation as a complement to usability testing: a case study in web domain. In: 2015 12th International Conference on Information Technology - New Generations (ITNG), pp. 546–551, April 2015
12. Paz, F., Pow-Sang, J.A.: Current trends in usability evaluation methods: A systematic review. In: 2014 7th International Conference on Advanced Software Engineering and Its Applications (ASEA), pp. 11–15, December 2014
13. Paz, F., Villanueva, D., Rusu, C., Roncagliolo, S., Pow-Sang, J.A.: Experimental evaluation of usability heuristics. In: 2013 Tenth International Conference on Information Technology: New Generations (ITNG), pp. 119–126, April 2013
14. Rusu, C., Roncagliolo, S., Rusu, V., Collazos, C.: A methodology to establish usability heuristics. In: 2011 4th International Conference on Advances in Computer-Human Interactions (ACHI), pp. 59–62 (2011)
15. Zapata, C.: Integration of usability and agile methodologies: A systematic review. In: Marcus, A. (ed.) Design, User Experience, and Usability: Design Discourse. Lecture Notes in Computer Science, vol. 9186, pp. 368–378. Springer International Publishing (2015)

Algorithm for Gaussian Integer Exponentiation

Aleksey Koval

Abstract In this paper we introduce a novel algorithm for Gaussian integer exponentiation that significantly improves its performance. We compare the performance of Gaussian integer exponentiation (using new and existing algorithms) to real integer exponentiation both analytically and experimentally. We demonstrate that the algorithms based on the Gaussian Integer exponentiation have significant advantages over the corresponding algorithms based on the real integer exponentiation. Moreover, we show that the new algorithm is significantly faster. Therefore, the new algorithm could speedup Public Key Discrete Logarithm Based cryptographic algorithms by about 40% with a possibility for further improvements.

Keywords Cryptography · Gaussian integers · Lucas sequences

1 Introduction

Gaussian integers can provide the alternative to real integers in Discrete Logarithm problem (DLP) based public key cryptography algorithms. In [1] the complexity of Gaussian Integer DLP is quantified and it is shown that a Gaussian integer cryptosystem modulo prime p is as secure as real integer cryptosystem modulo prime q, where q has twice as many bits as p (assuming Lucas sequence DLP is hard). In addition, we use the theorems presented in [1] to construct a novel exponentiation algorithm that uses Lucas sequences for Gaussian integer exponentiation. The theoretical analysis and experimental results show that Gaussian integer exponentiation modulo p is faster than real integer exponentiation modulo q. Furthermore, the Lucas sequence Gaussian exponentiation algorithm (the novel algorithm introduced in Section 4) is faster still.

A. Koval(✉)
Computer Science Department, New Jersey Institute of Technology, Newark, USA
e-mail: ak77@njit.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_93

2 Overview of Gaussian Integers and Lucas Sequences

A Gaussian integer is a complex number $a+bi$ where both a and b are integers: $Z[i]=\{a+bi : a,b \in \mathbb{Z}\}$. Gaussian integers, with ordinary addition and multiplication of complex numbers, form an integral domain, usually written as $Z[i]$. The **norm** of a Gaussian integer is the natural number defined as $|a+bi| = a^2+b^2$. We use vector notation for Gaussian integers: $A = a+bi$ is equivalent to $A=(a,b)$. We use the following definition of mod operation for Gaussian integers:

$$(a+bi) \mod p = (a \bmod p)+(b \bmod p)i \tag{2.1}$$

The order of $Z[i]$ modulo p is p^2-1.

It takes three multiplications to multiply two Gaussian integers:

Algorithm 2.1. Gaussian integer multiplication
Suppose $(a,b)(c,d)=(x,y)$. If $v_1=(a+b)(c+d)$; $v_2=ac$; $v_3=bd$, then $x=v_2-v_3$ and $y=v_1-v_2-v_3$.
It takes two multiplications to square a Gaussian integer:

Algorithm 2.2. Gaussian integer square

$$(a,b)^2 = \big((a+b)(a-b), ab+ab\big) \tag{2.2}$$

Lucas sequences are defined as sequences $U_k(P,Q)$ and $V_k(P,Q)$ ($P^2-4Q \neq 0$) by recurrence relations:

$$U_0=0; U_1=1; U_k(P,Q)=PU_{k-1}-QU_{n-2} \tag{2.3}$$

$$V_0=2;\ V_1=P;\ \ V_k(P,Q)=PV_{k-1}-QV_{n-2} \tag{2.4}$$

3 Gaussian Integer Multiplication vs. Real Integer Multiplication

There are numerous exponentiation algorithms and their variations. Most of them use integer multiplications and squaring as basic steps. Any such exponentiation algorithm can be used for Gaussian integer exponentiation. All we need to do is to replace real integer multiplications with Gaussian integer multiplications and real integer squaring with Gaussian integer squaring.

Since the size of the group of the exponentiation cyclic group of Gaussian integers is p^2-1, it is appropriate to compare Gaussian integer multiplication modulo p to real integer multiplication modulo q, where q is double the size of p. Suppose $n=\lceil \log_2 p \rceil$ (or in other words n is the number of binary bits of p). Let's select q to be the closest prime to p^2. Then the size of q in bits would be approximately $2n$ or $\lceil \log_2 q \rceil \approx 2n$.

With all efficient implementation of multiplication of large integers, the multiplication algorithm used varies with n. For small n the naïve multiplication method [2] with complexity of $\Theta(n^2)$ is most efficient. For larger n Karatsuba-Ofman [3] multiplication algorithm is appropriate with a running time of $\Theta(n^{1.585})$. For even larger n Toom-Cook 3-way (or Toom-3) [2] multiplication with a running time $\Theta(n^{1.465})$ is appropriate. As n increases further multiple levels of k-way Toom-Cook multiplication with a running time $\mathrm{O}\left(n^{\log(2k-1)/\log k}\right)$ [2] can be applied. For extremely large n algorithms based on Fast Fourier transforms (FFT), like Schönhage–Strassen algorithm [4]) or Fürer's algorithm [5] are more efficient. However, FFT algorithms are used for very large n because of overhead, far larger then the numbers used for public key cryptography. Therefore, we will not consider FFT algorithms in our subsequent discussion.

Suppose T_G is the time required for multiplication with reduction for Gaussian integers and T_r is the time required for multiplication with reduction for real integers. Since we excluded FFT multiplication algorithms from our discussion, we can say that

$$t_r(n) = \mathrm{O}(n^{\alpha}) . \tag{3.1}$$

where $t_r(n)$ is the multiplication time of two integers of size n.

After each multiplication or square we need to perform the reduction modulo prime. For Gaussian integers we will have to perform two modulo reductions, one for real and one for imaginary part using (2.1). For real integer we perform modulo reduction once, but the operands are twice the size. For reduction we can use modulo division or a more efficient Montgomery reduction *REDC*() operation [6]. Whether we use modulo division or Montgomery *REDC*() function, the speed of the reduction at each multiplication or square step can be expressed as

$$R(n) = \beta t_r(n) , \tag{3.2}$$

where $R(n)$ is the division time of an integer of size $2n$ by an integer (prime in our case) of size n and $\beta : 1 < \beta < 4$ is some constant. The value of β is heavily dependant on the implementation and varies among different platforms. With GMP, the speed of the modulo division is about 2 to 4 times slower than multiplication for moderately large integer sizes (section 15.2.3 of [7], [8]), while *REDC*() operation speed varies from about 1.2 multiplications to 2 multiplications. With GMP 6.0.0 the average β for *REDC*() is about 1.5 multiplications. Based on our assumptions, we can estimate the relationship between T_G and T_r **:**

$$\lim_{n\to\infty} \frac{T_G}{T_r} = \frac{3t_r(n) + 2R(n)}{t_r(2n) + R(2n)} = \frac{3 + 2\beta}{2^{\alpha}(1+\beta)} . \tag{3.3}$$

Integer squaring can be up to twice as fast as multiplication but in practice is much slower. The GMP manual [7] states that square is around 1.5 times faster than multiplication, if the library settings are optimized. Incidentally, the Gaussian integer squaring is also 1.5 times faster relatively to Gaussian integer multiplication on all platforms (refer to Algorithm 2.1 and Algorithm 2.2 in previous sections). Based on the assumption that the time for integer square is 2/3 times integer multiplication:

$$\lim_{n\to\infty}\frac{S_G}{S_r}=\frac{2t_r(n)+2R(n)}{.66t_r(2n)+R(2n)}=\frac{6(1+\beta)}{2^{\alpha}(2+3\beta)}. \tag{3.4}$$

where S_G is the time required for square with reduction for Gaussian integers and S_r is the time required for square with reduction for real integers.

Table 1 Summarized estimates of the ratio of the running time of Gaussian integer multiplication (square) over the running time of real integer multiplication (square), based on the formulas (3.3) and (3.4).

	Naïve α=2	Karatsuba α=1.585	Toom-3 α=1.465	Toom-4 α=1.4037
β=1	0.60(0.63)	0.80(0.83)	0.87(0.91)	0.91(0.94)
β=1.2	0.59(0.61)	0.79(0.82)	0.85(0.89)	0.89(0.93)
β=1.4	0.58(0.60)	0.77(0.81)	0.84(0.88)	0.88(0.91)
β=1.5	0.58(0.60)	0.77(0.80)	0.84(0.87)	0.87(0.91)
β=1.7	0.57(0.59)	0.76(0.79)	0.83(0.86)	0.86(0.90)
β=2	0.56(0.58)	0.75(0.78)	0.82(0.85)	0.85(0.88)
β=2.2	0.56(0.58)	0.74(0.77)	0.81(0.84)	0.84(0.87)
β=2.5	0.55(0.57)	0.74(0.76)	0.80(0.83)	0.84(0.86)
β=3	0.55(0.56)	0.73(0.75)	0.79(0.82)	0.82(0.85)
β=4	0.54(0.55)	0.71(0.73)	0.78(0.80)	0.81(0.83)

As we can see from Table 1 Gaussian integer exponentiation is faster on all platforms because the underlying multiplication and square operations (combined with reductions) are faster for Gaussian integers. The exact speedup would depend on platform, integer sizes, and exponentiation algorithm logic (i.e. number of multiplications and squares). Realistically, for the integer sizes used for Public Key cryptography (1000-4000 bits) we can expect speedup of around 20% with Gaussian integers. The rational for this is that the estimated speedup ratios for multiplications and squares for Montgomery reduction with $\beta=1.5$ and Karatsuba multiplication are 0.80 and 0.77 so regardless of the ratio of multiplications/squares in a particular exponentiation algorithm, the combined ratio would be under 0.80 (i.e. 20% speedup).

4 New Algorithm for Gaussian Integer Exponentiation

As shown in the previous section, Gaussian integer exponentiation is faster then real integer exponentiation when we replace real integers with Gaussian integers. However, we designed an even faster exponentiation algorithm for Gaussian integers. It is based on the relationship between Gaussian integers and Lucas sequences described in [1].

Algorithm 4.1. Lucas sequence Exponentiation of Gaussian integers (**LSEG**)

Inputs: (a,b) Gaussian integer

p – prime: $p \bmod 4 = 3$

n – exponent: $0 < n < p^2 - 1$

Output: Gaussian integer $(x,y) = (a,b)^n \bmod p$

1. $r := \left|(a,b)\right|^{\frac{p+1}{4}} \bmod p$
2. $(c,d) := r^{-1}(a,b) \bmod p$
3. $h := r^{n \bmod (p-1)} \bmod p$
4. $m := n \bmod (2(p+1))$
5. **if** $\left(r^2 == \left|(a,b)\right| (\bmod p)\right)$
6. Compute Lucas sequences $V_m(P = 2c, Q = 1)$ and $U_m(P = 2c, Q = 1)$
7. **else**
8. Compute Lucas sequences $V_m(P = 2c, Q = -1)$ and $U_m(P = 2c, Q = -1)$
9. **endif**
10. $x := hV_m \dfrac{p+1}{2} \bmod p$
11. $y := hU_m c \bmod p$
12. **return** (x,y)

Now we can explain the rational for the algorithm. We are given a Gaussian integer (a,b), prime $p : p \bmod 4 = 3$ and $n : 0 < n < p^2 - 1$. We are trying to compute $(a,b)^n \bmod p$. First we compute

$$r = \left|(a,b)\right|^{\frac{p+1}{4}} \bmod p . \tag{4.1}$$

r^2 is either $-\left|(a,b)\right| \bmod p$ or $\left|(a,b)\right| \bmod p$, depending on whether $|(a,b)|$ is QNR (i.e. has a square root) or not.

$$(c,d) = r^{-1}(a,b) \bmod p . \tag{4.2}$$

Notice that $\left|(c,d)\right| \bmod p = 1$ or -1 because:

$$\left|(c,d)\right| = \left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2+b^2}{r^2} (\text{mod } p) . \tag{4.3}$$

We factored (a,b) into a product of a real integer r and Gaussian integer (c,d):

$$(a,b) = r(c,d) \bmod p . \tag{4.4}$$

To get $(a,b)^n \bmod p$ we compute:

$$r^n = r^{n \bmod (p-1)} \bmod p . \tag{4.5}$$

$$(c,d)^n \bmod p . \tag{4.6}$$

To compute (4.5) we can use any available real integer exponentiation modulo prime algorithm (the order of real integer modulo p is $p-1$, so we can reduce n modulo $p-1$). To compute (4.6) we can use the relationship between Gaussian integers and Lucas sequences described in algorithm in [1]: compute $V_m\left(2c, \left|(c,d)\right|\right)$, $U_m\left(2c, \left|(c,d)\right|\right)$ and

$$(c,d)^n = (c,d)^m = \left(V_m / 2, U_m\right) (\text{mod } p) . \tag{4.7}$$

where $m = n \bmod (2(p+1))$ (the order of Lucas sequences with $Q=1$ is $p+1$ and with $Q=-1$ is $2(p+1)$). We can efficiently compute V_m and U_m using any published algorithm to compute Lucas sequences, such as [9] or [10]. The algorithms published in [9] and [10] would only compute $V_m(2c,1)$, however, they can be easily enhanced to compute $V_m(2c,-1)$, $U_m(2c,-1)$ or $U_m(2c,1)$. Finally we need few more multiplications to get the final answer:

$$(a,b)^n = r^n (c,n)^n (\text{mod } p) . \tag{4.8}$$

The speed advantage of the **LSEG** is due to fact that it does less work. Notice that the most expensive operations in this algorithm are two real integer exponentiations (lines 1 and 3) and one Lucas sequences computation (line 6 or line 8). The speed advantage of **LSEG** would vary depending on the exponentiation algorithm used and there are too many variations to consider. To illustrate the speed advantage of **LSEG**, let's compare it to the sliding window Montgomery Gaussian integer exponentiation algorithm.

As we already stated in Section 1, we will label the sliding window Montgomery reduction exponentiation algorithm for Gaussian integers as **SWG** (Sliding Window Gaussian). It is the same algorithm as the one implemented by GMP, but with real integers replaced with Gaussian integers modulo prime $p : p \bmod 4 = 3$. The prime p is n bits long. Suppose the window size is w and the exponent is

$e: 0 < e < p^2 - 1$. Suppose also that e is such that the number of multiplications is n / w (the best case for sliding window algorithm). This assumption is reasonable because we are dealing with random looking exponents for cryptographic applications. In case this assumption is not true, the **SWG** algorithm would be even slower than **LSEG.** Suppose $t_r(n)$ is the running time of one multiplication of two integers of n bits long and T_{SGW} is the running time of SGW algorithm (ignoring lower order operations like additions). For each Gaussian integer multiplication we have to perform 3 real integer multiplications $t_r(n)$ and 2 Montgomery reduction operations *REDC*(). For each Gaussian integer square we have to perform 2 real integer multiplications $t_r(n)$ and 2 Montgomery reduction operations *REDC*(). Each *REDC*() operation has a cost of $\beta t_r(n)$. We will ignore the precomputations required for the sliding window algorithm, it is larger for **SWG** algorithm, but becomes less significant as n grows. Thus:

$$T_{SGW} = n\left(2t_r(n) + 2\beta t_r(n)\right) + \frac{n}{w}\left(3t_r(n) + 2\beta t_r(n)\right) = \left(2 + 2\beta + \frac{3+2\beta}{w}\right) n t_r(n). \quad (4.9)$$

For **LSEG** we will use the same sliding window exponentiation with Montgomery reduction for real integer exponentiation. We have to do 2 real integer exponentiations with exponents less than have the size of e in bits. To compute Lucas sequences let's use the algorithm [9]. The size of the exponent for this algorithm is approximately one half of the size of e (i.e. n/2) and for each bit we have to do one multiplication, one square and two *REDC*() operations. Let's assume that the square takes 2/3 of the time of multiplication. Suppose T_{LSEG} is the running time of **LSEG,** ignoring lower order operations like additions. As with **SWG**, we ignore precomputation required for sliding window algorithm noting that it is smaller for **LSEG.** Finally:

$$\begin{aligned} T_{LSEG} &= 2\left(\frac{n}{2}\left(\frac{2}{3}t_r(n) + \beta t_r(n)\right) + \frac{n}{2w}\left(t_r(n) + \beta t_r(n)\right)\right) + \\ &+ \frac{n}{2}\left(\frac{2}{3}t_r(n) + \beta t_r(n) + t_r(n) + \beta t_r(n)\right) = n t_r(n)\left(1.5 + 2\beta + \frac{1+\beta}{w}\right) \end{aligned}. \quad (4.10)$$

Using this formula we can estimate the improvement based on various window sizes and values of β.

As we can see from Table 2, the **LSEG** offers an improvement approximately 18% over **SWG** for the window size relevant to real world cryptography applications (1000-4000 bits). Moreover, the improvement could be better if we use the algorithm published in [11] or any other improved algorithm to compute Lucas sequences. The real time could be significantly improved if we compute (4.5) and (4.6) in parallel. For the sake of brevity we will not discuss it any further in this paper and leave it for subsequent publications.

Table 2 The estimated improvement ratio of the running time of **LSEG** relative to the running time of **SWG** algorithm (T_{LSEG}/T_{SWG}) for various window sizes.

	w=1	w=2	w=3	w=4	w=5	w=6	w=7	w=8
β=1	0.61	0.69	0.74	0.76	0.78	0.79	0.80	0.81
β=1.2	0.62	0.70	0.75	0.77	0.79	0.81	0.81	0.82
β=1.4	0.63	0.71	0.76	0.78	0.80	0.82	0.82	0.83
β=1.5	0.64	0.72	0.76	0.79	0.81	0.82	0.83	0.84
β=1.7	0.64	0.73	0.77	0.80	0.81	0.83	0.84	0.84
β=2	0.65	0.74	0.78	0.81	0.82	0.84	0.85	0.85
β=2.2	0.66	0.74	0.79	0.81	0.83	0.84	0.85	0.86
β=2.5	0.67	0.75	0.79	0.82	0.84	0.85	0.86	0.87
β=3	0.68	0.76	0.80	0.83	0.85	0.86	0.87	0.88
β=4	0.69	0.77	0.82	0.84	0.86	0.87	0.88	0.89

We would like to draw your attention to the fact that it is not possible to modify a paper in any way, once it has been published. This applies to both the printed book and the online version of the publication. Every detail, including the order of the names of the authors, should be checked before the paper is sent to the Volume Editors.

5 Experimental Results

For the experiments we used the 5.0.1 release of GMP library on several platforms that vary widely in computing power: 32 bit laptop Intel Core2 Duo CPU T9400 @2.53GHz on Windows XP, 64 bit AMD Opteron 2218 @2.6 GHz Dual core RHEL Linux 4.2 kernel 2.6.9 (64 bit), 64 bit SunOS 5.10 on sun4u Ultra-4 with 2 UltraSPARC-II @296MHz processors. On each platform we installed the library and did the optimization step. We implemented the sliding window exponentiation with Montgomery reduction algorithm for Gaussian integers with optimal sliding window size (**SWG**).

We ran the experiments for bit sizes varying from 100 to 11500 bits. For each bit size we randomly generated a prime $p : p \bmod 4 = 3$ of bit size n and prime q of bit size from $2n$-1 to $2n$. We calibrated the number of trials N for each platform, so we could differentiate between the CPU time readings for lower bit sizes. For each bit size we randomly generated N Gaussian integers $(a,b) : 0 < a,b < p-1$, N exponents $e : 0 < e < p^2 - 1$ and N real integers $c : 0 < c < q-1$. For each of the N Gaussian integers we computed $(a,b)^e \bmod p$ using **SWG** and **LSEG.** Additionally, for each of the N integers c we computed $c^e \bmod q$ using **SWR**. For each algorithm we recorded total CPU time.

For real integer exponentiation we used mpz_powm that came with GMP library. It is implemented using Sliding Window exponentiation algorithm

(algorithm 14.85 in [12]) using Montgomery REDC reduction (section 15.4.2 [7]). We will label this algorithm for real integers as **SWR** (Sliding Window Real). We have tried to make our implementation of **SWG** as efficient as possible, however, it still has more overhead than GMP's "mpz_powm" function because extra additions required by Gaussian integer multiplication and comparative efficiency of GMP's implementation. Nevertheless, our implementation of **SWG** overtook **SWR** for bit sizes below 1000 on all platforms and showed the predicted in previous section 20% speed advantage.

We also implemented **LSEG** that uses **SWR** algorithm for real integer exponentiation and the algorithm published in [9] for Lucas sequences computations. For small bit sizes (500-2500 bits) it performed better than predicted in Table 2. This is due to the fact that it has much less overhead than **SWG**, and our estimates in Table 2 are biased towards **SWG** (make **SWG** look faster). For really high bit sizes (>4000) the experimental results confirm our estimates in Table 2 (show 15% speedup), but for the bit sizes that are practical for cryptography the results show improvement of 40-20%, which is much better than predicted 18%.

As expected, the nominal running time varied widely among, but the relative running times among the algorithms remained consistent. Below are the graphs of the results from the AMD Opteron platform. For the sake of brevity we didn't include the graphs for other platforms, which are very similar in shape.

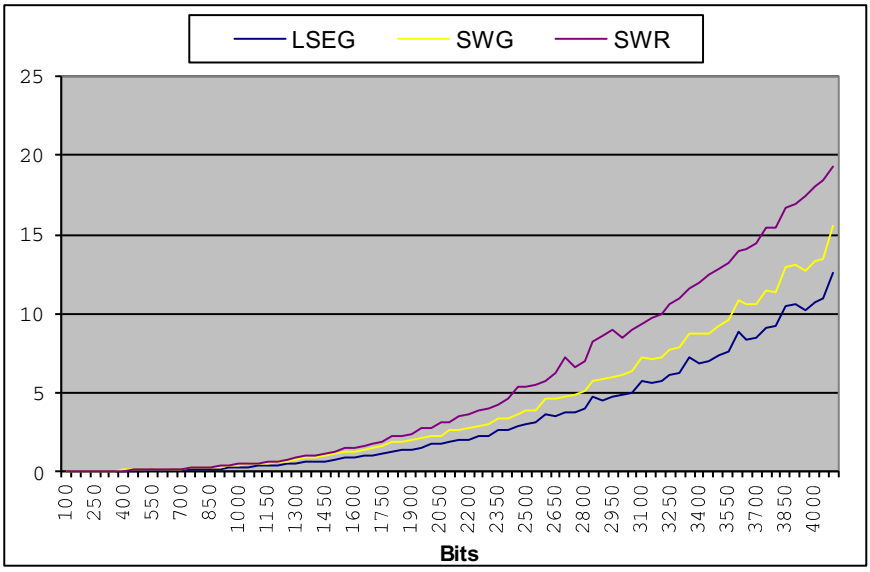


Fig. 1 The running time of the **SWR**, **SWG** and **LSEG** for various bit sizes.

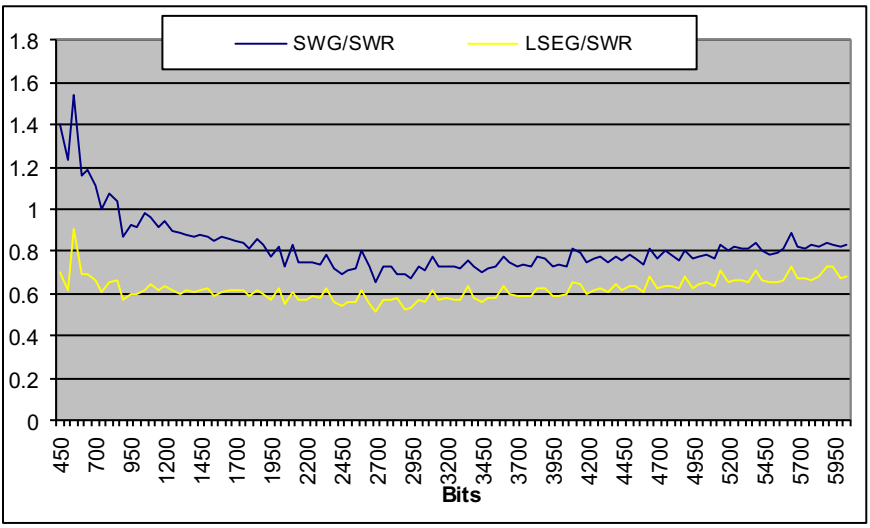


Fig. 2 The running time ratio **SWG/SWR** and **LSEG/SWR** side by side.

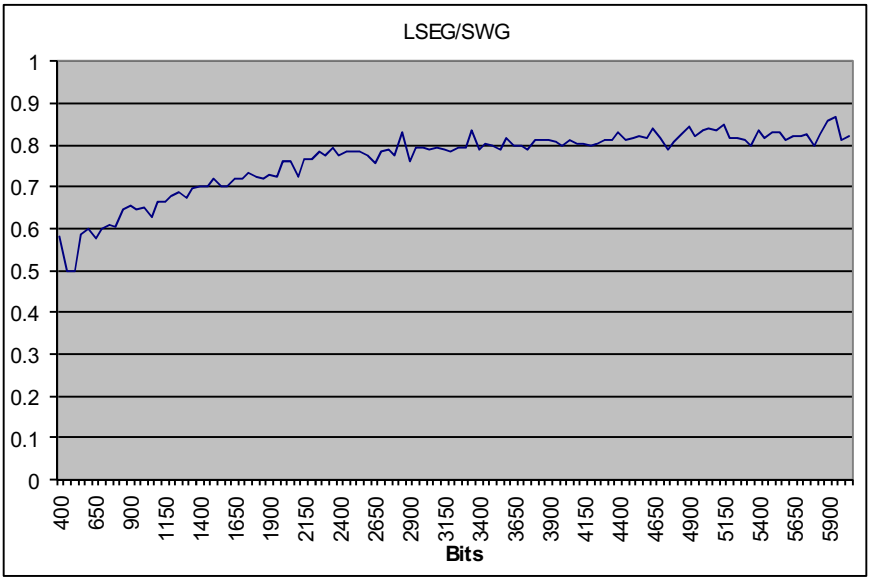


Fig. 3 The ratio of the running time of **LSEG** over **SWG**.

The experimental results confirm the estimates from Table 1 and Table 2**.**

6 Conclusion

We have showed in this paper that exponentiation can be done faster with Gaussian integers then with real integers of comparable size. We showed both theoretically and experimentally that if we simply replace real integers with Gaussian integers we would improve the speed by about 20%. Moreover, we presented a new exponentiation algorithm that improves the speed by additional 18% (on top of 20% which results in about 34% over real integer exponentiation). Consequently, if we use Gaussian integers and the new algorithm presented in this paper, we could immediately improve the speed of Public Key cryptography algorithms that are based on real integer Discrete Logarithm Problem by 34%. Further improvements are possible if we use a more advanced algorithm to compute Lucas sequences.

References

1. Koval, A.: Relationship between Lucas Sequences and Gaussian Integers in Cryptosystems. In: International Conference on Information Technology - New Generations (ITNG), Las Vegas, Nevada USA, pp. 229–233 (2015)
2. Knuth, D.E., The Art of Computer Programming. Addison-Wesley (1998)
3. Karatsuba, A., Ofman, Y.: Multiplication of Many-Digital Numbers by Automatic Computers. Proceedings of the USSR Academy of Sciences **145**, 293–294 (1962)
4. Schönhage, A., Strassen, V.: Multiplikation großer Zahlen. Computing **7**, 281–292 (1971)
5. Fürer, M.: Faster Integer Multiplication. Book Faster Integer Multiplication, Series Faster Integer Multiplication (2007)
6. Montgomery, P.L.: Modular Multiplication Without Trial Division. Mathematics of Computation **144**(170), 519–521 (1985)
7. Granlund, T.: The GNU Multiple Precision Arithmetic Library (2014). https://gmplib.org/gmp-man-6.0.0a.pdf
8. Hasselström, K.: Fast Division of Large Integers. Department of Numerical Analysis and Computer Science, Royal Institute of Technology, Stockholm, Sweden (2003)
9. Yen, S.M., Laih, C.S.: Fast algorithms for LUC digital signature computation. IEE Proceedings Computers and Digital Techniques **142**(2), 165–169 (1995)
10. Wang, C.-T., Chang, C.-C., Lin, C.-H.: A method for computing Lucas sequences. Computers & Mathematics with Applications **38**(11–12), 187–196 (1999)
11. Joye, M., Quisquater, J.J.: Efficient computation of full Lucas sequences. Electronics Letters **32**(6), 537–538 (1996)
12. Menezes, A.J., Van Oorschot, P.C., Vanstone, S.A.: Handbook of Applied Cryptography. CRC Press (1997)

Dynamic Simulation of the Flight Behavior of a Rotary-Wing Aircraft

Sebastião Simões Cunha Jr., Marcelo Santiago de Sousa, Danilo Pereira Roque, Alexandre Carlos Brandão Ramos and Pedro Fernandes Jr.

Abstract In this paper we seek to present the mathematical model relating to flight dynamics of a rotary-wing aircraft. The procedure starts by linearizing the translational and rotational dynamics and rotational kinematic equations of motion based on perturbation theory. Some procedures are simplified and are implemented here, due to the complex modeling process. The second step is to obtain the linear form, which is fundamental in order to describe the stability and response of the small movements of the helicopter around a specific attitude. Finally, the dynamic behavior of two helicopters (AS355-F2 Squirrel and BO105-S123) was simulated employing MATLAB Simulink v7.6 (2008) [1]. The input data, status and control derivatives from an existing helicopter were employed. It is noteworthy that the linearized model developed here is valid for applications where a deeper representation of the aircraft is not required.

Keywords Flight · Aerodynamic · Simulation · Modeling · MATLAB & simulink · Helicopter · Aircraft · AS355-F2

S.S. Cunha Jr. · M.S. de Sousa · D.P. Roque(✉)
Department of Mechanical Engineering,
Federal University of Itajubá, UNIFEI, Itajubá-MG, Brazil
e-mail: sebas@unifei.edu.br, marcelosousa.unifei@gmail.com,
danilo_roque@yahoo.com.br

A.C.B. Ramos · P. Fernandes Jr.
Department of Institute Exact Sciences,
Federal University of Itajubá, UNIFEI, Itajubá-MG, Brazil
e-mail: ramos@unifei.edu.br, Pedro.f.jr87@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_94

Nomenclature

- M_a - Mass of the helicopter [kg].
- X, Y, Z - External aerodynamic forces in x,y,z body axes [N].
- P_e, Q_e, R_e - Angular velocities around x,y,z body axes [rad/s].
- L, M, N - External aerodynamic moments in x,y,z axes [N.m].
- θ_0 - Main rotor collective pitch angle [rad].
- θ_{0T} - Tail rotor collective pitch angle [rad].
- θ_{1s} - Longitudinal cyclic pitch [rad].
- θ_{1c} - Lateral cyclic pitch [rad].
- u, w, v - Translational velocities components of helicopter along fuselage axes [m/s].
- p, q, r - Angular velocity components of helicopter.
- θ, φ, ψ - Euler angles defining the orientation of the aircraft relative to the Earth [rad].
- Ω_a - Main rotor angular velocity [rad/s]

1 Introduction

This paper relates to aerodynamic behavior of rotorcraft, flight controls and dynamic modeling of the helicopter. The study the dynamic modeling of helicopters was developed in order to verify the behavior of the aircraft in situations of a certain type of flight, for example, hovering (Hover).

A linearized model for flight dynamics can be implemented in simulator software, making it a very important tool in aviation. For example, the simulator can be a training tool for the instruction of future pilots. Another application is the analysis of real helicopter behavior, leading to the avoidance of errors, reducing costs and improving projects.

However, the dynamic modeling of helicopters has very complex characteristics. These are due, not only to the structure, but in the main to additional factors which give helicopters greater freedom of movement. The helicopter can move vertically, forwards, sideways, and backwards, and can hover; as well as a combination of all the movements described above.

Problems faced by helicopter flight dynamics are usually resolved through modern computers and software. However, in more complex scenarios, computer models cannot understand the problems and complexities of natural physics. In most cases, calculations can be simplified employing approximate formulas. For example, one calculation may simplify or ignore the effect of the turbulence and other disturbances.

2 Objectives

- Carrying out a review of current research to establish the present state of knowledge in relation to flight dynamic simulation behavior, specific to rotary-wing aircraft.
- Dynamic modeling of AS355-F2 Squirrel (Fennec).
- Simulating the modeling behavior in MATLAB & Simulink v7.6 (2008) [1].
- Disseminating the knowledge from results obtained.

3 Helicopter Flight Control

The helicopter flies using the same principles as an airplane. Lift is generated by the force resulting from the passage of air (relative wind) around the aerodynamic profile or an airfoil, i.e. wings of an airplane and rotary blades of helicopters.

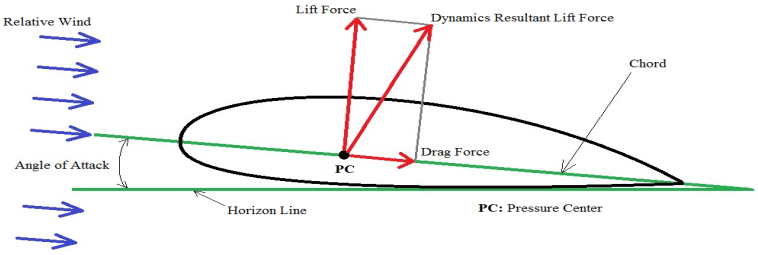


Fig. 1 Aerodynamic forces generated, shown in profile. (Author)

The helicopter presents six degrees of freedom in three spatial axes *x, y, z*. The system of coordinates applied here in order to study the dynamics of a helicopter is modeled on a rigid body (based on the fuselage reference points). Figure 2 shows the resultant forces and moments acting on the helicopter.

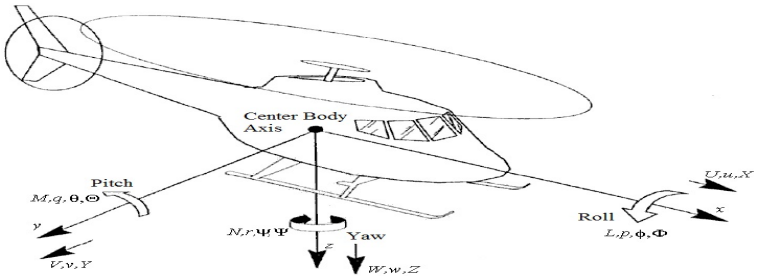


Fig. 2 Coordinate system of rigid body. *(Cruz [2])*

To operate the helicopter the pilot has three command levers (Cyclic, Collective and Pedals) which give control over the four degrees of freedom of movement (Pitch, Roll, Yaw and Flight ascending and descending).

3.1 Collective

The Collective control is responsible for upward and downward movement of the helicopter. This changes the collective pitch angle of all blades simultaneously. For example, if the command is upward (pull the Collective), the resulting angle of attack is increased; if it is downward (push the Collective) the resulting angle of attack is decreased.

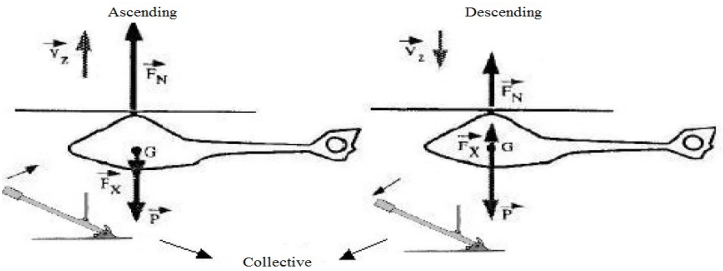


Fig. 3 Collective: Control ascending and descending. *(Firmino [3])*

3.2 Cyclic

The Cyclic control creates the longitudinal (Pitch) and side (Roll) movements of the helicopter. This changes the pitch of the blades in a different way, depending on the position over the course of rotation. For example, if the command is forward (Cyclic forward), the pitch of the blades decreases when they advance and increases when they retreat. If the command is to the left side (Cyclic left), the pitch on the blades is reduced when passing through the left side and increases when they pass through the right side.

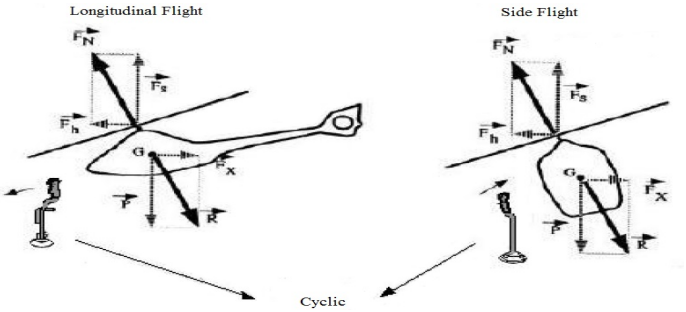


Fig. 4 Cyclic: Longitudinal and lateral control. *(Firmino [3])*

3.3 Pedal

The Pedal control is responsible for the rotational movement around the vertical z-axis (Yaw) of the helicopter. The tail rotor operates in the same way as the Collective on the main rotor, i.e. the step of the blades modifies simultaneously with the same angle of attack. The tail rotor of the helicopter is essential for stability and may neutralize the torque generated by the main rotor on the fuselage.

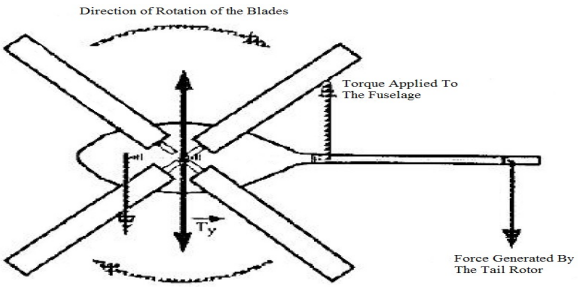


Fig. 5 Pedal: Yaw Control, helicopter nose right or left control *(Firmino [3])*

4 Linearization of Helicopter Motion Equations

The equations relating to motion, that is, the six degrees of freedom for the fuselage were developed from Newton's law, which were then applied to the reference axes of the helicopter. These take gravity into account (Padfield [4]).

$$\dot{x} = func(x, c, t), \tag{1}$$

The six degrees of freedom consist in the states of motion (x) and control (c), respectively.

$$x = \{u, v, w, p, q, r, \theta, \emptyset, \psi\}, \tag{2}$$
$$c = \{\theta_0, \theta_{1s}, \theta_{1c}, \theta_{0T}\} \tag{3}$$

The forms of non-linear equations of motion are as follows:

A. Force Equations:

$$\dot{u} = (rv - qw) + \frac{X}{M_a} - gsin\theta, \tag{4}$$
$$\dot{v} = (pw - ru) + \frac{Y}{M_a} - gcos\theta sin\emptyset, \tag{5}$$
$$\dot{w} = (qu - pv) + \frac{Z}{M_a} + gcos\theta cos\emptyset, \tag{6}$$

B. Moments Equations:

$$I_{xx}\dot{p} = \left(I_{yy} - I_{zz}\right)qr + I_{xz}(\dot{r} + pq) + L, \tag{7}$$
$$I_{yy}\dot{q} = (I_{zz} - I_{xx})rp + I_{xz}(r^2 - p^2) + M, \tag{8}$$

$$I_{zz}\dot{r} = \left(I_{xx} - I_{yy}\right)pq + I_{xz}(\dot{p} - qr) + N, \tag{9}$$

C. Euler angle Equations:

$$\dot{\emptyset} = p + qsin\emptyset tg\theta + rcos\emptyset tg\theta, \tag{10}$$
$$\dot{\theta} = qcos\emptyset + rsin\emptyset, \tag{11}$$
$$\dot{\psi} = qsin\emptyset sec\theta + rcos\emptyset sec\theta, \tag{12}$$

The equations of motion are linearized by a Taylor series, as proposed by Prouty [5] and Cooke & Fitzpatrick [6].

$$X = X_e + \frac{\partial X}{\partial_u}\Delta u + \frac{\partial X}{\partial_w}\Delta w + \frac{\partial X}{\partial_q}\Delta q + \frac{\partial X}{\partial_\theta}\Delta\theta + \frac{\partial X}{\partial_v}\Delta v + \frac{\partial X}{\partial_p}\Delta p + \frac{\partial X}{\partial_\emptyset}\Delta\emptyset + \frac{\partial X}{\partial_r}\Delta r + \frac{\partial X}{\partial_\psi}\Delta\psi +, \frac{\partial X}{\partial_{\theta_0}}\Delta\theta_0 + \frac{\partial X}{\partial_{\theta_{1s}}}\Delta\theta_{1s} + \frac{\partial X}{\partial_{\theta_{1c}}}\Delta\theta_{1c} + \frac{\partial X}{\partial_{\theta_{0T}}}\Delta\theta_{0T}. \tag{13}$$

The linearized equations can be differentiated in time in order to produce the derivatives for stability and control. Thus the equations can be described in matrix form.

$$\dot{x} = Ax + Bc, \tag{14}$$

D. State Matrix [A]:

$$A = \begin{bmatrix} \frac{X_u}{m} & \frac{X_w}{m} & \frac{X_q}{m} - w_0 & -g\cos\theta_0 & \frac{X_v}{m} & \frac{X_p}{m} & 0 & \frac{X_r}{m} + v_0 \\ \frac{Z_u}{m} & \frac{Z_w}{m} & \frac{Z_q}{m} + u_0 & -g\cos\phi_0 sen\theta_0 & \frac{Z_v}{m} & \frac{Z_p}{m} - v_0 & -gsen\phi_0\cos\theta_0 & \frac{Z_r}{m} \\ \frac{M_u}{I_{yy}} & \frac{M_w}{I_{yy}} & \frac{M_q}{I_{yy}} & 0 & \frac{M_v}{I_{yy}} & \frac{M_p}{I_{yy}} & 0 & \frac{M_r}{I_{yy}} \\ 0 & 0 & cos\phi_0 & 0 & 0 & 0 & -sen\phi_0 & \\ \frac{Y_u}{m} & \frac{Y_w}{m} & \frac{Y_q}{m} & -gsen\phi_0 sen\theta_0 & \frac{Y_v}{m} & \frac{Y_p}{m} + w_0 & g\cos\phi_0\cos\theta_0 & \frac{Y_r}{m} - u_0 \\ L'_u & L'_w & L'_q & 0 & L'_v & L'_p & 0 & L'_r \\ 0 & 0 & sen\phi_0 tg\theta_0 & 0 & 0 & 1 & 0 & \cos\phi_0 tg\theta_0 \\ N'_u & N'_w & N'_q & 0 & N'_v & N'_p & 0 & N'_r \end{bmatrix}$$

E. Control Matrix [B]:

$$B = \begin{bmatrix} \frac{X_{\delta_B}}{m} & \frac{X_{\delta_C}}{m} & \frac{X_{\delta_A}}{m} & \frac{X_{\delta_P}}{m} \\ \frac{Z_{\delta_B}}{m} & \frac{Z_{\delta_C}}{m} & \frac{Z_{\delta_A}}{m} & \frac{Z_{\delta_P}}{m} \\ \frac{M_{\delta_B}}{I_{yy}} & \frac{M_{\delta_C}}{I_{yy}} & \frac{M_{\delta_A}}{I_{yy}} & \frac{M_{\delta_P}}{I_{yy}} \\ 0 & 0 & 0 & 0 \\ \frac{Y_{\delta_B}}{m} & \frac{Y_{\delta_C}}{m} & \frac{Y_{\delta_A}}{m} & \frac{Y_{\delta_P}}{m} \\ L'_{\delta_B} & L'_{\delta_C} & L'_{\delta_A} & L'_{\delta_P} \\ 0 & 0 & 0 & 0 \\ N'_{\delta_B} & N'_{\delta_C} & N'_{\delta_A} & N'_{\delta_P} \end{bmatrix}$$

One may note that the ψ angle has been omitted since the flight direction in the horizontal plane has no effect on the aerodynamic or dynamic forces.

5 AS355-F2 Squirrel Data

The AS355-F2 Squirrel is a twin-engine lightweight helicopter (2.5 tons) manufactured by Aerospatiale in France. There are civilian and military versions. The military version, illustrated in Figure 6, is the subject of this research.

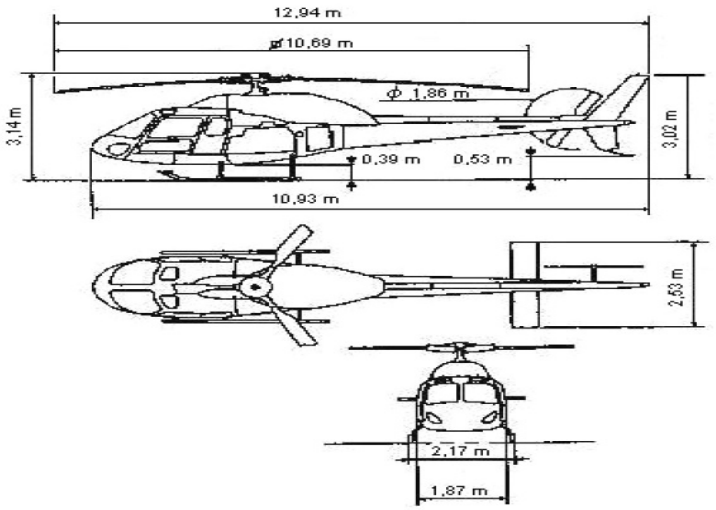


Fig. 6 Basic dimensions of the AS355-F2 Squirrel. *(Helibras [7])*

The derivatives of stability and control for the estimated parametric model are listed below:

Table 1 Derivatives of Stability:

$X_u = -0.1495$	$X_v = 0.0818$	$X_w = 0.1173$	$X_p = 7.1374$	$X_q = 2.6031$	$X_r = 2.3649$
$Z_u = -0.9061$	$Z_v = -0.1338$	$Z_w = 0.4057$	$Z_p = 6.060$	$Z_q = 8.6978$	$Z_r = 7.1931$
$Y_u = -3.3813$	$Y_v = -0.2218$	$Y_w = 0.8711$	$Y_p = -0.8151$	$Y_q = 0.9261$	$Y_r = -1.6897$
$L_u = -0.1356$	$L_v = -0.0508$	$L_w = 0.0379$	$L_p = -2.6332$	$L_q = 0.5898$	$L_r = 0.2696$
$M_u = 0.1335$	$M_v = 0.0038$	$M_w = -0.0812$	$M_p = -0.1946$	$M_q = -3.0734$	$M_r = -0.9387$
$N_u = -0.025$	$N_v = 0.0442$	$N_w = 0.0188$	$N_p = -0.6053$	$N_q = 0.6118$	$N_r = -1.2201$

Table 2 Derivatives of Control:

$X_{\theta_0} = 0.4582$	$X_{\theta_{1s}} = 0.2450$	$X_{\theta_{1c}} = -0.6233$	$X_{\theta_{0T}} = 0.1489$
$Z_{\theta_0} = -0.6980$	$Z_{\theta_{1s}} = -2.3090$	$Z_{\theta_{1c}} = -2.0854$	$Z_{\theta_{0T}} = -1.7667$
$Y_{\theta_0} = -8.4282$	$Y_{\theta_{1s}} = 1.1003$	$Y_{\theta_{1c}} = 0.858$	$Y_{\theta_{0T}} = -0.8089$
$L_{\theta_0} = -0.3824$	$L_{\theta_{1s}} = -0.1526$	$L_{\theta_{1c}} = 0.4121$	$L_{\theta_{0T}} = -0.0216$
$M_{\theta_0} = 0.1579$	$M_{\theta_{1s}} = 0.3456$	$M_{\theta_{1c}} = 0.1442$	$M_{\theta_{0T}} = 0.0508$
$N_{\theta_0} = -0.8907$	$N_{\theta_{1s}} = -0.1502$	$N_{\theta_{1c}} = 0.0674$	$N_{\theta_{0T}} = 0.2821$

Table 1 and Table 2: derivatives of stability and control. *(Cruz [2])*

5.1 Modelling Helicopter Flight Dynamics in MATLAB & Simulink

The state (A) and control (B) matrices were determined in MATLAB with the data of the dynamic forces and moments of a flying AS355-F2 Squirrel. In order to check the stability and rigidity of the matrix the command (damp) was applied.

```
>> damp (A)

      Eigenvalue                 Damping      Freq. (rad/s)

  4.44e+001                  -1.00e+000       4.44e+001
 -1.19e+000 + 2.78e+000i      3.93e-001       3.02e+000
 -1.19e+000 - 2.78e+000i      3.93e-001       3.02e+000
 -4.94e-001                   1.00e+000       4.94e-001
 -2.02e-001                   1.00e+000       2.02e-001
  1.76e-001                  -1.00e+000       1.76e-001
  7.78e-004                  -1.00e+000       7.78e-004
  8.10e-006                  -1.00e+000       8.10e-006
```

It can be observed here that positive values were obtained from the Eigenvalues, which indicate that the system is unstable. To resolve this problem a state matrix K was created by applying quadratic linearization (LQ), which provides feedback regarding the system and ensuring that it is stable.

$$\dot{x} = [A - BK]x + Bc, \tag{15}$$

$$y = Qx, \tag{16}$$

This can easily be solved using the function (LQR) in MATLAB. The matrix Q in equation 16 is the identity matrix 8x8.

```
C = diag([ones(1,8)]);
D=zeros(8,4);
[K]= lqr (A,B,C,diag([1 1 1 1]))
```

Checking the stability and rigidity of the resulting matrix again against the feedback, is carried out using the command (damp) in MATLAB, and resulted in:

```
>> damp(A-B*K)

      Eigenvalue                 Damping      Freq. (rad/s)

 -4.44e+001                   1.00e+000       4.44e+001
 -8.36e+000                   1.00e+000       8.36e+000
 -3.70e+000 + 3.55e+000i      7.22e-001       5.13e+000
 -3.70e+000 - 3.55e+000i      7.22e-001       5.13e+000
 -1.79e+000                   1.00e+000       1.79e+000
 -1.97e-001                   1.00e+000       1.97e-001
 -1.13e-003                   1.00e+000       1.13e-003
 -2.33e-005                   1.00e+000       2.33e-005
```

Note that all the real parts of the Eigenvalues are negative, indicating the stability of the system. In order to obtain the transfer function, the following controls are performed:

```
[K]= lqr (A,B,Q,diag([1 1 1 1]))
p=85;
[n1,d]=ss2tf(A-B*K,B,p*C,D,1);
[n2,d]=ss2tf(A-B*K,B,p*C,D,2);
[n3,d]=ss2tf(A-B*K,B,p*C,D,3);
[n4,d]=ss2tf(A-B*K,B,p*C,D,4);

figure(1)
tf(n1(1,:),d)
step(n1(1,:),d)
```

After performing some tests the selected weighting factor (p = 85) was chosen to reproduce the maximum vertical speed in the collective pitch angle.

In Figure 7, the schematic dynamic model implemented in Simulink, with the matrices A, B, C, and K calculated, can be observed.

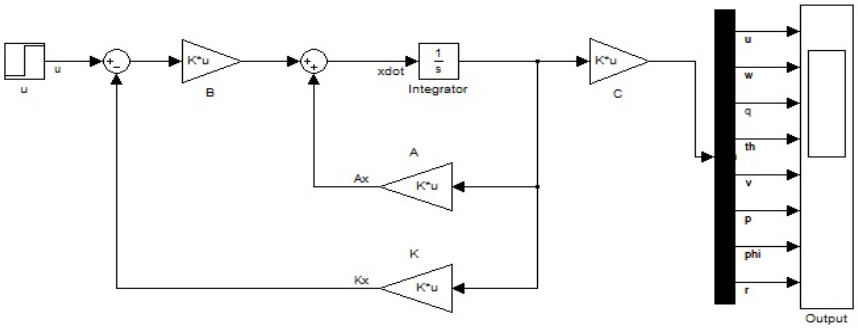


Fig. 7 Schematic Simulink. *(Author)*

The following two simulations are described below: The first one targets upward flight, reaching a maximum rate of climb of 1500 ft/min where 7.62 was achieved (see Figure 8). An input of collective angle of $\theta_0 = 7°$ resulted in disturbances on the other axes. To stabilize the angle of longitudinal and lateral tilt, the application of the longitudinal cyclic angle was changed to $\theta_{1s} = -2{,}75°$ and lateral cyclic angle to $\theta_{1c} = -2{,}5°$, as a result the maneuver was stabilized.

The second experiment was an example of forward flight. In order to achieve 100 knots forward speed (52m/s), the longitudinal cyclic input angle was changed to $\theta_{1s} = -10°$. As with the previous maneuver, the longitudinal cyclic angle was added and applied; a collective angle $\theta_0 = -2{,}5°$ was also applied in order to maintain altitude, and a lateral pitch angle of $\theta_{1c} = -5°$ was applied to counteract lateral movement of the aircraft. Figure 8 and 9 plot variations in the simulation generated for the AS355-F2 Squirrel.

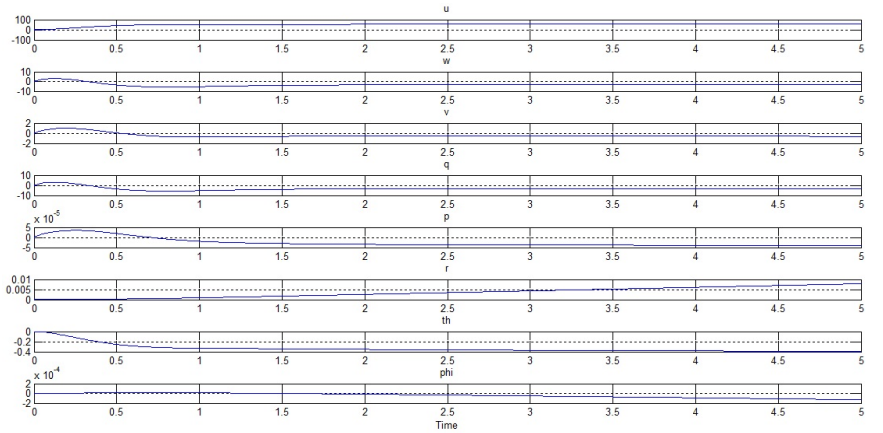


Fig. 8 Upward Simulation AS355-F2 Squirrel. *(Author)*

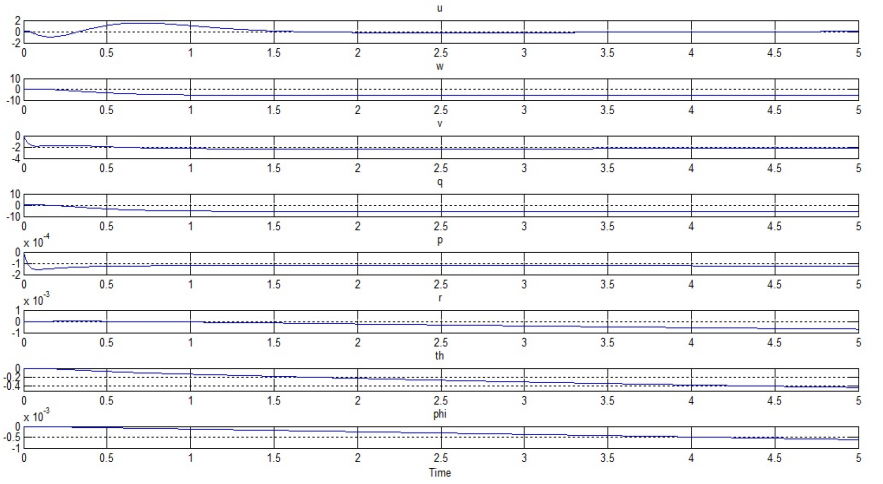


Fig. 9 Forward Simulation AS355-F2 Squirrel. *(Author)*

Unfortunately, further information is not available regarding best values in order to achieve the best performance for each set of output parameters. As further design and performance information becomes available, simulator modeling will improve, recreating more faithfully movements and controls relating to this helicopter.

5.2 BO105-S123 Data

In order to validate the information obtained from the AS355-F2 Squirrel, we have employed the same simulation procedures used for specifying stability and control in the BO105-S123 which is a parametric model described by Padfield [4].

The BO105-S123 was chosen for comparison because it falls into the same category (weight and number of crew members) as the AS355-F2 Squirrel. See Figure 10 for dimensions.

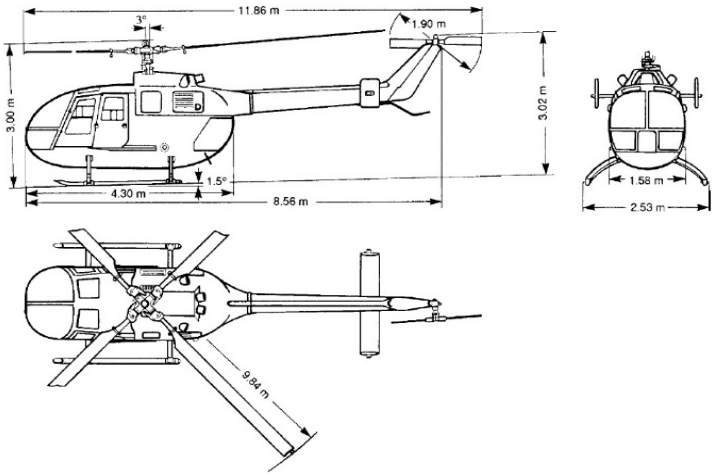


Fig. 10 Dimensions for the BO105-S123 aircraft. *(Padfield, Cap 4 pag. 264 [3])*

The two simulations for BO105-S123 are described below. For the upward flight and the maximum rate of climb of 1500ft/min the collective input the value of $\theta_0 = 5°$ causes disturbances in the other axes. This results in a longitudinal cyclic angle of $\theta_{1s} = 0{,}55°$ and lateral cyclic angle of $\theta_{1c} = 3°$. In the second experiment- forward flight- (100 knots), the longitudinal cyclic input angle was changed to $\theta_{1s} = -10.5°$ and, due to other factors, the collective pitch angle was adjusted to $\theta_0 = 0{,}25°$ and the lateral pitch angle to $\theta_{1c} = -3°$. Figures 11 and 12 show the results of the two simulations for the BO105-S123 described above.

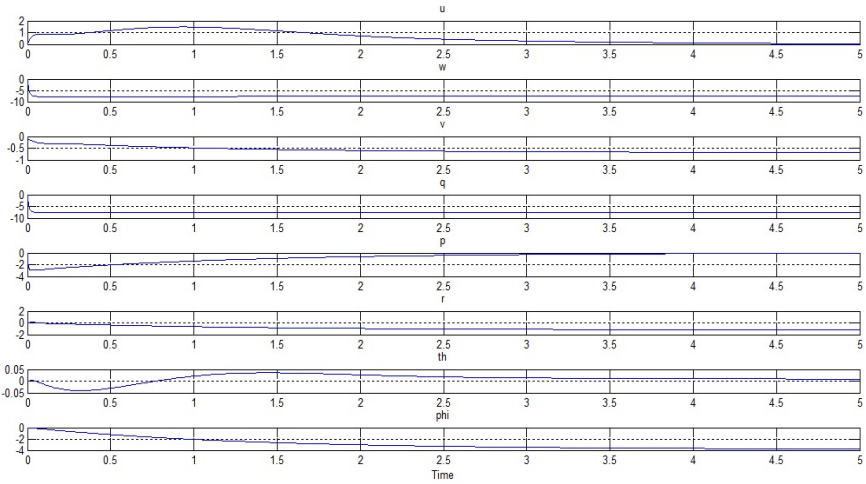


Fig. 11 Upward Simulation BO105-S123. *(Author)*

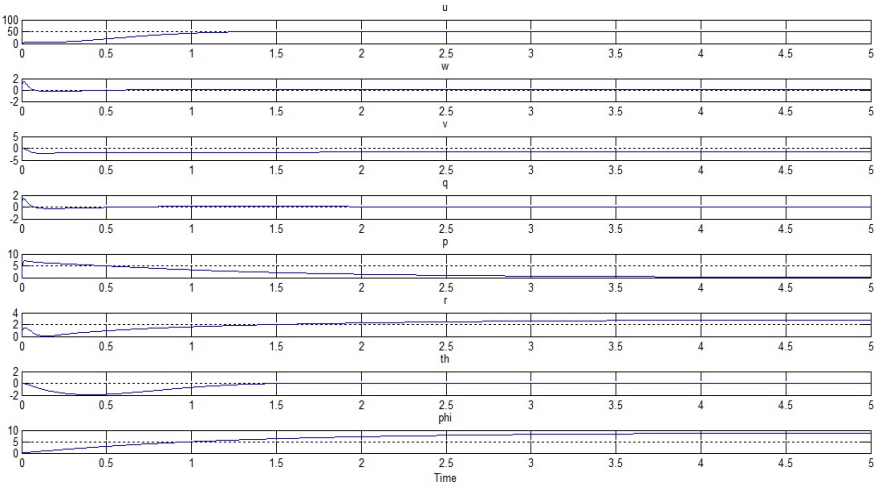


Fig. 12 Forward Simulation BO105-S123. *(Author)*

Please note that for both helicopters, the objectives and solutions were similar, especially in forward flight with a different angle -on the longitudinal Cyclic- the difference being only 0.5º between the two aircraft. This variation may be attributed to the different aerodynamic characteristics of the two aircraft, that is the forces and moments of inertia whilst similar are different, changing the attitude of the two helicopters. This confirms the choice of these aircraft for this study.

The AS355-F2 Squirrel appeared to be a dynamic helicopter with greater maneuverability due to its more aggressive aerodynamic design, which means that the aerodynamic influences are also higher, increasing the response rate due to disturbances.

On the other hand, the BO105-S123 presents a less streamlined aerodynamic design, especially in relation to the fuselage bubble shape, producing greater forces and smaller rates of disturbances.

6 Conclusion

This paper demonstrates procedures used in order to create a basic mathematical model that represents helicopter flight dynamics for the AS355-F2 Squirrel. The procedure started with linearizing the translational and rotational dynamics and rotational kinematic equations of motion using the small perturbation theory. The following step was the development of linearized equations to describe the stability in response to small movements in a helicopter when trim condition are applied.

In order to simulate the results, the input data and state and control derivatives from an existing helicopter were employed. The project, however, is not yet fully developed due to the lack of information on the performance specifications of AS355-F2 Squirrel. However, it is noteworthy that the linearized model developed here is valid for applications where a more detailed representation of the aircraft is not required.

References

1. MATLAB & Simulink 7.1 (2008)
2. Cruz, V.R.: Desenvolvimento de um Modelo Dinâmico para Simuladores de Helicoptero. ITA: São José Dos Campos, BR (2009)
3. Firmino, L.F.: Simulação e Controle de um Helicoptero a Partir de Modelos Linearizados em Sete Pontos de Operação. IME, Rio de Janeiro (2008)
4. Padfield, D.G.: Helicopter Flight Dynamics: The Theory and Aplication of Flying Qualities and Simulation Modeling. AIAA Education Series, Reston, USA (1996)
5. Prouty, R.W.: Helicopter Performance, Stability and Control. Malabar Krieger, IND, pp. 443–637 (1995)
6. Cooke, A.K., Fitpatrick, E.W.H.: Helicopter Test and Evolution. AIAA Education Series, Reston, USA (2002)
7. HELIBRAS: Manual de Voo do Helicoptero AS355-F2 Esquilo. Itajuba, BR (2005)
8. Homa, M.J.: Aerodinâmica e Teoria de Voo, 22th edn.. ITA, São José Dos Campos (2004)
9. Houghton, E.L., Carpenter, P.W.: Aerodynamics for Engineering Students, 5th edn. The University of Warwick, UK (2003)
10. Mahamood, F.A.A.: Constructing & Simulating a Mathematical Model of Longitudinal Helicopter Flight Dynamics. Article
11. Santos, R.A.L.: Modelo Computacional do Rotor Principal do Helicóptero em Men-on-the-Loop. Técnico Lisboa, Lisboa
12. Santos, W.V.: Modelagem, Identificação e Controle de Altitude deum Helicóptero em Escala Reduzida. UFRJ, Rio de Janeiro (2005)

Origami Guru: An Augmented Reality Application to Assist Paper Folding

Nuwee Wiwatwattana, Chayangkul Laphom, Sarocha Aggaitchaya and Sudarat Chattanon

Abstract Origami folders often find themselves unable to recall the entire sequences of foldings, and origami drill books have been the only resources for them to learn to fold. However, many people have difficulties understanding symbols and operations in drill books. The purpose of this research is to design and develop a mobile augmented reality application that acts as a technological aid to origami folders. Extensive user studies have been employed at every step of the implementation. The application recognizes the current state/shape of the paper and instantly overlays the diagrams and instructions of the next step(s) on the real paper through the smartphone's camera view. Evaluation results show that the application can successfully assist folders most of the time, and the qualitative mean score is about the same level of most mobile applications. This work will benefit a lot of origami practitioners such as those in geometry classes and intellectual development classes.

Keywords Augmented reality · Memory aid · Interaction design · Paper folding · Origami

1 Introduction and Motivation

The traditional art of folding flat papers into things is usually associated with Japanese culture and is generally known as Origami. Because of its entertaining and educational benefits, paper folding is often taught in early school years or young at home. In higher education, algorithms and data structures of folding and crease patterns are studied in mathematics and computer science. The principles of origami have also been used to solve real-world problems and inspired cutting-edge developments in science and engineering, such as origami stent grafts [1], nanoscale DNA origami [2],

N. Wiwatwattana(✉) · C. Laphom · S. Aggaitchaya · S. Chattanon
Computer Science Program, Department of Mathematics, Faculty of Science,
Srinakharinwirot University, Bangkok 10110, Thailand
e-mail: {nuwee,sc541010324,sc541011150,sc541010375}@swu.ac.th

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_95

Table 1 Summary of the preliminary user study

Observation/Question	User1	User2	User3	User4	User5
Have you ever folded a crane?	Yes	No	Yes	Yes	Yes
Did he/she consult the book?	Always	Always	No	Sometimes	No
Did he/she fold through trial and error?	No	No	Yes	Yes	No
Have you thought about the next fold before consulting the book?	No	No	N/A	Yes	N/A
Did he/she finish the crane?	Yes	Yes	Yes	Yes	Yes

folded automotive airbags, or folding solar panels for space satellites [3], to name a few.

There are an innumerable number of origami models foldable by hands ranging from basic and popular models (e.g., crane) to advanced and unimaginable models. Instruction books described in Yozhizawa-Randlett system [4] (also called origami drill books) have been the only resources to learn folding sequences. Until recently, the internet has become so widespread and origami instructions have been published on the web. At present, we are in the world where smartphones are becoming part of our organs. As a result, many drill books are built in the form of mobile applications. In the Apple's App Store and the Google's Play Store, there are over three hundred of such applications available for download (at the time of the preparation of this manuscript), which indicates significant interest in paper folding. Most of the applications are free; some of them are gamified; some of them are video-based; some are in 3D, while others are traditional 2D drawings.

It is not possible for everyone to be able to recall the entire sequences of particular origami models as folding papers are not everyday activities. To study how people fold a simple model and the extent to which the drill book aids the folding, five undergraduate students, two men and three women, were asked to fold a typical "crane" in our preliminary study. An origami drill book diagramming the crane model was provided from the beginning. Table 1 summarizes questions that were asked and observations that were made. Results from the preliminary study show that the drill book plays an important role in finishing the model regardless of folders' familiarity with the model. However, there are arguments we would like to note.

- It is not always plausible for folders to follow the instructions diagrammed in the books or applications. Although Yozhizawa-Randlett system provides a standard set of folds, operations, and symbols, not everyone has studied the system before.
- Some folders can not relate 2D portions in the books with the real objects they are folding, especially if the models are complex, as argued in [5].
- Instruction books and their mobile application counterparts are too similar to cheat sheets, and that would defeat the pleasure of paper folding.
- A paper drill book does not come in handy when it is needed.

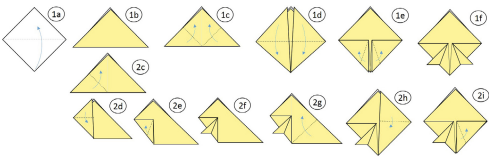


Fig. 1 Two possible paths of folding a samurai helmet: (1) 1a-1b-1c-1d-1e-1h-1i-1f (2) 1a-1b-2c-2d-2e-2f-2g-2h-2i-1f. The rest of the steps are omitted.

Fig. 2 Cue cards used in the second user study

Nobody has ever denied the fact that modern technology has greatly and positively assisted people in learning. Lately, a growing number of handheld augmented reality (HAR) applications are developed to improve the way a human learns. In 2015 NMC Horizon Reports, the renowned five-year horizon reports from an international not-for-profit consortium of learning-focused organizations, the expert panel listed augmented reality as one of the visualization technologies to help process information, identify patterns, and sense order in complex situations [6]. As a matter of fact, we thought that augmented reality technology could aid folders to learn and construct origami in a way.

In this research, we aim to develop an intelligent application utilizing the augmented reality to help people fold origami models. This paper is organized as follows. Starting from the ground up in Section 2, we establish requirements and develop designs from the mind process of the folders and folding sequences modeling. Section 3 addresses technical details of the implementation of the research prototype, and Section 4 evaluates the prototype. Subsequently, works related to our research are discussed in Section 5. Finally, Section 6 concludes our work and discusses future work.

2 Establishing Requirements and Developing Designs

We define our project as a mobile application that can aid origami folders, both novices and experts, to finish the models with the help of augmented reality technology. Our preliminary study has already addressed the needs for such application. Establishing requirements includes understanding the nature of the folding, i.e., the analytic process of the folder. In order to finish a paper model, the folder needs to know the sequence of the folding.

Let us first examine the sequences of folding several models. The technique of going from one (initial/immediate) shape to another (intermediate/final) shape might be a standard one—like a mountain fold, a valley fold, or a reverse fold—or complex one—like a squash fold or a rabbit ear fold. A folding operation consists of both the folding technique and the folded location. As seen in drill books, the folding

sequences are often captured in one fix path of sequences. However, many models can be accomplished in different paths. One example of these is folding a samurai helmet, shown partly in Fig. 1. There are two possible paths to fold the samurai helmet. The first path is to fold both the left and the right flap at the same time. The second path is to get an antler-like kuwagata, one flap at a time. Several models also share a number of standard bases as the first few steps in construction. All in all, a set of origami models folding sequences can be described by a state transition diagram, originally stated in [5].

2.1 State Transition Diagram (STD)

We formally define a state transition diagram of origami models' folding sequences as a quintuple:

$$STD(Q, \sum, \delta, q_0, F) \tag{1}$$

where:

- Q is a finite set of states, each state is a unique shape of (finished/unfinished) origami.
- $\sum$ is a finite set of folding operations.
- δ is the folding function, represented as arrows between states labeled by the folding technique and the folded location, $\delta \subseteq Q \times \sum \times Q$.
- q_0 is the start state, a.k.a initial shape, usually a flat rectangular paper, $q_0 \in Q$.
- F is a finite set of end states, a.k.a. completed origami models, $F \in Q$.

For folders who have never folded the model before, they generally imitate the instructions in drill books step by step. For folders who are accustomed to fold the model, they will use their memories to recall the folding operations—a conscious process of problem solving. We already knew from the preliminary study that when folding some people tried to fold through trial and error and some people thought about the next fold before consulting the book. We decided to continue studying the nature of folders' mind process by direct observation. We would like to know the stepwise thinking of folders and apply that thinking into the design. This leads to another round of user study.

2.2 The Second User Study

In the second user study, ten undergraduate students were recruited to fold the "crane" again. We prepared instructions in the form of stepwise cue cards as shown in Fig. 2. Unlike the preliminary study, we handed the next-step card to folders only when it is requested. The hypothesis is that, at any step, the folder would like to know one step

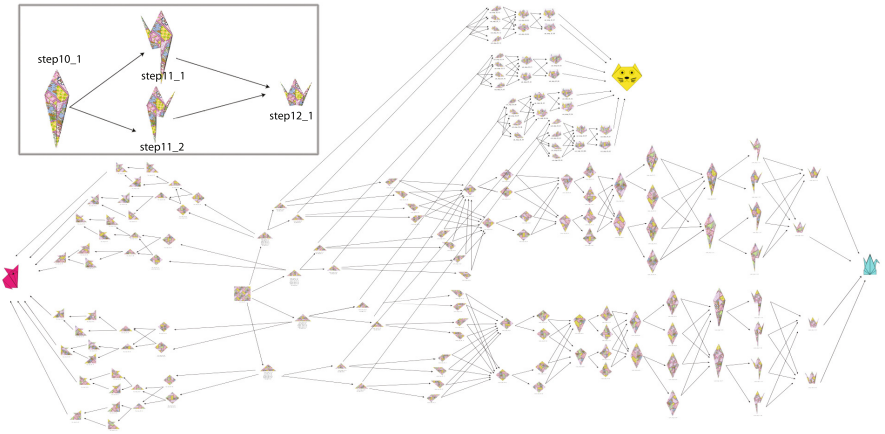


Fig. 3 The actual state transition diagram of three models: crane, fox , and cat face. The insert in the top left shows part of the three states of the crane. One sub-state can be folded into multiple sub-states, and multiple sub-states can be folded into one sub-state.

ahead. It turned out that nine students asked for the cards at least once (One person was very professional!). Five of them, who declared themselves as having folded the crane many times, asked to see the next step and the step after the next before continuing to fold. The rest of them asked for the next three steps. In addition, some of the students, especially those who rarely fold the crane, had difficulties in understanding the diagrams and had to ask for help sometimes.

The results from the second study confirmed the findings that parts of human memories are short-lived and that humans tend to use heuristics. They tend to carry out the folding by finding a short term solution, step by step. For this reason, our application is in a better form to support just enough information for technological intervention. This could be done with augmented reality to superimpose the additional information upon real-world objects in situ. To assist folders relate the real-world 3D papers to the diagrams, we can project the next-step diagram over the paper using the mobile screen and camera. Thinking about the situation when a folder is folding an origami, but finds that he cannot recall the next step. He can take out his mobile phone and points the camera to the currently unfinished paper. Instantly, the application suggests him what he should do next with the diagram and instructions floating right on top of the paper. Folders can readily map the current face, the current crease, or the current edge of the paper to the diagram and easily understand the instant direction of the next folding operation(s). Not only the next step, but we also added two more next operations to the application. Another consideration about perceptive cognition suggested that we should not use the same color/pattern of the diagram to that of the paper, otherwise folders may not be able to distinguish between reality and virtuality.

3 Implementation

We developed a prototype application on the Android platform using Unity [7] and Vuforia's development platform [8]. Incorporated the requirements and designs discussed earlier, the prototype has two main sections: the real-time model recognition and the static diagrams. The latter is an electronic version of drill books for beginners or for studying the entire folding sequences. The former is the augmented reality module, which we shall give more details in the following sub-sections.

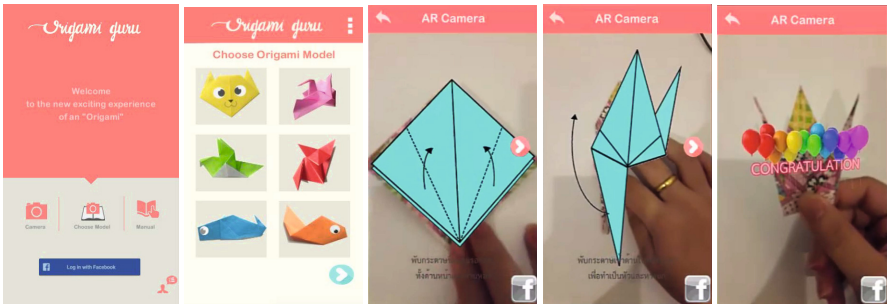
3.1 Architecture

The system consists of three layers of software performing all tasks in an Android application unit, which are made possible by Unity and Vuforia. The bottom layer is a driver to control and send real-time feed from the camera. The middle layer is an augmented reality (AR) engine. The AR engine performs a live image analysis of the camera feed, recognizes the paper model, estimates the 3D position of the paper (pose estimation), and renders the virtual diagrams and instructions. The top layer is an interactive interface and live screen that users see. Users can interact with the application such as rotating the camera and zooming the camera.

In Vuforia, each real object that the developer would like to recognize is called a target (a marker in a broad sense). The image targets are kept in the Vuforia's device database. To be a valid target, and to be able to perform a high quality pattern recognition, an ample amount of feature points has to be extracted. We found that plain color papers do not produce a high enough number of feature points. Thus, we needed to use a sophisticated pattern of papers with a lot of feature points. The pattern is shown in the middle of Fig. 3. A target is associated with one or more virtual diagram to be superimposed on. All targets and virtual diagrams are to be registered with the AR engine.

3.2 Data Structures

A state in the state transition diagram is, however, not a one-to-one mapping to an AR target. This is because of two main reasons. (1) An origami model is a 3D statue that has an unlimited number of perspectives. There is no automatic way, as of yet, to directly recognize an arbitrary 3D object in Vuforia. (2) An origami paper, says a rectangle with 4 sides: a-b-c-d, can be folded in different directions. For example, a valley folding from side a and c is the same as a valley folding from side b and d. If the paper pattern is not symmetric, like ours, there will be multiple image patterns for a given shape. Therefore, we had to create multiple targets for each state of the origami model. We call each of them a sub-state. A sub-state is folded to one or more



(a) Home page (b) Choose page (c) Step 6 (d) Step 11 (e) Last step

Fig. 4 Screenshots of the mobile application before and during the folding of a crane

sub-states of the next state. Fig. 3 depicts an actual state transition diagram that we built. Out of 8 models available in the application, only 3 models—the crane, the fox, and the cat face—are shown here due to space limitation. The insert on the top left is a close-up of the crane model. A sub-state of state no. 10 (step10_1) can be reverse-folded on the left leg to the first sub-state of state no.11 (step11_1), or can be reverse-folded on the right leg to the second sub-state of state no.11 (step11_2).

Each image target has a corresponding identifier. Each sub-state in the state transition diagram is a JavaScript Object Notation (JSON) object. Each JSON object consists of the model identifier, the image target identifier, the state identifier, the next-state virtual diagram object identifier, the second-next-state virtual diagram object identifier, and the third-next-state virtual diagram object identifier. Three next states are maintained because, according to the user study, this is the number of steps folders generally requested at a time. At present, although a sub-state can be folded into several next sub-states, we only encoded one path.

3.3 User Interface and Functions

Our application is called "Origami Guru". The user interface is designed using the material design principle. On the homepage, users can login with their Facebook accounts. Users can choose between three modes of functions, shown in Fig. 4a. The first mode opens the camera view and let users point the camera to the paper model they are folding. If there are multiple models that belong to the shape that is being recognized, as are the cases for the first few folds, the application asks users to choose the model they are folding. The virtual diagram (with instructions) of the next step is superimposed on the paper model with an arrow button to superimpose the step after the next, and so on.

The second mode lets users choose the model they are folding first before opening up the camera view. Choosing the model page is shown in Fig. 4b. A partial screen

flow of recognizing the crane model is shown in Fig. 4c, Fig. 4d, and Fig. 4e. Users can also share their model with the Facebook community via the share button. The third mode is the static diagram page.

4 Evaluation

To measure the quality of our application, we conducted two evaluation studies and derived two scores: the average accuracy of recognizing the folded paper, and the overall usability score. The LG G2 smartphone with Android 4.4 (KitKat) was used in these studies. The first study measured the average accuracy of recognizing 6 models across all foldable shapes and was conducted by our team of researchers. For each model, we performed ten rounds of foldings. For each shape, we calculated the consistency percentage, which is the percentage the application was able to recognize as the correct model(s) within 5 seconds using the second mode. Fig. 5 shows the comparison between these 6 models. Detecting the pigeon's folded papers yields the highest accuracy of 88 percent, followed by the cat face, the crane, the fox, the fox face and the fish respectively. This is because the pigeon's folded shapes have very rich natural features such as corners and crease patterns. The less, the harder, such as the fox face or the fish. On average, out of 10 times, there are about 7 times that the application were able to detect models correctly. However, these numbers can vary depending on the phone camera, the environment (lighting and contrast), and the augmented reality platform.

We recruited 20 participants, 9 men and 11 women of age between 15 and 25, to evaluate the application. We asked each participant to select one origami model that he/she has folded before, gave him/her the paper, and explained the basis of the application. Due to its evidence as a reliable method for web and mobile usability assessment, the Software Usability Scale (SUS) [9] was employed in our study. The standard version of SUS was translated into Thai and distributed to the participants after they finished the test. As shown in Fig. 6, the mean usability score is 69.72 (SD = 13.14)

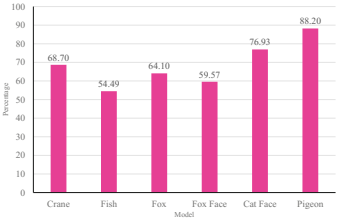


Fig. 5 The average accuracy of recognizing each origami model across all foldable shapes

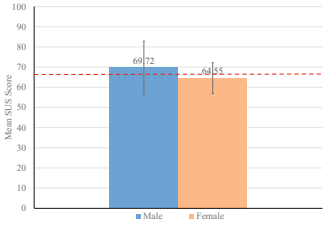


Fig. 6 The mean score for male and female participants according to the Software Usability Scale (SUS)

for men and 64.55 (SD = 7.65) for women. However, this does not show evidence of a statistically significant gender difference (p-value = 0.32, by the two-sample t-test with unequal variance). Results involving all participants are acceptable with an overall mean score of 66.88 (SD = 10.51) in red dashed line, about the level of the mean for web and mobile interface type, a "C" grade, according to [10].

Comments from participants suggested that the application is innovative and surprised them about how augmented reality can be combined with folding origami. One participant even encouraged us to add more models to the application. Another participant suggested superimposing animated instructions in place of the current static picture and suggested that we should opt for hands-free devices.

5 Related Work

Shimanuki et al. [11] argued that some people had difficulties with the origami drill books because paper books could not show all portions of the shape. Their research proposed a vision mechanism to extract diagrams from the drill books and converted them into digital graphics that can be explored in 3D spaces. The most-related researches are the interactive system to support folding operations by Watanabe and Kinoshita [5] and their earlier works [12], and [13]. Nonetheless, their system uses an algorithmic approach to recognize folding processes via a web-camera system while ours is an enhancement in design to incorporate augmented reality technology to a smartphone application.

Nobu Kobayashi [14] introduced Goal Structuring Notation (GSN) to the process of origami folding. His goal was to study the implicit knowledge of folding to "make it beautiful" and to make the knowledge explicit by GSN. We are not concerned about folding beautifully, but rather to finish it, although GSN can be used in the same manner for our work. Furthermore, if folders follow the diagram precisely, shape recognition is easier.

Different works have provided practical examples of technology support for memory capacity limits. Only a few of them are listed here. Cook's Collage [15] provides surrogate memory support for memory-slipped cooking tasks using a display of six most recent actions in a temporal sequence. Geo-reminder, a type of mobile applications to remind its users based on location as well as time, is often shipped with the smartphone platform.

6 Conclusion and Future Work

In this paper, we are the first to propose a mobile application to aid origami folding using the augmented reality technology. We investigated how folders consult the drill books for guidance of steps through a series of user studies. Results from the requirement studies were carefully used in the design of our prototype. The prototype

application we developed are able to recognize 8 origami models. Once recognized, the diagrams and instructions of the next folding operations are superimposed on the real paper model in the camera view. Experimental evaluation has proved that the application is able to perform correctly most of the time. The user satisfaction is quite positive. Our application is expected to come in handy for people who have problems recalling origami folding sequences and those who have problems reading diagrams and instructions.

The application is a preliminary prototype and therefore still has many more rooms for improvement in terms of usability and technologies used. In order to make the application more usable, we have to overcome the limitation of the fixed paper pattern, possibly by implementing an origami recognition that is specific to detect creases, faces, corners, and edges; for instance, works by [5] and [16]. An even harder question is how can we identify wrong foldings and suggest the correction. It is interesting to know whether the time-to-completion when using our application is better than searching through a drill book, and if yes, in which circumstances. Yet, less time-to-completion does not always mean more pleasure because sometimes struggling and recalling in between processes are the elements of satisfaction, not only the finished product. In games, we feel fun because of the playing we attempted, not just because the boss are finally killed.

Acknowledgment This research was supported by a grant from Srinakharinwirot University, Thailand, under grant no. 778/2558. The authors express our warm thanks to Dr. Thitima Srivanatakul for her comments and to Srinakharinwirot University for the travel grant.

References

1. Kuribayashi, K., Tsuchiya, K., You, Z., Tomus, D., Umemoto, M., Ito, T., Sasaki, M.: Self-deployable Origami Stent Grafts as a Biomedical Application of Ni-rich TiNi Shape Memory Alloy Foil. Materials Science and Engineering: A **419**(1–2), 131–137 (2006)
2. Smith, L.M.: Nanostructures: The Manifold Faces of DNA. Nature **440**(7082), 283–284 (2006)
3. Nishiyama, Y.: Miura Folding: Applying Origami to Space Exploration. International Journal of Pure and Applied Mathematics **79**(2), 269–279 (2012)
4. Randlett, S.: The Art of Origami; Paper Folding, Traditional and Modern, 1st edn. E P Dutton, June 1961
5. Watanabe, T., Kinoshita, Y.: Folding support for beginners based on state estimation of origami. In: Proceedings of IEEE Region 10 Conference TENCON, pp. 1–6 (2012)
6. Johnson, L., Brown, M., Becker, S.A., Cummins, M., Diaz, V.: NMC Horizon Reports (2015)
7. Unity - Game Engine. https://unity3d.com/
8. Qualcomm Vuforia Developer Portal. https://developer.vuforia.com/
9. Brooke, J.: Usability Evaluation in Industry, chap. SUS: A Quick and Dirty Usability Scale, pp. 189–194. Taylor and Francis, London, June 1996
10. Sauro, J.: A Practical Guide to the System Usability Scale: Background, Benchmarks and Best Practices. CreateSpace Independent Publishing Platform (2011)
11. Shimanuki, H., Kato, J., Watanabe, T.: Recognition of folding process from origami drill books. In: Proceedings of Seventh International Conference on Document Analysis and Recognition, pp. 550–554 (2003)
12. Kinoshita, Y., Watanabe, T.: Estimation of Folding Operation Using Silhouette of Origami. IAENG International Journal of Computer Science **37**(2) (2008)

13. Kato, J., Watanabe, T., Nakayama, T., Guo, L., Kato, H.: A model-based approach for recognizing folding process of origami. In: Proceedings of Fourteenth International Conference on Pattern Recognition (1998)
14. Hiranabe, K.: How to make it better? - Extracting implicit knowledge in Origami (paper folding) via GSN. http://astahblog.com/2014/08/24/
15. Tran, Q.T., Calcaterra, G., Mynatt, E.D.: Cook's collage: deja vu display for a home kitchen. In: Proceedings of Home Oriented Informatics and Telematics (HOIT 2005), York, UK (2005)
16. Zhu, K., Fernando, O.N.N., Cheok, A.D., Fiala, M., Yang, T.W., Samani, H.A.: A SURF-based natural feaure tracking system for origami recognition. In: Proceedings of 20th International Conference on Artificial Reality and Telexistence, Adelaide, Australia (2010)

Augmented Reality Approach for Knowledge Visualization and Production (ARAKVP) in Educational and Academic Management System for Courses Based on Active Learning Methodologies (EAMS–CBALM)

Helen de Freitas Santos, Wanderley Lopes de Souza, Antonio Francisco do Prado and Sissi Marilia dos Santos Forghieri Pereira

Abstract The education has configured itself by the transfer of teacher's knowledge to the student, without the proper criticism or reflection by the student. Thereby, the main objective of the education, the student's learning, is not being achieved and it has taken people to a constant questioning of this traditional education and giving space to the Active Learning Methodologies (ALM). Such methodologies use the questioning as the strategy of the learning, where the student builds his own knowledge, by problems that given to him, making this student a being reflexive and critical. In recent decades Information and Communication Technologies (ICT) are used more and more on education, being even more useful when in ALM environment. In this sense, the objective of this paper is to present an Augmented Reality Approach for Knowledge Visualization and Production (ARAKVP) in Educational and Academic Management System for Courses based on Active Learning Methodologies (EAMS–CBALM).

Keywords Augmented Reality · Active Learning Methodologies · Learning Environment · Augmented Reality Approach

H. de Freitas Santos
Department of Computing, Federal Institute of Education,
Science and Technology of São Paulo, Campus Birigui, Birigui, SP, Brazil
e-mail: helen@ifsp.edu.br

W.L. de Souza · A.F. do Prado · S.M. dos Santos Forghieri Pereira
Department of Computer Science and Department of Medicine,
Federal University of São Carlos, São Carlos, SP, Brazil
e-mail: {desouza,prado}@ufscar.br, sissimarilia@uol.com.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_96

1 Introduction

Augmented Reality (AR) can be defined as an overlaying of digital data, generated by computer, visualized over image obtained from real environment [1]. Therefore, when we use AR technologies is possible to blend real objects with virtual objects, putting the suitable information on real environments, augmenting the real world environment. There are several environments where the AR can be inserted and one of those is the learning environment (LE). The teaching environment with traditional methodology became obsolete, uninteresting and discouraging. Searching for a more attractive teaching learning process, new methodologies, like Active Learning Methodologies (ALM), emerge, and occupy a prominent place. On the other hand, new technologies like AR started to be seen as an ally of education to promote the innovation in LE. It allows to converge areas of education and entertainment, creating new opportunities to support teaching and learning in formal and informal contexts [2]. Thereby, the New Media Consortium (NMC) has assumed the role of mediator between education and technology, trying to modify the human behavior, making effective the use of technology and trying to bring important innovations over the coming years to the field of education [3]. AR is one of the technologies listed in NCM report as an innovation in education in the coming years when talking about wearable technologies [4]. Therefore, based on the NMC 's instructions, trying to join concepts and aspects of computer science in education around AR, specifically for courses based on ALL, this study presents an Augmented Reality Approach for Knowledge Visualization and Production (ARAKVP) in Educational and Academic Management System for Courses based on Active Learning Methodologies (EAMS–CBALM).

2 Active Learning Methodologies (ALM)

2.1 Presentation of ALM

When traditional learning methodology is used the trend is that teachers teach concepts and students simply reproduce them when performing theirs learning tasks. However, this methodology fails because it underuses the human capability, this situation worsens when the teaching-learning process also requires association between problems and actions. To improve this scenario, it has been proposed ALM, an educational concept that encourages critical reflection on teaching-learning process, raises the student's participation level and go in the opposite direction to the methods and techniques that emphasize the transmission of knowledge. The educator participates as a mediator and the student becomes a critical being.

2.2 ALM Examples at Brazil

At Brazil, courses from Federal University of São Carlos (UFSCar) and Institute for Teaching and Research of the Sirio Libanês (IEP) use ALM. Such courses are referenced here as CBALM, and support on the experiences and theoretical foundation of problem-based learning, using the constructivist spiral as a reference (Fig 1) [5][6], which is formed by six movements, driven in a anticlockwise.

2.3 Curricular Structure of CBALM

The CBALM curricular structure is formed by levels, where the inner level identifies the educational units (EU), which represent the axes through which the curriculum is developed and where the class load is defined. Differentiated educational activities (EA) are prepared for the EU and their purpose are development of cognitive, attitudinal and psychomotor skills. Examples of EA are: problem situation, narrative, team based learning, plenary, workshop, travel, distance education, self-directed learning (SDL), portfolio of accomplishments, paper for end of the course and an action research called applied project. The EA have learning triggers (LT), which are problems that simulate or reflect the daily activities to be practiced in the future, by the CBALM egress. The LT initiate the constructivist spiral movements, then the students are encouraged to exercise reflection and consequently they are stimulated to develop the cognitive, attitudinal and psychomotor skills. From each LT the students have to identify the problem, formulate explanations, prepare question, look for new information, build new meanings and evaluate the process, through all the motions of constructivist spiral.

Fig. 1 Constructivist Spiral of teaching-learning process from the exploration of a trigger (extracted from [5],translated and adapted from Lima, V.V. learning issues raised by students during PBL tutorials compared to curriculum objectives. Chicago, 2002 [Master Dissertation – University of illinois at Chicago. Department of Health Education].

2.4 Community of CBALM

The teaching-learning community of CBALM is formed by students, teachers, course coordinator and educational project coordinator. Teachers can play the

following roles: legislators, experts, facilitators and learning manager. The educational project coordinators can play the role of general coordination, pedagogical and administrative. Beyond these actors, the community can also be formed by representatives of educational institutions' outside agencies such as Ministry of Health and National Council of Health Secretaries (CONASS) [5], organizations involved where students can exercise their professional practice.

2.5 Process of CBALM

The teaching-learning process CBALM is developed through the following steps: 1) the educational project coordinator defines the offer for each course; 2) the learning manager prepares school calendar and the offer curricular structure; 3) experts define the EA, as well as its programming, and LT; 4) facilitators choose the content to capabilities development for each LT, such contents are selected from the profile for each EU and from previous knowledge of students; 5) facilitators guide the EA's LT during the school period; 6) students participate in the meetings, realize the EA, develop their cognitive, attitudinal and psychomotor skills, produce and share their knowledge, working individually or in groups and covering the movements of constructivist spiral; 7) experts draw up assessments for each AE; 8) students and facilitators evaluate the teaching-learning process of each EA; 9) learning manager, subsequently, detects improvements to be implemented in the next offering of the course. Students have reserved and protected periods during the week, which are considered spaces of SDL, where students can devote themselves to the search for new information. The trajectory of each student, as well as the facilitator, during the construction process, production and acquisition of knowledge should be consolidated into portfolios.

2.6 CBALM and Evaluation

Evaluation is an ongoing activity in CBALM that allows the monitoring of the teaching-learning process by identifying to improve processes, products and results. It is performed by all of those involved in the process who express their perceptions and indicate the relevant aspects, those that need to be improved, reworked or replaced. There are two types of evaluations: formative and summative, and possible concepts to be received/given are satisfactory, needs improvement and unsatisfactory [6]. Evaluations are consolidated by applying a set of assessment tools, and they are applied over time in the teaching-learning process develops. The assessment tools used are: Performance Evaluation in Teaching-Learning Process (PETLP), Reflective Portfolio (RP), Progress Test (PT), Objective and Structures Assessment of Professional Performance (OSAPP), Problem Based Exercise (PBE) [6]. The student who does not reach a satisfactory concept in the ratings must undertake improvement plans, re-test or recovery, proposed by the facilitator [6].

3 EAMS–CBALM

The Medicine Department (DMed) of UFSCar uses ICT to record their academic and educational activities through two systems: Moodle (Modular Object-Oriented Dynamic Learning Environment) and Integramed. The Moodle is a free Learning Management System (LMS) designed to educators, administrators and students to create personalized learning environments [7]. The Integramed is a proprietary software and it was developed by former students of the computer science course, from the Computer Department (DC) of UFSCar. The academic activities recorded in Integramed should be transferred, subsequently and manually, to an Integrated Academic Management System called SIGA, under responsibility of UFSCar. It is possible to imagine the difficulties faced by DMed as the integration of all academic and educational information: rework, data inconsistency and community discontent are just a few examples. In addition, some specific pedagogical activities of ALM are not covered by Moodle. Therefore, the DC, the Dmed and the IEP have joined to build a custom system to comply the specifics of CBALM: curricular organization and improvement of the technological environment. This system is called EAMS–CBALM and its architecture is described in the following sections.

3.1 Architecture of EAMS–CBALM

The EAMS–CBALM consists of two modules: academic and pedagogical. The academic module records the course structure and curriculum. The course structure allows you to create one or more courses. Each course can be offered at various times, composing offers and the offerings can be subdivided into one or more classes. First students must be registered at EAMS–CBALM, later they are enrolled in the courses, and then they are allocated in groups. The curriculum framework allows you to create levels for the curriculum representation, which must have a minimum of two and a maximum of five levels. The curriculum must always have the levels of curricular activity and educational action.

The pedagogical module must be understood as an overlay of the curricular structure to each class of the course, both previously recorded in the academic module EAMS–CBALM. From this view the activities of the teaching-learning process are programmed and they originate the educational environment on which students and teachers perform and record their activities. The programming starts with the planning of level educational action, the more specialized level of the curriculum. The educational action is materialized through meetings. Each meeting comprises a movement of constructivist spiral. Finally, the programming ends with the preparation of evaluation. Students and teachers perform and record their activities in the educational module. The EAMS–CBALM offers discussion forums, sending text documents, images, video or audio, the creation of individual authorship document or collaborative authorship with other students and versioning of documents. All these activities help the capability development and has its

peak at the knowledge production. Another feature of EAMS–CBALM regards the data visualization. All educational content produced during the teaching-learning process is provided in curricular trajectory. Thus, the curricular trajectory includes since the teaching contribution until the knowledge production of students, and also include their own or shared production. The curricular trajectory is viewed in two ways, the timeline or associated with the EU, and allows students and teachers to view all practices developed during the teaching-learning process more adequately at every moment. In addition, the production of knowledge is indexed by keywords, facilitating the location of any content produced throughout the teaching-learning process.

3.2 The Strengths and Weakness in the EAMS–CBALM

In this section, we analyze the strengths and weaknesses of EAMS–CBALM. The strengths for students are: indexing of knowledge, individual or collaborative document editing and versioning of documents. The strengths for teachers are in carrying out the assessment with concepts and building improvement plans. For both roles, students and teachers, activities viewed as a form of curricular trajectory is also a strong point. The weak points are: 1) limit in the levels of the curricular structure, 2) lack of system intelligence to bring students and teachers to the point of curricular activity when accessing the system and 3) a few formats of questions available for the preparation of ratings. Endowing the EAMS–CBALM of certain intelligence avoids a lot of clicks and waste of time.

4 Discussion

Two aspects deserve attention: students must produce their knowledge when using ALM and AR is an emerging technology for data visualization and it should be available in the coming years. Integrating these two assumptions is interesting and a challenge arises: allow students without prior knowledge of computational aspects to develop their skills and share it, making use of AR technology. Thus, here is the scenario: students produce knowledge using the AR; then share the knowledge produced; students and docents visualize the knowledge produced/shared through immersion, interaction and involvement in the real scene; a new scene is generated and it has virtual objects inserted by computing increasing the real world and then the theory becomes more real and contextualized. Some questions arise when we try to include some AR technologies in ALM environments: 1) how do the students, in general and specifically ALM students, produce their knowledge making use of AR, 2) how do we make AR content production so simple to be produced by those who has not computing knowledge and 3) how do we analyze the true AR contribution to CBALM?

Recently, a systematic review of AR in medical education said that we need to deepen the study of the use of AR in health education since students need to

experience and that interaction is not always possible with real objects; technology itself is not enough to promote learning, even though it was presented as an engine for education reforms; new researches on AR should be held in the medical field trying to find the AR model and appropriate educational projects, besides identify the most effective way for the use of AR in health education and before assert the great potential of AR is important to have a clear understanding of its impact on learning [8].

Therefore arises the ARAKVP approach in EAMS–CBALM that will identify people, processes, structures, strategies, techniques and results, which serve as a source for practical and principles of using AR in CBALM.

5 ARAKVP in EAMS–CBALM

ARAKVP uses AR technology, computer graphics and Ubiquitous Computing. As for hardware, it makes use of laptop, TV, Kinect and mobile devices. The approach is based in EAMS–CBALM, images library of projects like Visible Human Project (VHP), Visible Korean Human (VKH), Chinese Visible Human (CVH) and a software to support the CBALM academic community's knowledge acquisition and production (Fig 2). This section presents the approach and software architecture that supports it.

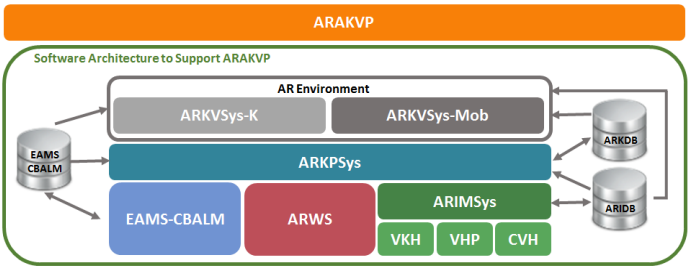


Fig. 2 ARAKVP Overview

5.1 ARAKVP

Each CBALM has its particularities and specificities. The differences and similarities observed in such courses as curriculum guidelines, educational complex, individualized beings and their profiles and environments were identified, studied and analyzed to propose this approach. In addition, studies of AR, AR for education systems, AR's impact on medical education, comparison between AR features, AR advantages and disadvantages in CBALM and expected results with the use of AR also contributed to the development of this approach.

The approach has been built from the selection of grounds for its composition. These grounds are the basis for the software architecture implementation, which will allow assessment of the approach and AR impact study on the CBALM

teaching-learning process. The choice of these grounds was based on the study of the environments that apply ALM, the nowadays hardware and software technologies and the studies by [3] and [8]. The ARAKVP consists of 15 grounds, listed in table 1.

Table 1 ARAKVP Grounds

Ground	Description
Approach Overview	Use of AR in CBALM
Learning Theory	Problem Based Learning
Learning Strategy	SDL
Approach Objective	Prove that AR can contributed to the CBALM teaching-learning process
AR Role	Enrich the current environment of the teaching-learning process by inserting virtual objects (images, videos, sounds and texts), generated by the computer over the real scene
How? (methodological)	Students create stories that represent their reflection about criticality and competence acquired during the teaching-learning process
How? (technological)	Students without prior knowledge of computing are able to justify their learning in the face of questioning using AR as visualization technology. This occurs through the production of mixed scenarios, consisting of real scenario and virtual objects, previously prepared and stored in a database.
Actors	Learning manager, legislator, expert, facilitator and student
Prerequisite	EAMS–CBALM, images library of projects like Visible Human Project (VHP), Visible Korean Human (VKH), Chinese Visible Human (CVH)
Hardware Technology Used	Laptop, computer, data server, application server, smartphone, tablet, mobile internet, kinect, data show, video, TV
Software Technology Used	AR, Computer Graphic, Ubiquitous Computing, Grails, Java, Angular, Javascript, PostgreSQL, ARToolkit, C#, Visual Studio, Unity 3D
Cost Efficiency	This approach is financially accessible to CBALM community and educational institutions because it uses low cost hardware infrastructure
Expected Results	Own style of study, time learning individually discovery for each student creates his knowledge, ALM strengthening, teaching-learning process strengthening. These are hypotheses to be confirmed after ARKVSys-K and ARKVSys-Mob evaluation
User Acceptance	The academic community is made up of young people who already use smartphone and Kinect, mostly . Thus, the user acceptance is achieved. These are hypotheses to be confirmed after ARKVSys-K and ARKVSys-Mob evaluation
AR Impact on the Teaching-Learning Process	SDL performed anywhere, unlimited resources to perform SDL in the sense of supply any number of students, popularization of AR in CBALM , use of new data visualization techniques in CBALM. These are hypotheses to be confirmed after ARKVSys-K and ARKVSys-Mob evaluation

5.2 Software Architecture to Support ARAKVP

Some software elements should be developed to validate and evaluate the ARAKVP. One of those elements was developed and presented in section III. The others are classified in databases, web services, content management system and AR system. The software architecture has, in total, seven elements, presented below.

First, the computational model, also called AR images, must be created. The AR images are produced from images of VHP projects, VKH and CVH. It is made through the CMS ARIMSys (Augmented Reality Images Management System) responsible for cataloging, indexing and storing each image in ARIDB (Augmented Reality Image Data Base). Moreover, it must exists an EAMS–CBALM (described in section III) for the software elements work to support ARAKVP. The EAMS–CBALM should be extended, giving rise to ARKPSys (Augmented Reality Knowledge Production System). The ARKPSys aims to store the knowledge produced by the student during the SDL activity. Web services and APIs (Application Programming Interface) must be developed to facilitate the implementation of ARKPSys, forming a set of classes and services called ARWS (Augmented Reality Web Services). The knowledge produced by the student is stored in ARKDB (Augmented Reality Knowledge Data Base). From that moment the student can make use of the AR environment to view your knowledge produced using any of the AR systems: ARKVSys–K or ARKVSys–Mob. The ARKVSys–K (Augmented Reality Knowledge Visualization System using Kinect) presents the knowledge produced by the student in an environment whose infrastructure provides the use of Kinect while ARKVSys–Mob (Augmented Reality Knowledge Visualization System using Mobile Device) uses a technology with a lower cost: the student's own smartphone. With those systems the student selects the knowledge that he wants to show to the academic community. It consists of a scenario formed by AR images and notes, previously prepared and recorded in ARKDB. Then, the AR content is generated: virtual objects are mixed with the real environment, the student interacts with the virtual objects, shares his knowledge, contextualizes with virtual images very close to reality and finishes a teaching-learning process cycle enriching the learning environment.

6 Related Works

This research is supported by some previous works that use AR in education. Some studies found provide the learning content to the student, but this idea is contrary to the CBALM proposal. A study aimed to investigate whether AR with haptic interface would be an effective tool for learning anatomy [9] [10]. Other proposed a bone anatomy learning system with AR through a system called Bone Puzzle AR. It uses AR and puzzle to teach bone anatomy [11]. Mirracle, an AR Magic Mirror is a prototype that allows the visualization of virtual anatomical information about the actual body of the user and the user controls the Mirracle

through hand gestures and voice commands [12]. BodyExplorerAR is a system that provides interaction methods in mannequin simulators and shows the anatomy of the human body as selection made by the student on the dummy [13]. In addition to these studies, two articles made systematic review of the literature on AR in educational environments. They point out that AR is just beginning in education and that more research needs to be conducted about the subject AR in education [14] [15].

7 Conclusion

This paper presented an AR approach for EAMS-CBALM and a software architecture to support such approach. For this, we studied CBALM processes, strategies, results, people, structures and techniques of UFSCar and IEP courses that uses ALM. The approach consists of 15 grounds to define the scope of AR in the teaching-learning process. These grounds cover the gap in the use of AR in health education studies and they describe one appropriate educational project to use AR in health educational environment. Beyond approach, this paper also presented the software architecture to support de ARAKVP. This architecture is formed by the elements: ARIMSys, ARIDB, EAMS-CBALM, ARWS, ARKPSys, ARKDB, ARKVSys–K and ARKVSys–Mob. As future work will be developed each one of the software elements. After that, we will evaluate the AR environments with UFSCar and IEP students and then issue an opinion about the impact of the use of AR in CBALM.

Acknowledgment Our thanks to Federal Institute of Education, Science and Technology of São Paulo for financial sponsorship.

References

1. Owen, M., Owen, S., Barajas, M., Trifonova, A.: Pedagogic Issues and Questions from the Science Centre to Go, Augmented Reality, Project Implementation. In: EDEN 2011 – Open Classroom Conference. Proceedings of the “Science Centro To Go” Workshops, October 13-30 2011. Ellinogermaniki Agogi, Athens (2011). Haller, apund
2. Owen, M., Owen, S., Barajas, M., Trifonova, A. Pedagogic Issues and Questions from the Science Centre to Go, Augmented Reality, Project Implementation. In: EDEN 2011 – Open Classroom Conference. Proceedings of the “Science Centro To Go” Workshops. October, 13-30 2011. Ellinogermaniki Agogi, Athens (2011). Salmi, apud
3. New Media Consortium. http://www.nmc.org/about (Accessed 21 October 2015)
4. Johnson L., Adams Becker, S., Estrada, V., Freeman, A.: NMC Horizon Report: 2013 Higher Education Edition. Austin, Texas: The New Media Consortium (2014)
5. Tempski, P. et al.: Caderno do Curso Educação na Saúde para Preceptores do SUS, Instituto Sírio Libanês de Ensino e Pesquisa (2014)

6. Ribeiro, A.C. et al.: Caderno do Curso de Medicina UFSCar, Universidade Federal de São Carlos
7. Moodle 2015. https://docs.moodle.org/29/en/About_Moodle (Accessed October 21, 2015)
8. Zhu, E., Hadadgar, A., Masiello, I., Zary, N.: Augmented reality in healthcare education: an integrative review. PeerJ **2**, e469 (2014)
9. Yeom, S-J.: Augmented Reality for Learning Anatomy. In: Proceedings Ascilite 2011 Changing Demands, changing directions, 4-7 December 2011, Tasmania, pp. 1377–1383 (2011). ISBN 978-1-86295-644-5 [Refereed Conference Paper]
10. Yeom, S., Choi-Lundberg, D., Fluck, A., Sale, A: User Acceptance of a Haptic Interface for Learning Anatomy. In: Proceedings of the International Conference e-Learning 2013, 23-25 July 2013, Prague, Czech Republic, pp. 239-246 (2013). ISBN 978-9-72893-988-5 [Refereed Conference Paper]
11. Stefan, P., Wucherer, P., Oyamada, Y., Ma, M., Schoch, A., Kanegae, M., Shimizu, N., Kodera, T., Cahier, S., Weigl, M., Sugimoto, M., Fallavollita, P., Saito, H., Navab, N.: An AR edutainment system supporting bone anatomy learning.VR2014, pp.113–114
12. Fallavollita, P., Blum, T., Eck, U., Sandor, C., Weidert, S., Waschke, J., Navab, N.: Kinect for interactive AR anatomy learning. In: Proceedings of 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 277–278 (2013). doi:10.1109/ISMAR.2013.6671803
13. Samosky, J.T., Nelson, D.A., Wang, B., Bregman, R., Hosmer, A., Mikulis, B., Weaver, R.: BodyExplorerAR: enhancing a mannequin medical simulator with sensing and projective augmented reality for exploring dynamic anatomy and physiology. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction, pp. 263–270, February 2012. ACM
14. Park, J.S., Chung, M.S., Hwang, S.B., Shin, B.S., Park, H.S.: Visible Korean Human: its techniques and applications. Clinical Anatomy **19**(3), 216–224 (2006)
15. Bacca, J., et al.: Augmented reality trends in education: a systematic review of research and applications. Journal of Educational Technology & Society **17**(4), 133–149 (2014)

Designing Schedulers for Hard Real-Time Tasks

Vasudevan Janarthanan

Abstract While synthesizing real-time schedulers on single processor systems using priority-based supervisory control of timed discrete-event systems (TDES), we came across the problem of state space explosion. In order to over-come it, a modified form of symbolic modeling methodology along with the pre-stable algorithm has been utilized in this work. The main contribution through this paper has been the development of an informal procedure for uniprocessor scheduler design with reduced state space for hard real-time tasks.

Keywords Real-Time tasks · Real-Time schedulers · Supervisory control · Timed Discrete-Event systems · Uniprocessor systems

1 Introduction

Real-time systems are a form of control systems having constraints on the execution time of their tasks. These constraints are expressed as real-time constraints. A real-time constraint can be defined as a condition on the timing of enabling, firing, initiation and termination of system events. A real-time constraint can be expressed as a Boolean condition on the values of clock variables, whose values increase with time. A real-time constraint of a task could be either the specification of its deadline or its complete execution. A real-time system can be designed as a set of tasks that can be differentiated based on their timing requirements as hard real-time, soft real-time and non-real time tasks. A hard real-time task is defined as one whose timely and logically correct execution is considered crucial for the normal operation of the entire system.

Real-time scheduling is defined as assigning the exact execution times for a set of real-time tasks such that all the temporal constraints are satisfied. In a real-time system, the purpose of the scheduling algorithm is to determine the sequence of execution of the real-time tasks, thereby ensuring their adherence to resource and timing constraints. While designing a real-time system, the choice of an appropriate scheduling algorithm or policy depends on factors like task synchronization methods,

Vasudevan Janarthanan(✉)
Department of Information Technology, Fairleigh Dickinson University, Vancouver, Canada
e-mail: v_janart@fdu.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_97

number of processors available in the system and the priorities of the tasks. The crucial aspect of real-time scheduling is to make sure that these tasks satisfy their temporal constraints and that the overall system performs correctly according to its specification.

Given this fact, a framework for designing such schedulers for hard real-time tasks upon Uniprocessor systems based on Supervisory Control Theory (SCT) for timed discrete-event systems was proposed in [7, 9]. In [7, 9], it has been shown that SCT of timed discrete-event systems could be applied to the scheduling of hard real-time systems. In particular, a formal framework was presented for the synthesis of real-time schedulers on uniprocessor systems using priority-based supervisory control of TDES. The execution of a set of tasks was modeled as a Discrete Timed Automaton (DTA). Then the supremal controllable sublanguage [4, 12] of timing constraints with respect to task DTA subject to a priority relation was computed to find all executions in which no deadline was missed.

Since discrete time is represented explicitly, model size for each task is considerably large and proportional to its period. The complexity in the synthesis of a scheduler using supervisory control [12] stems from the fact that, with the synchronous product, the number of states of a composite TDES increases exponentially with number of real-time tasks. As far as complexity in both time and space is concerned, the procedure followed namely TDES3 = sync (TDES1, TDES2), which computes the synchronous product of two TDES, has complexity proportional to the product of the state sizes of the two machines. Theoretically, the number of states of the synchronous product of two TDES is less than or equal to the product of their number of states. But in reality, it is often much less than their product for a nontrivial system. Therefore, we often need to allocate much more space than is actually required to store the result. The approach to confine the state explosion problem, which arouse while designing schedulers for uniprocessor and multiprocessor systems, relies on the symbolic representation of sets of states, which in essence computes the set that satisfies a formula as a fixpoint of a functional on state predicates.

The paper is organized as follows. Section 2 describes previous work from literature. Section 3 introduces basic concepts about timed automata and region graph. In Section 4, the modeling of real-time systems and specifications is presented based on timed automata for pre-emptive condition. Finally Section 5 provides the conclusion.

2 Previous Work

The authors in [3] have proposed an algorithm combining the symbolic and on-the-fly approaches. Their algorithm performs an on-the-fly exploration of the symbolic graph. They name the resulting graph as simulation graph, and they show in their paper that simulation graphs are much smaller than region graphs. They also perform model-checking for a temporal logic formula using simulation graph in [3]. The algorithm proposed in [3] has been incorporated in this work.

In [5], the authors have employed on-the-fly and space-efficient model-checking methods to solve the state explosion problem. Using the first method on a real-time program, they explore only the regions needed for checking the satisfaction of a

specification. The second method is used to store only necessary and minimal information in the memory. But the main drawback in [5] is that in the worst case situation, the on-the-fly method explores the entire region graph.

In [6], the authors illustrated the working of a symbolic model checking algorithm that works on a quotient of the region graph that depends on the formula being checked. They have shown how a symbolic fixpoint approach can be used to test if a guarded-command real-time task is non-zero and, if not, how it can be converted into an equivalent non-zero task. This algorithm, however, fails to terminate if no such quotient exists. Also, this algorithm requires that the explicit representation of discrete structure of the automaton be constructed a priori.

In [10], the author has introduced an approximation scheme that further reduces the complexity related to timing. The work in that paper is based on the observation that not all the timing information in the description of a timed system is usually needed to guarantee the satisfaction of a given property. This approximation scheme has been used in the model-checker RT-Cospan [11]. In [11], the underlying untimed description of the system is composed with an automaton representing the time bounds, and only the bounds that are necessary to verify the given property are introduced in the composition.

3 Basic Concepts

3.1 Timed Automata

Timed automata [1] use finitely many real-valued clocks to keep track of timing constraints, and serve as the common way of modeling the behaviour of real-time tasks. A type of timed automata called Discrete Timed Automaton (DTA) [2], which is used to model real-time systems and to specify requirements on their behavior, including timing constraints, has been used in this work. Time is assumed to be discrete (as opposed to dense [2]).

A DTA is a four-tuple L: = $(X, x_0, \sum^t, \delta)$, where X is the state set, x_0 is the initial state, $\sum^t$ is an alphabet, and $\delta: X \times \sum^t \rightarrow X$ is a partial transition function. It is assumed that $\sum^t := \sum \cup \{t\}$, where $\sum$ is the alphabet of system events and "t" is a special symbol which denotes the passage of one unit of time, or one tick of the global digital clock. Tickcount is denoted as a function over the set of all strings that returns the number of ticks in a string:

$$\text{tickcount}: \sum^{t*} \rightarrow N : s \rightarrow \text{no. of t's in s}$$

Time is measured in discrete steps; if $x = \delta(x_0, s)$ and tickcount(s) = $n \in N$ then at state "x" discrete time is equal to "n" while the value of real time could be anywhere in the real interval [n, n + 1). As before, it is assumed that the system events are instantaneous. For a DTA to model a real-time system, time always has to have a chance to advance. DTA is said to satisfy the time-progress property if:

$$\forall x \in X. \exists u \in \sum^*. \delta(x, ut)!$$

In this work, the attention is restricted to the class of DTA that satisfies the time-progress property. Let L be a prefix-closed language over $\sum^{t}$. L is called as a timed language if:

$$\forall s \in L.\exists u \in \sum^{*}. sut \in L$$

3.2 Region Graph

Definition: Let TA be a symbolic timed automaton, S be a finite set of states, C be a finite set of timer variables with |C| =: N, and T_c be the set of timing constraints over C. Let k_i be the (largest) constant to which $c_i \in C$ is compared, i = 1, 2, … , N. The region graph of TA, denoted by RG(TA), is a finite graph, in which nodes are regions given as r = (s, $\underline{v}$), where $s \in S$ and $\underline{v} \in \prod_{i=1}^{N}[0, k_i]$ is the vector of values of variables in C, and edges are transitions between the nodes which denote the passage of one unit of time (a tick of the global clock).

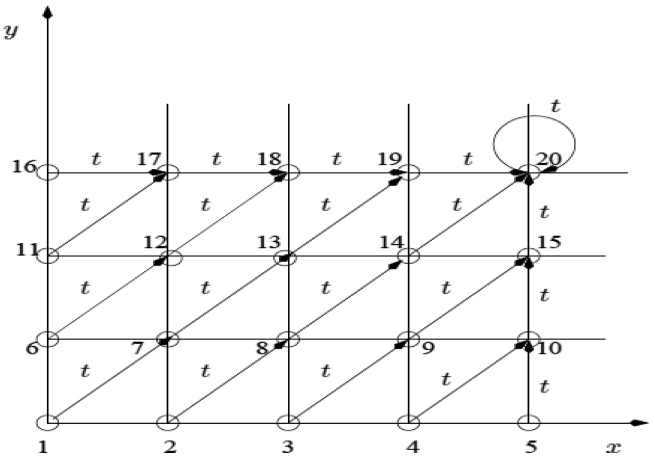


Fig. 1 Region graph of task T

The timer variables of task T, namely, execution time and period denoted by x and y, form the axes in the region graph. The different regions in the graph denote the values of timer variables in the symbolic DTA. For example, region 6 denotes x = 0 and y = 1, meaning that with one time unit has already gone by, T has been executed for one time unit. As Fig. 1 indicates, a tick of the global clock moves the graph from region 6 to region 12. As another example, in region 16, the timer variable y has reached its limit, and thus subsequent ticks do not increment the value of y.

4 Task Modeling

Fig. 2 illustrates the automaton of real-time tasks T_1 and T_2 Event "s" is used when a task gets preempted while its execution is not yet complete. Event "s" is called

a suspend event. When event "s" occurs, the current execution value of the task is stored in a variable. Variable "z" is used to store the value of preempted task's execution time; in other words, it records how much of the task has been executed thus far. Event "r" is used when a preempted task resumes its execution and is referred to as a resume event. When event "r" occurs, the stored value of the task's execution is retrieved from "z" and the task continues its execution until its completion.

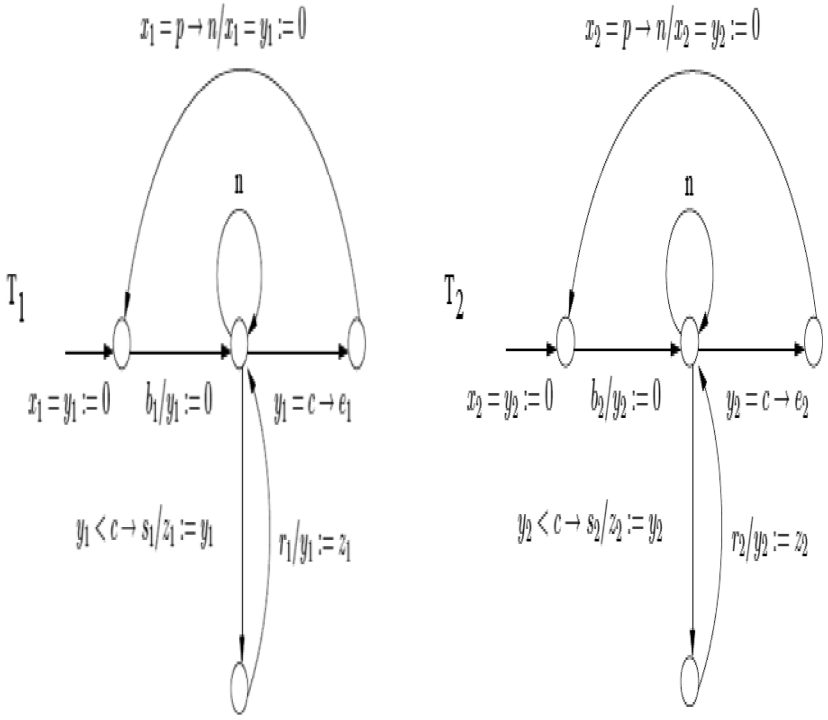


Fig. 2 Automaton of real-time tasks T_1 and T_2

The specification automaton for preemptive tasks is given in Fig. 3. It states that when the execution of a task is started (b_i) or has just been resumed (r_i), the other task cannot begin its execution until after the execution of the first task is complete (e_i) or is suspended (s_i).

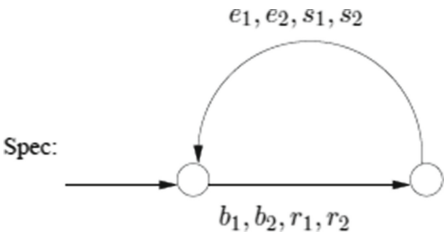


Fig. 3 Specification automaton for preemptive tasks

The automaton that could be used to design schedulers for tasks T_1 and T_2 with reduced state space is obtained by taking the synchronous product of specification with the composed task model, which is depicted in Fig. 4. TS: = meet (T, S), where T is composed automaton model and S is specification for task's preemption.

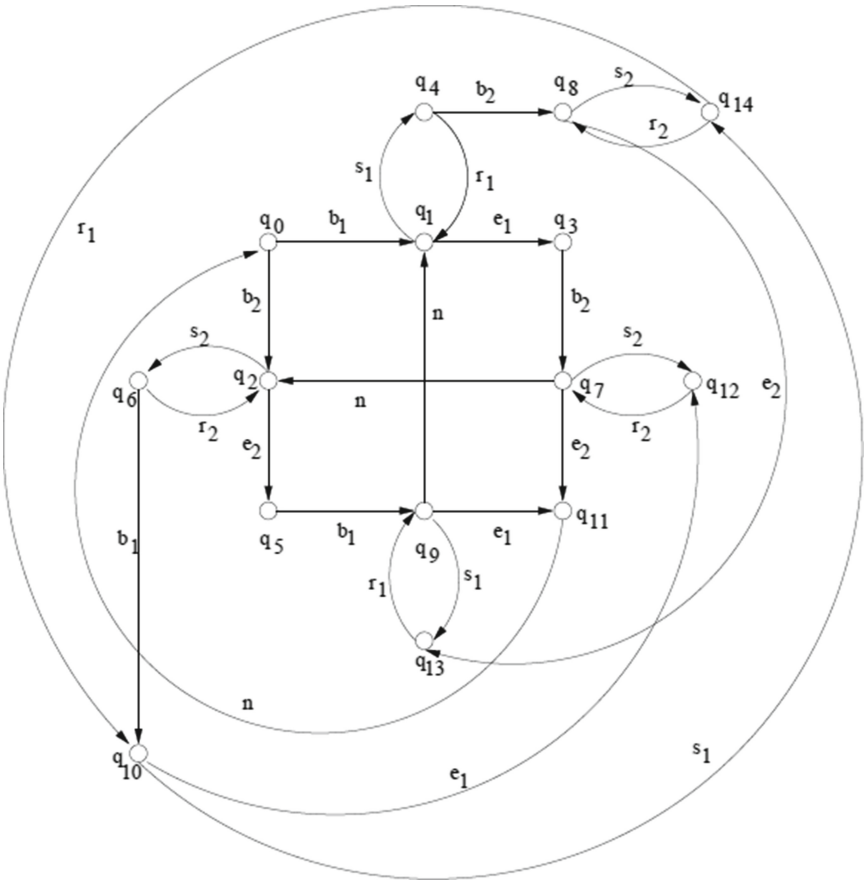


Fig. 4 Automaton used to design schedulers

In Fig. 5, the symbolic graph of composed model of the two tasks under preemptive assumption is enumerated. It consists of the various states as in Fig. 4 along with the values of the timer variables for period and execution time for the two tasks.

For example, $q_0$0000 means that at state q_0, the timer variables of tasks T_1 and T_2, namely, x_1, y_1, x_2 and y_2 respectively, are equal to zero i.e. they have just been reset. The symbolic graph is then obtained by considering the different transitions that could be traversed based on the various events of Fig. 4 while satisfying the timing constraints. The region sets in Fig. 5 denoted by R_0, R_1, etc., have nodes with time (portrayed by the tick event) passing explicitly between them. One can observe that

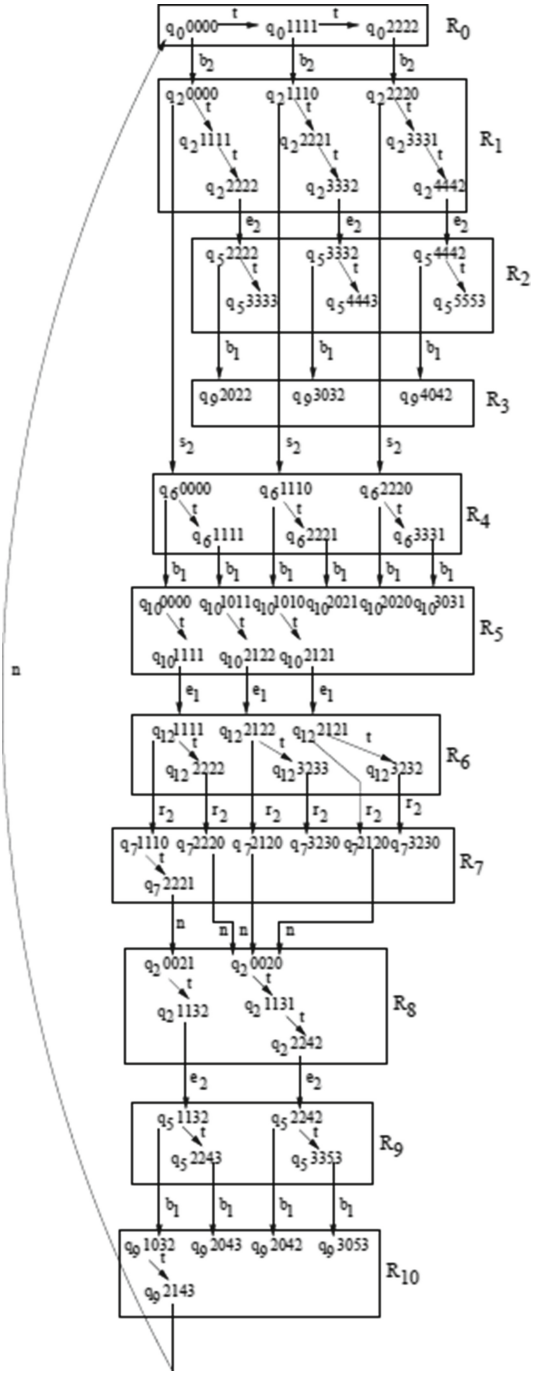


Fig. 5 Symbolic graph of composed model

transitions from some of the nodes inside these region sets are not portrayed beyond a certain point. For example, consider node $q_9 2022$ in region R_3 of the figure. In that node, the timer variables for task T_1, namely, period (x_1) and execution time (y_1) have values 2 and zero, respectively, while the timer variables of task T_2, namely, x_2 and y_2 have values 2 and 2, respectively (meaning that task T_2 had been executed for two units of time).

From that node, no more transitions have been illustrated because task T_1 had already missed its deadline, meaning that the path traversed from that node would not provide us with a feasible scheduler. Also, in Fig. 5, from region R_5, event s_1 could also be executed but has not been depicted in the figure. This is because, suspending T_1 would obviously lead to the execution of task T_2, thereby missing its deadline.

Next, the simulation graph is obtained by enumerating nodes as region sets where only task transitions are explicit, while time passes implicitly inside nodes as shown in Fig. 6.

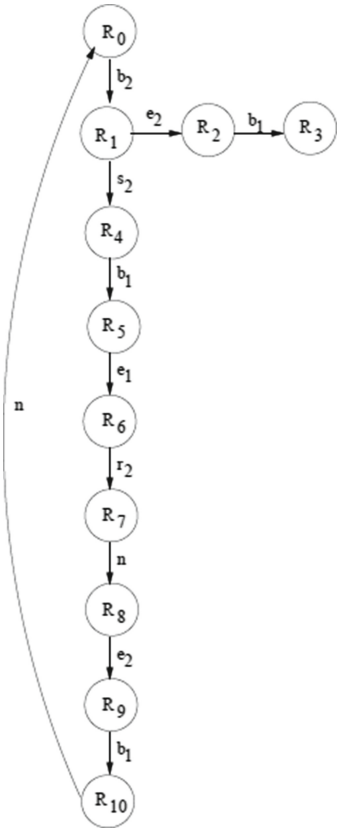


Fig. 6 Simulation graph

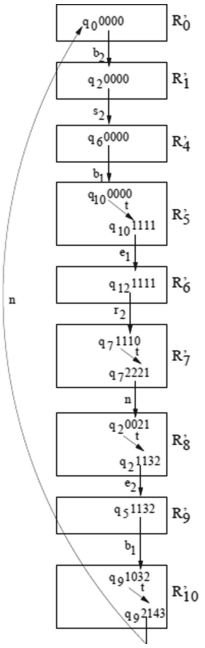


Fig. 7 Pre-stable symbolic graph

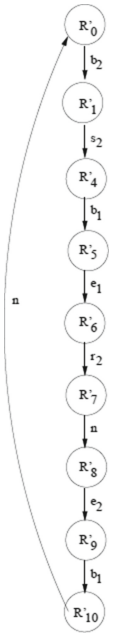


Fig. 8 Pre-stable simulation graph

The pre-stable symbolic graph of Fig. 7 is a scheduler for the two tasks with a reduced state space. The region sets obtained in the pre-stable symbolic graph are subsets of the corresponding region sets in the symbolic graph. The corresponding pre-stable simulation graph (shown in Fig. 8) is obtained by removing explicit time transitions and also representing region sets by nodes.

5 Conclusion

In this paper a modified form of symbolic graph called simulation graph has been utilized and the pre-stable condition [3] has been applied on it. The scheduler design procedure has been extended under the preemptive assumption. On comparison with the universal scheduler obtained in [7, 9], it is found that the pre-stable symbolic scheduler obtained in this work to be far smaller in size, thereby handling the state space explosion problem that occurred in [7, 9].

References

1. Alur, R., Dill, D.: A theory of timed automata. Theoretical Comp. Science, 183–235, (1994)
2. Alur, R., Dill, D.L.: Automata for modeling real-time systems. Lecture Notes in Computer Science, pp. 322–335. Springer (1990)
3. Bouajjani, A., Tripakis, S., Yovine, S.: On-the-fly symbolic model checking for real-time systems. In: IEEE Real-Time Symposium, pp. 25–34 (1997)
4. Brandin, B., Wonham, W.M.: Supervisory Control of Timed Discrete-Event Systems. IEEE Transactions on Automatic Control, 329–342 (1994)
5. Henzinger, T., Kupferman, O., Vardi, M.: A Space-Efficient On-the-Fly Algorithm for Real-Time Model- Checking. In: CONCUR 1996 (1996)
6. Henzinger, T., Nicollin, X., Sifakis, J., Yovine, S.: Symbolic model checking for real-time systems. Information and Comp., 193–244 (1994)
7. Janarthanan, V., Gohari, P.: Universal Scheduler Design on Uniprocessors in Supervisory Control of Discrete-Event Systems Framework. In: IEEE Conf. on Control Appl., pp. 916–921 (2005)
8. Janarthanan, V., Gohari, P.: Supervisory Control-Based Design of Real-Time Schedulers for Uniform Multiprocessor Systems. In: IASTED Intl. Conference on Intelligent Systems and Control, pp. 488–493 (2005)
9. Janarthanan, V., Gohari, P., Saffar, A.: Formalizing Real-time Scheduling Using Priority-Based Supervisory Control of Discrete-Event Systems. IEEE Transactions on Automatic Control, 1053–1058 (2006)
10. Balarin, F.: Approximate reachability analysis of timed automata. In: RTSS, pp. 52–61 (1996)
11. Alur, R., Kurshan, R.P.: Timing analysis in COSPAN. LNCS, vol. 1066, pp. 220–231 (1995)
12. Wonham, W.M.: Supervisory control of discrete-event systems. Professor Wonham's page on University of Toronto

An Autonomous Stair Climbing Algorithm with EZ-Robots

Jason Moix, Sheikh Faal, M.K. Shamburger, Chris Carney, Alex Williams, Zixin Ye and Yu Sun

Abstract Every day we witness greater advances in robotic technologies appearing around us. From the creation of military robots built to travel across hostile territories to the implementation of drones in a domestic setting by Amazon.com, these robots continue to become a more important part of our lives every day. Due to this fact, the importance of investigating new methods for ensuring the safe travel from one location to another continues to grow as well. In the spirit of researching these methods, this research aims to develop the functionality for a robot to climb a set of stairs. The robot being used for this project is manufactured by EZ Robot and is provided with its functionality through the execution of scripts and animations created in an integrated development environment. Using this development environment, we developed the autonomous climbing algorithm to allow the robot to successfully climb a set of stairs.

Keywords Algorithmic traversal · Stair climbing · SIX robot

1 Introduction

The world is changing and the development of robots to complete everyday tasks is increasing every day. From providing the functionality to cook hamburgers to providing the functionality to put complex machines together, the possibilities for routes of development are endless. After the disaster at Fukushima in 2015 where a tsunami interrupted the power supply to the cooling systems of three nuclear reactors resulting in a nuclear meltdown, the world found itself in need of a way to

J. Moix · S. Faal · M.K. Shamburger · C. Carney · A. Williams · Y. Sun(✉)
Computer Science Department, University of Central Arkansas, Conway, AR, USA
e-mail: yusun@uca.edu

Z. Ye
Central High School, Little Rock, AR, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_98

get to areas where no human can. After the disaster the area was filled with toxic radiation. With no way to get to ground zero, we were left with no clue as to the current condition of the reactor nor as to if radiation was constantly being leaked into the ecosystem. Robots seemed to be the answer, and the world flocked to the development of the functionality to allow humanoid robots to traverse varying terrains in disaster areas in order to not only survey the area but also perform repairs, perform rescue operations, transport supplies, and perform any other task a human on the scene would be able to do.

Since the eruption of this research field onto the mainstream, several companies have developed their own robots for the task of traversing varying terrains. Boston Dynamics, a major robotics lab located in the United States, has developed their humanoid robot, Atlas. He is able to negotiate rough terrain with 28 degrees of freedom that allow him to avoid obstacles, recognize dangerous areas on which not to step, and climb up steep using his hands and his feet [1]. Furthermore, Japan has produced constant progress in their work with the robot Asimo since 1980, long before the research in this field became so popular. Not only can Asimo run and negotiate varying terrains, he can also recognize speech patterns, communicate with humans, recognize postures and emotions [2, 3].

The University of California Berkeley has also joined the race to drive the advancement of robots that are created to be faster, stable, and more maneuverable than anything we have worked with to date. Unlike Atlas and Asimo, however, the robots that they have been working with have proved much less complex than the robots other companies have worked with. Ariel, for example, is designed with a body much like that of a crab. The simplified body design allows the body crawl over rocks, through ridges, and even across surfaces underwater. It is also able to sense its body's current orientation. For example, waves strike Ariel from the side or perhaps Ariel tumbles down a ridge. Because it can detect its current orientation, it can reorient its leg in order to ensure the restoration of its balance on land. At the moment, Ariel has been outfitted with sensors and AI capabilities that allow it to travel through surf zones, locating dangerous mines [4].

Another one of UC Berkeley's robot, Rhex, is one of the most maneuverable robots built to date [5]. Rhex, about the size of a shoebox, has six legs located along a horizontal plane, much like the body of a cockroach. Its six springy legs allow it to travel across varying terrains while employing reflex technology that allow it to adapt to sudden changes in uneven terrain. The makers have described it as resembling a hyperactive terrier as it travels [6].

A revolutionary robot at UC Berkeley in conjunction with Stanford is Sprawl. Sprawl is revolutionary because his entire body is made up of a single piece [7]. They manufacture each sprawl body with specific indicators of strength, durability, and flexibility in order to ensure the robot's adaptability under certain conditions. Its legs are piston driven and cause a kind of bouncing movement over objects like that of a grasshopper. This bouncing movement, combined with the customized, reactive style that the body was manufactured with, allow it to travel over varying terrains without the need to sense the state of the environment around it.

The success that has occurred as a result of these companies's never-ending effort is inspiring, and it has driven us towards our research into how to allow the EZ robot, "SIX", to climb a set of stairs. There are many obstacles and pitfalls standing between us and the top of the stairs, however, it is our belief that SIX should climb!

The rest of this paper is organized as follows. Section 2 introduces system overview and description. Section 3 presents our proposed algorithm and methods. Section 4 discusses the performance evaluation. Section 5 concludes the paper and discusses future work.

2 System Overview and Description

Our robot shown in Fig. 1, "SIX", is a hexapod robot. That means it has six legs which gives it more stability and freedom to move around. It needs only three legs to balance while standing and it can walk using just four legs. SIX is assembled with EZ-bits which makes it very easy to customize. All of its joints are powered by twelve heavy duty servo motors, allowing a wide range of different movements. The removable EZ-bit joints and top mounted camera enhance SIX's ability to interact with the environment around it. The camera features advanced vision learning, detection, and tracking abilities. SIX can interact with different objects, colors, faces, motion, and QR codes as well.

Fig. 1 "SIX" Robot.

The EZ-builder software makes it very easy to make SIX do different movements and interactions. Within the EZ-builder software there is EZ-script. It lets you make your own code to make SIX do different things instead of using the preloaded motions. This is what we used for this project. You can add behaviors such as camera tracking, speech recognition, GPS, and more. The Auto Position in the EZ-builder software is used to animate SIX. You can make SIX do certain things when it sees a face or hears a command. You can even control it from your phone with the EZ-builder app.

The brain of SIX makes this all possible. It is the EZ-B v4. The EZ-B v4 boasts two 32-bit CPUs at 200mhz (ARM Cortex M3 & Microchip PIC32), Energy Efficient Digital Switching Power Supply, Wi-Fi Connectivity with security, Embedded Web Server, Amplified Digital Audio with Speaker, 3 x i2c ports, 3 x UARTs, 24 x Multi-Function I/O, 73 x Servos (Dynamixel & PWM), 8 Analog

ports, and Integrated Video. This piece of hardware makes it possible for SIX to do what want it to do [8].

While all of these features combine to make a great product, there are some pitfalls in the design. The camera is a narrow lens receiver without the ability to pan up or down. Because of this, successfully measuring a stair when the robot is standing in front of it is very difficult. Imagine walking while wearing a visor under your eyes, without the ability to see where your feet are stepping. Now imagine wearing leg braces! SIX boasts 12 degrees of movement, however, we found that we really needed more in order to develop an optimal climbing algorithm. In order to counter this problem, we built our own set of steps.

In designing the steps that SIX was to climb we not only had to take height into account, but the step surface dimensions as well. A regular step wouldn't do because SIX has stance that spans a circle with a diameter of about 13 inches. Of course it may have been possible for him to climb the steps while maintaining a diagonal stance on the edges of his current and next step, but this runs the risk of SIX suddenly tumbling down the steps without warning. This risk is a result of SIX's inability to measure his body's current orientation. With gyroscopic sensors, this implementation may have been a little more realistic. However, since we found our self without those advanced sensors, each step was designed to have a surface area of 16x16 inches, thus, allowing SIX to stand on each step with room to budge. The height of each step was decided based on the overall height of SIX's reach. When his claw is placed at the 1 degree mark it reaches to a height of about 3.5 inches. For this reason, we made each step 3 inches high.

While designing the steps, we also had to take into account how able SIX was to maintain a grip on an object at any given time. His claws are relatively stationary, like the shins of someone wearing leg braces. Furthermore, his body is made of hard slick plastic. This fact was certainly less than optimal for accomplishing our task of allowing SIX to climb. Therefore, we began to investigate methods that may allow SIX to get a better grip on the ground throughout the climb. Our conclusion was to use a Yoga mat to pad the surface of each step. Yoga mats are designed to hold grip on bodies undergoing strain in opposite direction. For this reason, they were the perfect solution to our problem of grip. Fig. 2 shows the picture of our developed steps.

Fig. 2 The developed steps.

Even though these stairs managed to solve some of our problems involving the stair climb, we still had many issues to face. For example, now that we could no longer trust our camera for step measurement, how were we going to implement this functionality? The answer came in the form of an ultrasonic distance sensor. While this sensor is not a factory part of SIX's body, it proved invaluable in the implementation of our algorithm. However, as a consequence of it not being a factory part, there was no real place to mount it. Therefore, we were forced to fashion our own ultrasonic distance sensor mount out of cardboard and painters tape. This process took a way a good chunk of our development time, since faulty reading resulted in a constantly malfunctioning algorithm. However, with a lot of tweaking and a lot of tape we were able to successfully find a place to mount the sensor. Unfortunately, not having as many degrees of freedom as we needed to accomplish the climb still plagued us. This problem and a few other suggested features are discussed in the performance evaluation section.

3 Algorithms and Methods

Because of our inability to successfully implement the camera in our stair detection algorithm and because of the camera's narrow lens and inability to pan up or down, SIX makes use of an ultrasonic distance sensor to detect obstructions by reading sound waves. This process, shown in Fig. 3, involves sending a pulse of sound from a small speaker located in the sensor. Then, the sensor reads the sounds using an onboard microphone as they bounce back in order to calculate the distance to obstructions, just like echo-location.

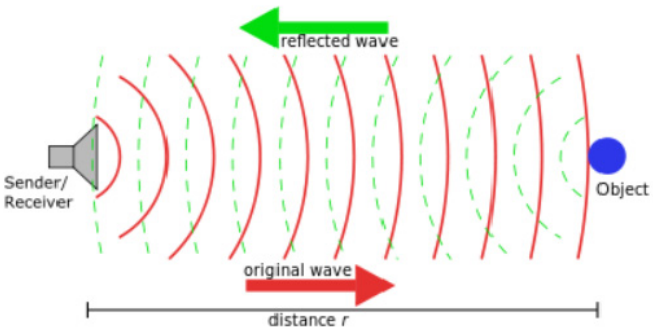


Fig. 3 Ultrasonic Distance Sensor Operation [9].

Once we were able to accurately detect obstructions in front of SIX, we were able to implement our algorithm. The algorithm is designed to work on its own, absent of human control. Fig. 4 illustrates the overview of our proposed step algorithm.

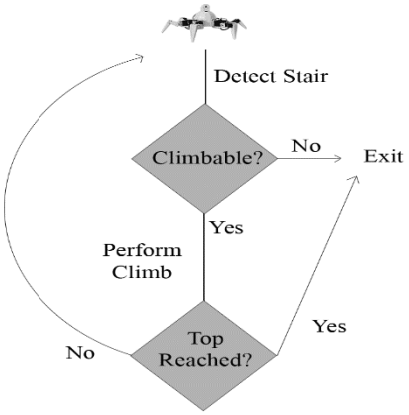


Fig. 4 The Diagram of the Proposed Climb Algorithm.

The step by step description of our proposed algorithm are presented below:

1. SIX walks forward, keeping track of its distance from obstructions at 800ms intervals. If the current distance equals the new distance, then SIX has been unable to move. Thus, SIX has encountered his first obstruction.
2. Next, SIX checks the distance value of the last gathered data in order to determine if the step is climbable. Since the distance sensor is designed to peak over a climbable step, if the distance value is less than 10, the obstruction appears too close. Therefore, the stair is not climbable. If the value is more than 10, the sensor is successfully peaking over the stair, thus making it climbable.
3. The SIX performs a climb action that was developed in the EZ Builder IDE. This action was very difficult to implement and most likely took the most time. While SIX's legs are very moveable, they aren't really designed for performing the movements required to climb a set of steps. With a little work, however, we were able to successfully develop and action that works some of the time, and implemented the functionality to play a little music to liven things up.
4. After the step has been climbed, SIX again checks the distance to the next obstruction. If the distance is greater than 20, the sensor doesn't detect any further steps and a celebration action can be executed. However, if there is another step, the algorithm loops back to step 1 until the top is found.
5. Finally, once the top has been found, SIX can finally celebrate. He performs an action to turn himself around and enters an "Action Pose." Then, the celebration music is played, marking the end of our script run.

While this algorithm often results in SIX reaching the top of a stair, many things can go wrong for many reasons. These things will now be covered in the performance evaluation section.

4 Performance Evaluation

After a long and arduous development process, we have successfully implemented the above proposed algorithm. Our experimental results have demonstrated that SIX can climb a set of stairs. However, the climb is not always successful for several reasons due to:

- The legs of SIX don't really have enough degrees of freedom to make a good climbing algorithm. While we did our best with what we had, depending on weight distribution and contour, SIX may not make it up a lot of steps.
- Speaking of weight distribution, SIX has no way of detecting his. Imagine walking up a set of stairs without the ability to feel your legs and the action you are performing. Because of this, fine tweaking the algorithm was necessary in order to allow for the fewest possible variations in weight distributions. While this was accomplished after many iterations, SIX will often land a little sideways on the stair, thus affecting the next iteration of our algorithm.
- Friction also plays a big part in a successful climb. You aren't going to be as successful climbing a hill of ice as you may be one of dirt. SIX's body is designed much like the ice, in that it is very slick. On most surfaces, he has trouble gaining traction just to walk forward. Taking climbing a stair into account, the differential in friction on varying surfaces gave our algorithm a big hit.
- SIX has no default location to place his distance sensor. As a result, we had to fashion a makeshift mount and tape it his body. Because of its unstable placement and the amount of movement that occurs through the execution of a step forward, distance readings can vary greatly. This fact heavily impacts the performance of our algorithm.
- More sensors! In order to accurately measure a step and SIX's placement on the step, we definitely need more sensors. With the current set-up SIX has no way to determine exactly where on the step that he is nor if he is parallel to the next step to be climbed. Because of this fact, SIX is always on the verge of attempting to climb right off the side of a step.

While all of these problems plagued our development process, we drove through. We managed to design an algorithm that allowed SIX to successfully climb the steps most of the time, a task we see as a great success.

5 Future work

Now that we have developed the algorithm for SIX to climb a simple set of stairs, we can look into all the other functions that we can implement as a result of this

research. For this project, we focused on building an algorithm for a predetermined, custom built set of stairs. The question remains, however, with what frequency to you come across a situation where things are exactly what you expect them to be. The answer for most people would be not often at all. For this reason, it is important to build on our basic algorithm to allow for variation in step size, angle, number, and surface. At the moment, as covered in the section on performance and evaluation, we are lacking both the equipment and sensor capabilities necessary in order to successfully implement such a flexible algorithm.

Pressure sensors beneath each of the feet would allow SIX to monitor his distribution of body weight. This is important because it is necessary to modulate our body weight distribution when on surfaces that are found to be uneasy or yielding, such as sand or snow. Picture a runner making his way across a beach. The deliberate strides that runner makes while measuring how hard to let his feet come into contact with the ground optimize the distance that runner can travel with every step. Now picture that same runner going across the beach with no attention paid to those kinds of considerations. The runner's feet would hit the sand a predetermined angle with a predetermined force. Sure, every squirrel will get his nut and the runner may make a perfect stride at times on the surface he is traveling across. However, statistics predicts that many of his steps may merely graze the sand gaining no kinetic friction against the grains of sand. Other steps may strike too deep, forcing the runner's foot to become engulfed by sand or completely throwing the runner off balance in the absence of sand that yields. Under both of these conditions, the runner will make menial progress with each step he takes all while risking injury to his body. The ability to detect pressure is what allows us to avoid such catastrophe.

Gyroscopic sensors would work in pair with the pressure sensors in order to ensure that SIX kept his footing as he climbed varying terrains. Going back to the example of the runner on the beach losing balance, imagine each foot hitting the sand expecting the give with each step. Then imagine what would happen if one foot suddenly landed on a flat rock without the runner seeing it coming. Without quickly adjusting to allow for the great increase in leverage provided by his foot pushing off a hard, flat surface instead of sand, the runner would push himself up and to the side with great force. The result of this action would be a complete loss of balance and ultimately a fall to the ground. However, runners have the ability to monitor their equilibrium. If they are tipping to the point of falling over they can easily realize that they are about to fall and employ countermeasures in order to prevent such a result. This functionality is provided through gyroscopic sensor and would be invaluable to SIX as he climbed varying terrains. For example, he could catch himself should one foot step off of the stair and his body started to tilt. This act would greatly reduce the risk of harm to the equipment.

Once these sensors are added to SIX's configuration, research could spread to many different avenues. One involves the optimization of pressure that SIX applies to each step when the surface yields with a varying degree. This research would involve both the pressure sensors and gyroscopic sensors. Pressure sensors

would be able to detect how much pressure was currently being forced on a surface while the gyroscopic sensors would be able to detect how much the surface yielded as a result of this pressure. By developing a function between the two and optimizing the pressure to surface yield ratio you can optimize the rate at which SIX travels and the success of a particular climb.

Once the research on optimizing the pressure to surface yield ratio is completed, the algorithmic development can finally begin on how to build the functionality for SIX to travel across many different surfaces, from steps to hills to dunes. This research can then be applied to not only SIX, but every other robotic drone that uses land traversal as its main mode of transportation.

6 Conclusion

This project was definitely more difficult than we thought it would be, mostly due to the problems we faced during the development process as a result of how SIX is built. We countered these problems by building our own set of steps that were customized to allow for SIX's large body size, lack of equilibrium detection, and small reach. We also built our own custom mounts for the ultrasonic distance sensors that pointed the sensor exactly where we needed it to point. Then we went through the painstaking process of designing each deliberate move that SIX should make as he climbs a set of stairs. As a result, we successfully managed to develop an algorithm to allow SIX to climb a set of steps and perform a celebration dance at the top. However, there is a lot of room for future work. The degrees of freedom of motion that we have to work with in regard to SIX's body are pretty much set, unless we can add a third set of servos which may not be realistic at all. However, with the implementation of more sensors and the modification of the surface of SIX's legs and feet, I believe we can design a much more efficient climbing algorithm that would allow SIX to climb many other types of objects. Then, research can spread to what SIX is really capable of and exploration into of those functions.

Regarding future work, we will continue to implement different surface materials onto the body of SIX. At the moment, his shell is much too slick to allow for any effort into developing algorithms that provide with the kind of consistency that would be necessary for an algorithm of this type. At the very least, some sort of foot pads should be implemented to counteract the lack of friction found as SIX traverses across varying landscapes.

Acknowledgment This work was partially supported by NASA EPSCoR Grant 2012 (No. NNX13AD32A).

References

1. Feng, S., Whitman, E., Xinjilefu, X., Atkeson, C.G.: Optimization based full body control for the ATLAS robot. In: Proceedings of the 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Madrid, Spain, November 2014
2. Chestnutt, J., Lau, M., Cheung, K.M., Kuffner, J., Hodgins, J.K., Kanade, T.: Footstep planning for the honda ASIMO humanoid. In: Proceedings of IEEE International Conference on Robotics & Automation (ICRA), Barcelona, Spain, April 2005
3. Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J.K., Kiesler, S.: Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In: Proceedings of Human-Robot Interaction Conference (HRI 2006), March 2006
4. Full, B.: Unlocking the Secrets of Animal Locomotion, University of California, Berkeley, April 2015. http://www.berkeley.edu/news/media/releases/2002/09/rfull/home.html
5. Altendorfer, R., Komsuoglu, H., Buehler, M., et al.: RHex: A Biologically Inspired Hexapod Runner. Autonomous Robots **11**(3), 207–213 (2001). Kluwer Academic Publishers
6. Saranli, U., Buehler, M., Koditschek, D.E.: Rhex: A simple and highly mobile hexapod robot. International Journal of Robotics Research **20**(7), 616–631 (2001)
7. Zarrouk, D., Pullin, A., Kohut, N., Fearing, R.S.: STAR, a sprawl tuned autonomous robot. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, Germany, May 2013
8. EZ-Robot Manual, from EZ-Robot.com, April 2015
9. Wiora, G.: Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Sonar_Principle_EN.svg

Toward Indoor Autonomous Flight Using a Multi-rotor Vehicle

Connor Brooks, Christopher Goulet and Michael Galloway

Abstract The objective of the paper is to detail progress toward an indoor autonomous micro aerial vehicle (MAV) that uses a multi-layered hardware and software control architecture. We build a quadrotor MAV with design considerations for indoor use and document the design process. Our design uses a three-layered control system in both the hardware and software components. We focus on using a modular control architecture to allow for individual development of layers. For initial progress toward autonomous flight, we create an automated altitude control program that enables the MAV to hover stably at a given height. Experiments include hovering and directed flight through hallways with obstacles. Initial results provide optimistic feedback for future experiments concerning traversing indoor environments using an autonomous MAV with a three-layered control architecture.

Keywords Autonomous robotics · Indoor UAVs · MAV control

1 Introduction

Navigation and coverage algorithms provide a growing area of study, especially in autonomous robots. Autonomous robots and the coverage algorithms they use are more prevalent than ever, especially in military use [4], and are being seen in various consumer products such as automated indoor vacuum cleaners, automated lawnmowers, and personal drones. Aerial robots possess certain advantages over ground robots in dealing with rough ground terrain as they can simply fly over obstacles, but also face increased difficulties including mapping and navigation in three-dimensional environments, limited payloads, and constant motion [1].

The development of autonomous aerial vehicles has rapidly progressed in the last decade, and the miniaturization of existing technologies now allows for research

C. Brooks(✉) · C. Goulet · M. Galloway
Department of Computer Science, Western Kentucky University, Bowling Green, KY, USA
e-mail: connor.brooks499@topper.wku.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_99

into autonomous aerial vehicles that can operate indoors. These vehicles may be used for surveillance and monitoring, exploration, search and rescue, or military use. Generally, for applications involving small environments such as navigating an indoor space, micro aerial vehicles, or MAVs, are used.

For our research, the control structures being investigated are specialized for indoor navigation of a MAV. The indoor algorithms must rely solely on proximity sensor data for information about the surrounding environment, while typical analogous outdoor algorithms have the availability of GPS and barometric pressure sensors. The scope of the research will include design and construction of an indoor MAV, interfacing with the MAV, controlling the MAV using sensor data, and, finally, work towards autonomous flight and navigation. We introduce a multi-layered hardware and software control architecture for indoor MAV control that allows for abstraction of the individual layers from one another, thereby creating a platform for further research into various indoor MAV control algorithms and strategies.

Fig. 1 Top-down view of our MAV

2 Related Work

Designing a MAV with an embedded computer enables it to compute some needed algorithms mid-flight. With several research projects conducted on autonomous indoor MAVs in recent years, many different control architectures have been used that implement embedded computers. Due to computational limitations, several of these projects require some processing that is done on a base station off-board the MAV [1, 2, 3, 7]. Our project does not rely on any off-board processing for successful autonomous control of the vehicle, instead using distributed computing that is fully on-board the vehicle.

Other recent projects have also required fully on-board processing. Some projects that allow for all on-board processing divide computation between two components. Satler et al. [6] split their software architecture between two units - a low-level unit which monitors attitude and controls communication with the actuators,

and a high-level unit that handles the rest of the computing. Nieuwenhuisen et al. [5] also split computing in this way, and further modularize their software architecture by creating a layered approach within this high-level processor. However, neither of these projects separate the computation of all critical time-of-flight operations (including obstacle avoidance) onto separate hardware components from that which handles mapping, localization, and other computationally complex tasks or explicitly abstract the software components from one another. Tomic et al. [8] also employ a low-level and high-level unit conceptually, but do use a three-layer division of components between perception, cognition, and action layers. These layers do not correspond to particular on-board computers and are not explicitly abstracted from one another.

We contribute a multi-layered software and hardware architecture specialized for autonomous indoor MAV control that allows for the complete abstraction of layers from one another. This is done through the alignment of a three-layered distribution of computing with a three-layered distribution of required processing tasks, taking into consideration the requirements of an autonomous indoor MAV in the division of these tasks. This contribution allows for the individual development of each layer and the potential interchangeability of both the hardware and software components of a single layer without affecting the other layers.

3 Design of the Multi-rotor Vehicle

There are two general categories of MAVs: fixed-winged and rotary-wing. In the case of trying to build a MAV for indoor use, a rotary-wing design is an obvious choice as it offers more stability and maneuverability than its fixed-winged counterpart. We implement a quadrotor MAV as the basic framework for our MAV.

Size is also a factor in the design of the MAV. When considering indoor environments, the limiting access point is typically a door frame, as this is the smallest space through which an entity navigating the environment will have to fit. However, the MAV must still be large enough to hold all necessary sensors and computers. Taking that into account, we choose to design a MAV with a wingspan (calculated from the end of one rotor to the end of the opposing rotor) of 35 centimetres, giving it plenty of room to make it through typical door frames.

Using the size parameters and estimating the weight of components, we design a MAV with a focus on being as modular as possible. The outline of the design includes four brushless motors, four speed controllers, a 2800 mA/hr lithium-ion polymer battery, and a receiver with corresponding remote-controller. Three separate computers are also carried on board to allow for three separate hardware layers, consisting of a preprogrammed flight-control board, an Arduino MEGA, and a Raspberry Pi 2 Model B.

4 Three-Layered Control Architecture

4.1 Division of Tasks

With the completed MAV designed and built, we focus on moving the MAV towards autonomy. We recognize three distinct levels of tasks as divided by their complexity and required runtime, similar to the distinctions used by Achtelik et al. [1]. In our first group of tasks we include those tasks which require the highest rate of completion - stabilization of pitch and roll must occur at a rapid frequency to ensure stable flight. In the second group we include those tasks which must be completed in real-time and do not require complex computation, but are not as urgent or require more computation than those in the first group. This group contains altitude stabilization, local obstacle avoidance, and specific task execution (landing and take-off, for example). Finally, the last group contains all tasks which are highly complex or are not required to be real-time, including mapping and data recording.

Using this division, we create three computational layers that act as separate modules. A primary requirement of our project is to create a MAV with all computing done on-board so that flight capabilities do not depend on external computation whatsoever. To achieve this while also creating our modularized layers, we use distributed computing on-board in the creation of our autonomous control structure.

4.2 Division of Computing

For the lower layer hardware component, we use a preprogrammed KK Flight Control Board that monitors stabilization and processes all output to the actuators - the first group of tasks. An embedded gyroscope and accelerometer allow this board to detect the current pitch and roll of the vehicle, and it corrects these values toward its calibrated goal values by making fine adjustments to the actuator output through the use of a PID controller. Similar low-level units have been used by other projects [6]. Next, the middle layer computational component is an Arduino MEGA board. This layer is responsible for the second group of tasks. As knowledge of the local environment is required for these tasks, the Arduino MEGA communicates with range sensors to read in values while also applying algorithms to control the MAV's positioning. Finally, a Raspberry Pi 2 Model B is used for the upper layer, which is responsible for the third and final group of tasks. The upper layer sends its output in the form of flight tasks to be executed to the middle layer, which implements those instructions as part of its algorithms before sending thrust, pitch, roll, and yaw commands to the lower layer to be processed and sent to the actuators. Thus, the upper layer never directly communicates with the lower layer.

This three-layered architecture allows for the individual development of each computational layer. Layers can be tested and optimized individually, beginning with the lower layer and working up to the higher layer. Different control strategies can potentially be implemented within separate layers without adverse affects on the overall MAV's performance, provided each layer maintains sufficiently correct output. Additionally, separating the middle layer and upper layer functions allows for simple loop-based, reactive algorithms to be used on the middle layer, and ongoing complex algorithms to be used on the upper layer. This allows for clean and intuitive implementation of repetitive and fast processes, while more complex computations can be done in a completely separate module, negating any chance of causing flight failure because of a delay in actuator output that occurred due to an issue in a more complex computational task.

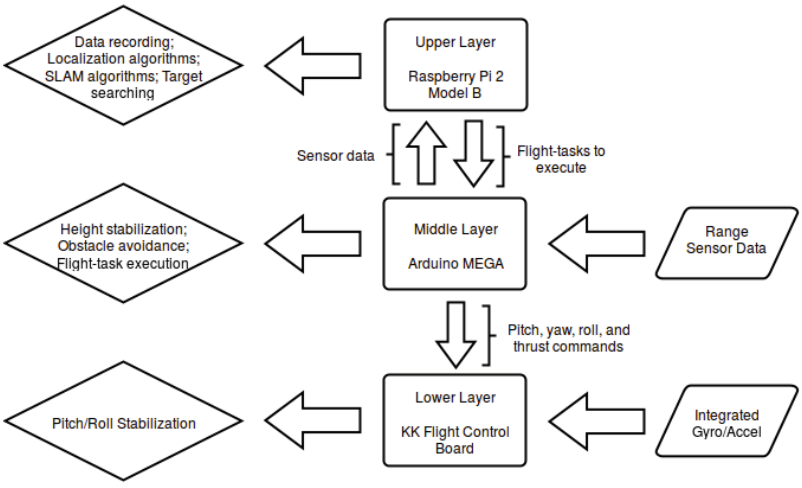


Fig. 2 Flowchart for three-layered control

5 Autonomous Flight Algorithms Implementation

5.1 Lower Layer Implementation

Since we use a preprogrammed flight control board in our implementation of this three-layered control structure, our lower layer was completed simply by tuning the PID controllers used by the flight control board to stabilize the vehicle. This board uses its own embedded sensors to control the pitch and roll of the vehicle. Consequently, our primary implementation details begin with the middle layer.

5.2 Middle Layer Implementation

5.2.1 Middle Layer Sensors

Considering the tasks within the middle layer, we implement range sensors facing downwards and in each of the four cardinal directions into our middle layer. We use two ultrasonic range sensors pointed toward the ground to obtain altitude, and one ultrasonic sensor facing each of the left, right, and back of the MAV. As our front sensor will be the most critical for obstacle avoidance, considering the MAV will primarily move in a forward direction, we use a laser range sensor in this direction that gives a more accurate reading with a much greater max range than that of the ultrasonic sensors (40 metres compared to 4 metres). We also consider the MAV's remote controller (which allows for manual flight) as a sensor to the middle layer - both manual and autonomous flight require the stabilization found in the lower layer, so it is the middle layer that must determine whether manual or autonomous flight is taking place.

5.2.2 Middle Layer Algorithms

The overall architecture of the code running on the middle layer can be split into two fundamental divisions - code for manual flight and code for autonomous flight. The division of code to use is determined by the value of a switch on the remote controller. If manual mode is turned on, the middle layer reads in the thrust values from the remote controller, applies a scaling algorithm, and outputs these values to the lower layer. If autonomous mode is activated, algorithms are then used to control the altitude of the craft while all other degrees of freedom are still monitored by the remote controller. In future work, the activation of autonomous mode will result in the full activation of the middle and upper layers as described previously, but the rest of this system has not yet been implemented.

The altitude-controlling algorithms within the autonomous flight section we developed use information from the downward-facing ultrasonic range sensors to control altitude and vertical velocity. Pings are sent out from both sensors: if exactly one sensor returns an out of range reading, the other sensor's data is returned. Otherwise, the readings from the two sensors are averaged together and rounded to an integer to return a stable reading. The data from these sensors gives us not only vertical position, but also a vertical velocity estimate by subtracting the previous height value from the current height value. An altitude stabilization algorithm running on the middle layer then utilizes this information to reach and maintain a given height.

5.2.3 Take-off Task Execution

Within the autonomous altitude control mode on the middle layer, there are three major task divisions: take-off, hover, and land. The take-off task is immediately

activated whenever there is a switch from manual mode to autonomous mode, which is signified by using a boolean flag. Take-off mode starts the thrust at a low level, then continually increases it until a height value slightly above the sitting height value for the MAV is achieved. At this point, the current thrust value is stored as the approximate amount of thrust required for hovering, take-off mode is turned off, and hovering mode is turned on. This take-off process allows the MAV to calibrate the approximate amount of thrust required to hover at a stable altitude.

5.2.4 Hover Task Execution

For the hovering section of the algorithm, the height readings are obtained from the ultrasonic sensors as described previously, then approximate vertical velocity is obtained by comparing the newest height value with the previous value. These values are used as parameters for a PI controller and a P controller, respectively (these controllers were implemented from an existing Arduino library). The PI controller monitoring height corrects toward a predefined goal height, which can be altered by other functions as needed. The P controller monitoring velocity corrects towards a goal that depends upon the MAV's current height. A positive goal is used when the height is more than 10 centimetres below the goal height to encourage upward velocity, and, conversely, a negative goal is used when the height is more than 10 centimetres above the goal height (although this negative goal is very small to prevent crashing to the ground). Finally, if the MAV is within 10 centimetres of its goal height, a velocity goal of 0 is used. This P controller is given a much smaller range of values than the PI controller, as it merely complements the PI controller to offer additional stability. The values produced by the two controllers are added to the hovering thrust calibration variable created by the take-off sequence to get the final thrust, thereby adjusting the height while also monitoring velocity.

5.2.5 Landing Task Execution

The third and final mode of our autonomous algorithm is landing mode. While a transition to this mode will in the future be triggered by output from the upper computational layer, it is currently activated by a secondary switch on the remote controller. This allows us to conduct flight tests of an indefinite length prior to finishing the implementation of the upper layer architecture.

Once landing mode is activated, the basic hovering algorithm continues to be used, but the goal for the PI controller correcting for height decreases by 1 for each loop of the program. This allows reuse of the same functional processes used by the hovering algorithm while slowly adjusting the height values which the algorithm seeks until eventually the goal height is 0 centimetres.

5.3 Upper Layer Implementation

At current status of our research, the function of the upper layer consists solely of data recording. Each loop of the autonomous flight algorithm on the middle layer, the upper layer records both the current output to the actuators as well as the most recent readings from all sensors. These values are stored in a log file that can be accessed once the MAV has landed for analyzation and visualization of flight data.

Additionally, we implement a web based interface for the MAV that enables real-time flight monitoring. This web interface takes flight data and displays it on a public web page so that any web enabled device can be used to witness data visualizations at time of flight. This interface includes capabilities for a live video stream (in anticipation of a camera being added to the MAV in the future), console output from the upper layer computer, and graphical representation of sensor data. Live updating of data is enabled by a Wi-Fi module on the upper layer computer that communicates with the server during flight.

6 Experiment Setup

6.1 Controller Balancing

Our first experiments with the MAV involve basic balancing of the PI and P controllers used for height stabilization in the middle layer. While a large amount of overshoot would be unacceptable as it would result in severe oscillation of the MAV's height and possibly a crash, optimum balancing of the controller is not within the scope of this project. For our purposes, we accept a sufficiently balanced controller to have been found when the MAV is able to stay within 15 centimetres of the predefined goal height.

To help monitor the calibration of this controller as well as record flight data, we use the upper layer controller to record the height value every iteration of the algorithm's main loop. Then, we implement a Python script that converts these values into a list of data, which is read into Wolfram Mathematica to make visual plots of the height over loop iterations. This process allows for a visual representation of flight data using programs native to the Raspberry Pi of which our upper layer consists.

6.2 Floor Obstacle Avoidance

After we were able to achieve sufficiently stable flight over a flat surface, we experiment with flying the MAV over obstacles that change the height of the surface over which the MAV is flying. To do this, we place the MAV in a hallway with three wooden

benches arranged horizontally across the hall. After automated altitude control mode successfully reaches the goal height of 100 centimetres, we then use the remote controller to direct the MAV down the hallway, flying it over the three benches.

7 Preliminary Results on Autonomous Flight

7.1 Controller Balancing

In these experiments, we were able to reach sufficient stabilization of the MAV at various goal heights using only the reactive algorithms on the middle layer. We defined sufficient stabilization as the MAV rapidly reaching the goal height, then remaining within 15 centimetres of the goal height while minimizing the amount of large oscillations. We wanted the MAV to be able to achieve these criteria at goal heights up to 100 centimetres, as this height value allows for a large amount of floor obstacles to be avoided while still giving plenty of space for upward movement inside a typical room (which may be needed to avoid taller obstacles). Fig. 3 shows the plot of the height data from a full flight using our current algorithm. The goal height for these tests was 100 centimetres. Though some height oscillation still occurred in the final test, most of these oscillations were due to horizontal drift and subsequent corrections, and nearly all oscillations were within 15 centimetres of the goal height once it was reached.

7.2 Floor Obstacle Avoidance

The MAV was able to successfully fly over the benches in the hallway using a set goal height of 100 centimetres. The plot of the height values shows a sharp dip in height followed by a sharp peak in height at the three instances in which the benches were flown over, as demonstrated in Fig. 4. The dip corresponds to the MAV beginning to fly directly over top of the bench - the MAV remained at the same absolute altitude at this time, but the sensor readings dipped sharply due to the MAV's lower height relative to the bench. This dip in the relative height of the MAV is followed by elevation to the goal height relative to the bench. Such an approach allows for the MAV to easily traverse floor obstacles, treating each obstacle as the new "floor".

Once the MAV has finished flying over the bench, there is a sharp peak in the distance readings corresponding to the relative increase in distance to the actual floor. The MAV then corrects to its goal height relative to the floor once again. The presence of these dips and peaks followed by a correction back to the correct relative height demonstrate the MAV's ability to successfully handle floor-based obstacles. The sharp dips and peaks produced by flying over obstacles will be used by the upper layer controller in the future to map floor based obstacles and multiple building levels.

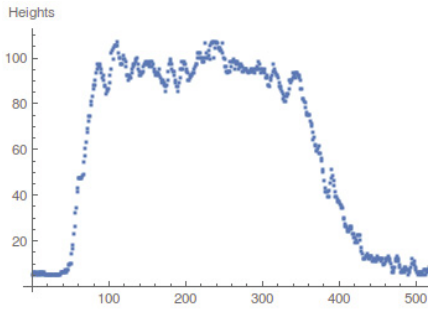


Fig. 3 A full flight with takeoff, hovering, and landing

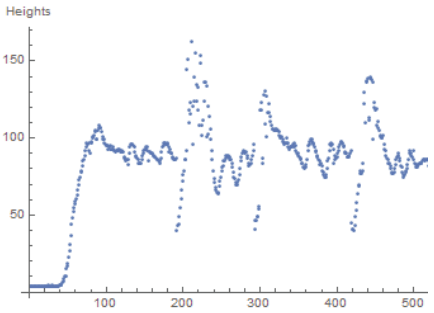


Fig. 4 Height values from flight over benches

8 Future Work

Future work will begin with finishing all algorithms for the middle layer controller, including obstacle avoidance and the execution of more sophisticated tasks. We plan on investigating the implementation of a subsumption architecture on this middle layer. After this work on the middle layer is completed, work will be done on navigation and coverage algorithms using the upper layer controller. Finally, we will finish specifying an interface to be used for communication between layers, and will experiment with the interchangeability of various hardware and software layers utilizing this interface to demonstrate the full abstraction of layers.

9 Conclusion

To date in our project, we have developed a MAV with a three-layered control architecture. This MAV is currently capable of autonomous altitude control and stabilization within 15 centimetres of a given goal height, including altitude adjustment and stabilization over floor obstacles. This MAV provides a platform for the development of mapping and navigation algorithms that do not have to consider critical time-of-flight operations in their design, due to the separation of layers by complexity.

Acknowledgements The research is based upon work supported by a Western Kentucky University FUSE Grant under Grant #15-FA245 and multiple Scholar Development Grants from Western Kentucky University's Student Government Association.

References

1. Achtelik, M., Bachrach, A., He, R., Prentice, S., Roy, N.: Autonomous navigation and exploration of a quadrotor helicopter in GPS-denied indoor environments. In: 1st Symposium on Indoor Flight, International Aerial Robotics Competition (2009)
2. Beul, M., Krombach, N., Zhong, Y., Droeschel, D., Nieuwenhuisen, M., Behnke, S.: A high-performance mav for autonomous navigation in complex 3D environments. In: International Conference on Unmanned Aircraft Systems (ICUAS) (2015)
3. Grzonka, S., Grisetti, G., Burgard, W.: A fully autonomous indoor quadrotor. IEEE Transactions on Robotics **28**(1), 90–100 (2012)
4. Gupta, S.G., Ghonge, M.M., Jawandhiya, P.M.: Review of unmanned aircraft system (uas). International Journal of Advanced Research in Computer Engineering & Technology **2**(4), 1646–1658 (2013)
5. Nieuwenhuisen, M., Droeschel, D., Beul, M., Behnke, S.: Autonomous navigation for micro aerial vehicles in complex gnss-denied environments. Journal of Intelligent & Robotic Systems, 1–18 (2015)
6. Satler, M., Unetti, M., Giordani, N., Avizzano, C.A., Tripicchio, P.: Towards an autonomous flying robot for inspections in open and constrained spaces. In: 2014 11th International Multi-Conference on Systems, Signals Devices (SSD), pp. 1–6, February 2014
7. Tang, Y., Li, Y.: Realization of the flight control for an indoor uav quadrotor. In: 2013 IEEE International Conference on Information and Automation (ICIA), pp. 1278–1283, August 2013
8. Tomic, T., Schmid, K., Lutz, P., Domel, A., Kassecker, M., Mair, E., Grixa, I.L., Ruess, F., Suppa, M., Burschka, D.: Toward a fully autonomous uav: Research platform for indoor and outdoor urban search and rescue. IEEE Robotics Automation Magazine **19**(3), 46–56 (2012)

Using Tweets for Rainfall Monitoring

Luiz Eduardo Guarino de Vasconcelos, Eder C.M. dos Santos, Mário L.F. Neto, Nelson Jesuz Ferreira and Leandro Guarino de Vasconcelos

Abstract In Brazil, the summer season is the wettest period in which many disasters can happen, such as landslides and floods. In recent years, even with the rainy season, the metropolitan region of São Paulo (Brazil) suffers a severe water crisis. Given this scenario, monitoring of rainfall is fundamental for taking preventive actions and planning in the various business branches. Thus, the use of computers to develop tools that assist the rainfall monitoring can help extend the coverage of the existing solutions. Moreover, it is known that, every day, the number of social media users is increasing, and consequently increases the amount of content published in these medias. The objective of this study is to analyze the contents of the Twitter social media, especially the tweets related to rainfall events in order to determine whether this information can contribute to the monitoring of rainfall events in Brazil. More than 1 million tweets published in Brazil related to rainfall were collected in a period of 30 days. Gathered tweets were analyzed and evaluated taking into account the data collected by automatic weather stations (AWS or EMA). The results were satisfactory and indicate a relationship between the geolocated tweets and data from AWS.

Keywords Twitter · Hashtags · Geolocation · Social media analysis

1 Introduction

The number of people connected to the Internet has increased. According to Internet World Stats [1], 3.2 billion people in the world have connected the

L.E.G. de Vasconcelos(✉) · M.L.F. Neto · N.J. Ferreira · L.G. de Vasconcelos
Brazilian Institute for Space Research (INPE), São Paulo, Brazil
e-mail: du.guarino@gmail.com, le.guarino@gmail.com

L.E.G. de Vasconcelos · E.C.M. dos Santos · L.G. de Vasconcelos
FATEA, Lorena, Brazil
e-mail: sep4eder@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_100

Internet. Among these people, more than 2 billion have profiles on social media and messaging services [2]. This number represents 27% of the world population, which is approximately 7.2 billion people in 2015. In Brazil, the population is approximately 204 million people [3] and 54% of the brazilian population access the Internet. Furthermore, there are 96 million active social networking accounts in the country.

Through content and data posted on social media, it is possible to collect, analyze and extract information and knowledge that can be useful to public institutions, companies and any citizen. In addition, the majority of people feel the desire and / or need to publish content on the Internet, both to show a moment of personal life as moments related to occupation.

Due to the success of social media in Brazil, especially Twitter, there was the hypothesis to monitor rainfall through tweets. The hypothesis arose from the need to improve rainfall estimation accuracy in the Brazilian current scenario, after all, this estimate is crucial in several business fields, such as water management, agriculture, disaster prevention, and others.

The most used tools in the rainfall monitoring are rainfall networks and remote sensing. Surface networks represent the intensity of rainfall on time with high reliability, but they have problems to represent its spatial variability [4, 5, 6]. Measures with rain gauges can be influenced by exposure to wind, which can cause up to 20% underestimation of the measurements [7].

Remote sensing is an alternative. Weather radars are able to represent the spatial structure of rainfall systems, but they have several sources of error inherent in the rainfall intensity [8]. The main problems are related to droplet spectrum, the presence of bright band, sampling problems for gate, refractive index of the atmosphere (anomalous propagation), etc. [9, 10, 11]. Also, the radars are costly and do not cover the whole national territory. Surface measures networks and radar rainfall estimates can be combined to reduce the magnitude errors [6].

Given the above, this work aims to analyze the contents of the Twitter social media, especially the tweets related to rainfall events in order to determine whether this information may contribute to the rainfall monitoring in Brazil. For this, an experimental application was developed by the Division of Satellites and Environmental Systems (DSA) of Weather Forecast and Climate Studies Center (CPTEC), under the Brazilian Institute for Space Research (INPE). The DSA meets the demand for information from weather satellites in Brazil. The application developed gathers Twitter data for specific hashtags and/or search terms. Tweets which have geolocation information are displayed on a map in order to facilitate visual analysis.

To validate the experiment, the rain gauge data are shown on the map, collected in near real time, provided by the National Institute of Meteorology (INMET) [12]. There was also a visual analysis with radar data provided by the Meteorological Network (REDEMET) of Aeronautics [13]. The results of initial experiments are satisfactory and are presented in Section 3.

2 Theoretical

2.1 Social Networks

A social network is composed by people or organizations connected to a computer network [14].

Recuero [15] also reported that two elements define a social network: actors (individuals, institutions or groups) and their connections (interactions or social ties).

Social networking sites like Facebook, YouTube and Twitter have users or company profiles, which can be considered actors in a social network. It is known that there is a lot of content generated from the connections between the actors. The interactions can be exemplified by comments made by some users from posting or even by the actions linked to the post, like "retweet" and "favorite" of Twitter (which mean respectively replicate a post from a user to a list of followers, and show that you liked the post content, leaving it in a list of posts with same criteria).

Social ties can be exemplified when there is interaction between two most users of any social media; when all the communication generated by these users becomes frequent and generates a connection between them.

According Recuero [15], social ties are classified into two types: associative and dialogical. Associative occurs when there is interaction between various authors on a social network, such as sharing a link to a news story by a tweet. Dialogic occurs when it does not depend on interactions, requiring only belong to a group or a community, for example decide to follow a user on Twitter.

As reported by Seron [16], the social networking sites are very important in disseminating information, where users have the opportunity both to send and to receive this information. In addition, you can reach thousands of users per second, only with a message posted on Twitter or Facebook.

2.2 Twitter

Every social media has a purpose. It can be said that Twitter is intended to convey information quickly, since each tweet is limited to 140 characters. Furthermore, Twitter allows ease of use of the information through its API (Application Program Interface), mainly through hashtags, resource created by own microblog.

As mentioned, there is a growth in the use of social media. On Twitter, for example, there are more than 316 million monthly active users and over 500 million tweets are sent per day [17].

Even with all the publicity, advertising and marketing of Twitter, beyond good and quick way to personal interaction among users, the microblogging is far from social media most used by Internet users in the world. With Facebook on the list of worldwide market share of social media, with just over 85%, Twitter is in third place with just over 3%, just behind the Pinterest, with nearly 4% (data updated in June 2015) [18].

2.3 Social Network Analysis

With the large volume of data and information provided in social media, it is possible to study patterns of interactions between the actors of social networks [15].

The term "Social Network Analysis" (SNA) has existed for a long time and their use has grown in recent years due to the widespread use of social media.

Researchers from various fields of knowledge are very important in SNA [19]. A large part of social media has powerful tools that analyze user profiles. For example, it can get information about the best days and times for content of publications, based on the estimation of content viewing; view graphics on the effect of any content posted; find out how much people viewing some content; among others.

2.4 Weather Information

For the acquisition of meteorological data in Brazil, automatic weather stations (AWS or EMA in portuguese) and meteorological radars are used.

An automatic weather station collects, minute by minute, weather information (temperature, humidity, atmospheric pressure, rainfall, wind speed and direction, solar radiation) representing the area where it is located. Every hour, these data are available for transmission, via satellite or mobile telephony, to the headquarters of INMET in Brasilia. The set of received data is evaluated by a quality control and stored in a database.

In addition, the data is freely available in real time via internet [12] for the development of weather forecast and the different meteorological products of interest to industry users and the public in general, and for a wide range of applications in research meteorology, hydrology and oceanography.

Another important feature in meteorology is the use of weather radar. The radars are used to perform the meteorological monitoring on vulnerable cities the occurrence of floods, flash floods and geological events such as landslides. Radars produce information necessary for the preparation of warnings about possible disasters related to rainfall.

Weather radar in Brazil are the responsibility of RedeMet [13] which aims to integrate the meteorological products focused on civil and military aviation, in order to make access to this information faster, efficient and safe.

Furthermore, in Brazil there is the monitoring of rainfall using satellite images. The satellite-based measurements establish relationship with meteorological variables estimated by existing devices on satellites. The accuracy of the information depends on the resolution of the images of the satellite. In this paper will not be covered comparative analyzes with satellite images.

3 Case Study

This section presents the solution developed by the Division of Satellites and Environmental Systems (DSA) for monitoring rainfall through the use of tweets.

We present the architecture of the solution, the storage solution, data distribution and the data analysis.

To collect the tweets is necessary to consume the data from the Twitter API, which provides a defined set of request and response messages. This API is well documented and enables the development of various applications such as websites, widgets and other projects that can interact with Twitter. The API also allows you to use almost all the features that are available on the site. For example, the user information, the timeline, their friends and followers, tweets and retweets, search for tweets, among others. It isn't possible, for example, consume information from a profile that has no relationship with the application that collects the data. Only public information can be collected through the APIs, that is, information that users without access restriction. The application communicates with the API over HTTP requests and receives data in JSON format (JavaScript Object Notation). To query information is used the GET method and to send information is used the POST method. Several error messages can be obtained of the API to verify if is possible to get information of the tweets [20].

The Twitter API allows extract various information of the social network, such as: friendly relations among users, who users are following, friends, the user profile and posted public tweets. The disadvantage of this API is the limit of requests, which is 180 requests per 15 minute intervals depending on the resource being requested. In addition, each request returns up to 100 tweets in one request.

For access the API's information, the application should be authenticated through the OAuth protocol [21]. There are two types of requests: refers to public information and user data modification. To manipulate data of a user, as well as authorization to access the API, it is also required user authorization. To develop an application that accesses data public must obtain a token that is provided on the Twitter site. Thus, the application can get the JSON responses of the API.

The Twitter API provides various information related to a particular tweet, such as date and time of creation, location, among others. Information relating to the user's location are disabled by default in Twitter. If user enable this option, the application can get the information about tweet location. This information includes the latitude and longitude of the user.

3.1 Application Architecture

The Twitter's data stream is continuous, however the API sets limits for data consuming, and it was necessary to develop a tool that collects and store the tweet's information and the user who made the publication. The application architecture can be seen in Figure 1. Briefly, the information of the tweets are collected through the Twitter API (a) and rainfall information are collected from AWS (b). The information from these sources are stored in a database (c) and are consumed through a web application (d).

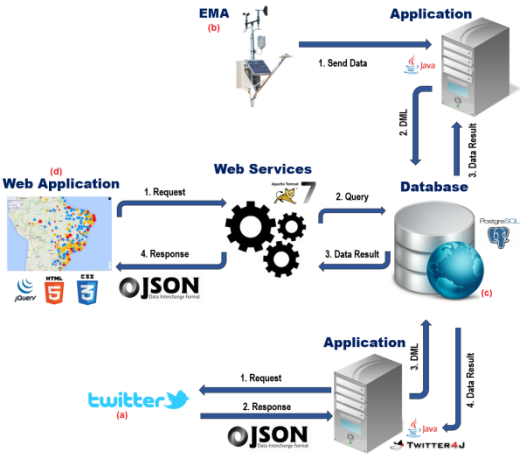


Fig. 1 Application Architecture

In the application were used:

1. The Java programming language, along with the Twitter4J, which is an abstraction of the Twitter API.
2. Database PostgreSQL to the storage of all data collected by the application;
3. Web service in Java, hosted on Apache Tomcat 7 Server;
4. Web application that consumes and uses the web service to view the data on a map.

The application uses the Twitter information, by Twitter4J, using a request with some parameters such as: amount of tweets to be collected, default language of the posted tweets and tweets from a date. These parameters are processed by Twitter4J, being sent to the Twitter API and then are returned in JSON files.

All tweets collected by Twitter4J are returned in JSON files, received by the application and then are entered into the database.

In addition, another application receives the data from the AWSs, stores them in the database and makes the rainfall data are available in JSON files.

The web application consumes the web services that return information, for example, what are the tweets of a particular hashtag, what are the tweets of a date; what are AWSs with millimeters of rainfall greater than zero; what are AWSs with millimeters of rainfall greater equal zero. The information returned are shown on a map using latitude and longitude of each information (tweet or AWS).

3.2 Experiments and Results

Initially, we evaluated several words that users can used to represent rainfall. For this, we used the site TopSys [http://topsy.com/analytics] that shows the approximate amount of tweets for a period and word. The words used in the search are in Table 1. This search occured between 6 August 2015 to 06 September 2015.

Table 1 Words used to rainfall Monitoring

Word (in Portuguese)	Word (in English)	Amount of tweets	%
chuva	rain	793,132	71,15
chuvisco	drizzle	786	0,07
trovoada	thunderstorm	10,544	0,95
nevoeiro	fog	5,310	0,05
tempestade	storm	76,231	6,84
raio	lightning	60,087	5,39
garoa	drizzle	4,934	0,44
trovão	Thunder	19,012	1,71
temporal	temporal	112,505	10,09
relâmpago	lightning	26,158	2,35
dilúvio	flood	23,790	0,21
chuvarada	rain	4,219	0,38
precipitação	rainfall	1,456	0,38

After the Table 1 analysis, it was defined #chuva (rain) hashtag and the search term “chuva.” We collected 1,328,221 tweets (posted by anybody) between days 08 September 2015-08 October 2015. The collection was done through the application.

From the collected data, it was found that only 0.1% of tweets had the latitude and longitude in the tweets. To increase the amount of tweets with geolocation, the city defined in the user profile was used which is available in each tweet. When the user city is used, a attribute in database is setted to indicate that latitude and longitude have been adjusted by the application. For this, they were stored all geolocation information (latitude and longitude) of Brazilian cities. These are available in the Brazilian Geodesic System (SGB), of the Brazilian Institute of Geography and Statistics (IBGE) [22]. For each tweet without geolocation information, were assigned for tweet the latitude and longitude of the user city provided by SGB. This allowed 23.78% of tweets were geolocation information.

From this information, the web application was developed for viewing the location of the tweets on the map. This application makes use of the Google Maps API. Figure 2 shows the map of Brazil with the collected tweets (marker in red color), the AWS who had a change in rainfall (indicating that there was rainfall on site - Umbrella icon), and the AWS who had no change the rainfall (in green marker). Each marker on the map represents a tweet with their respective coordinates (latitude, longitude).

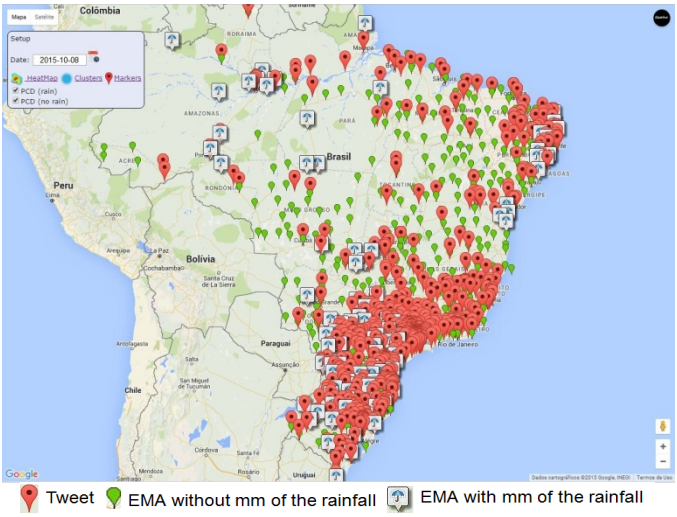


Fig. 2 Chart with markers - October 8, 2015

In Figure 3 is shown a heat map, which allows viewing of the places with more tweets. Red marker indicates the region that had more published on the subject, while that the marker light blue represents the region with less tweets.

The information in Figures 2 and 3 are of the day 8 October 2015, which accounted 48.500 tweets about rainfall. In addition, 102 AWS received mm of rainfall and 439 AWS had no rainfall that day.

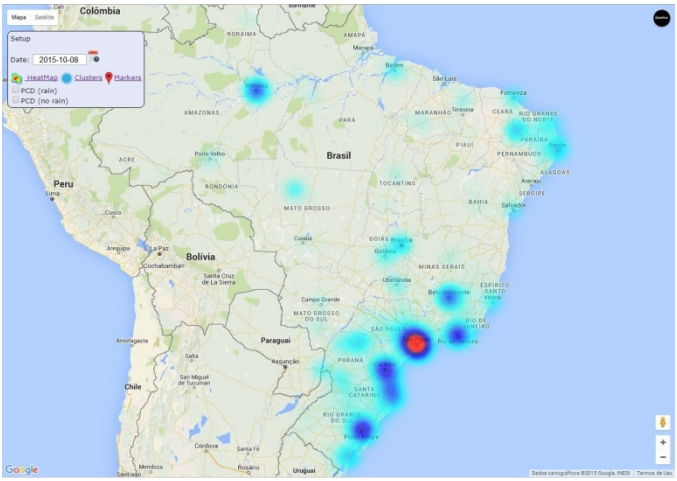


Fig. 3 Table 2. Heat Map chart - October 8, 2015

In Figure 4 is shown a region of the map of Brazil. In addition, markers that are close to each other are grouped, forming clusters. This makes it easy to display on the map. If the user zooms in application, clusters are dismembered. Cluster in red color indicates greater number of counters in the same area and the blue color indicates a smaller amount. Within each cluster is shown the amount of the markers that were grouped. In this figure, it also is possible to see the distance between the tweets and AWS. In Figure 5 is shown the same region of Figure 4, but without clusters of tweets. Red markers represent the tweets. It can view a good concentration of tweets close to the AWSs.

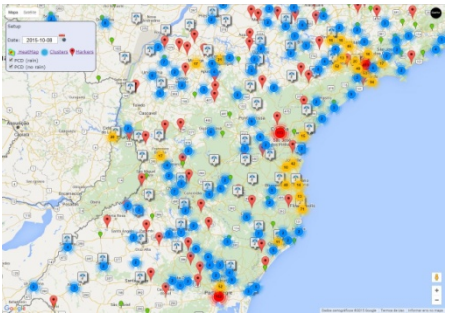


Fig. 4 Chart with tweets (clusters) and AWS's of the Southern Brazil region – October 8, 2015

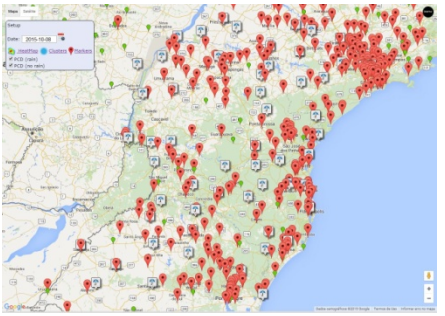


Fig. 5 Chart with tweets (markers) and AWS's of the Southern Brazil region – October 8, 2015

In figure 6 its is possible to visualize the relationship between the amount of tweets and the amount of mm of the rainfall per day. It can be seen that there is a relationship between these two variables. Blue line refers to the tweets and the red line refers to amount of mm of the rainfall.

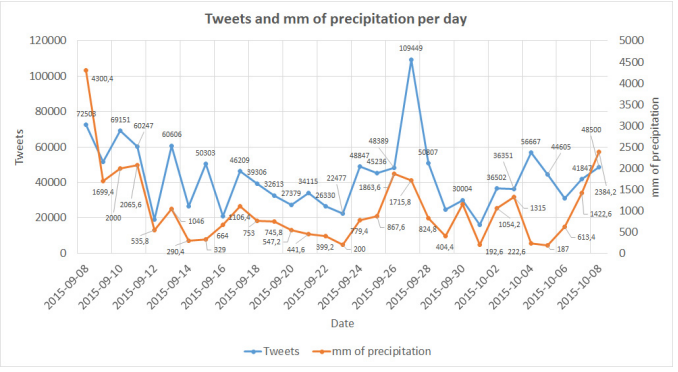


Fig. 6 Amount of tweets and mm of the rainfall per day

4 Conclusion

In this work, we performed social network analysis of Twitter, in particular, the tweets related to rainfall events in order to determine whether this information could contribute to the monitoring of precipitation events in Brazil. It can be seen that a close relationship exists between the amount of tweets and the amount of mm precipitation detected by AWS.

The solution used in this study is a tool that can assist in monitoring rainfall events. Through it, there is the possibility, for example, to provide warnings about certain weather conditions and especially supplement the information provided by other monitoring means (eg AWS, weather radar, satellite imagery).

It is also concluded that the study and analysis of social networks and their media is very important, not only to the field of technology, but also for other fields of knowledge.

Considering possible future work, it is possible to evaluate the possibility of dismembering the information of tweets to analyze content, making sure the content is positive or negative; analyzing feelings of tweets by checking the behavior of users through the content posted; and expand the analysis considering evaluate other social media, such as Instagram, which has the use of interactions through hashtags, similar to those of Twitter, with the difference being only through photos or videos.

Moreover, given the success of the experimental application, CPTEC should make the release of a specific hashtag so that people can inform precipitation events. The suggested hashtag is #CPTECCHUVA.

Other suggested work are (i) examine minimum and maximum distance of tweets regarding the AWS with and without precipitation statement; (ii) monitor hourly tweets relating them to rain gauges and radars; (iii) visually compare the tweets map with satellite images processed by DSA/CPTEC/INPE.

References

1. Internet World Stats. http://www.internetworldstats.com
2. We Are Social. Digital, Social & Mobile Worldwide in 2015. http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/
3. IBGE. Brazilian Institute of Geography and Statistics. Projeção da população do Brasil e das Unidades de Federação. http://www.ibge.gov.br/apps/populacao/projecao/
4. Rocha Filho, K.L.: Modelagem hidrológica da bacia do Rio Pirajuçara com TopModel, telemetria e radar meteorológico. 138 f. Dissertação (Mestrado em Meteorologia) – Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo (2010)
5. Silva, F.D.S.: Análise objetiva estatística da precipitação estimada com radar e medida por uma rede telemétrica. 2006. 101 f. Tese (Mestrado em Meteorologia) – Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo (2006)

6. Pereira Filho, A.J., Crawford, K.C., Hartzell, C.: Improving WSR-88D hourly rainfall estimates. Weather and Forecasting **13**(4), 1016–1028 (1998)
7. Legate, D.R., Deliberty, T.L.: Measurement biases in the United States rain gauge network. Water Resource Bulletin **29**, 855–861 (1993)
8. Calvetti, L., Beneti, C., Pereira Filho, A.J.: Integração do radar meteorológico doppler do SIMEPAR e uma rede pluviométrica para a estimativa da precipitação. In: Simpósio Brasileiro de Sensoriamento Remoto, Belo Horizonte. Anais do Simpósio de Brasileiro de Sensoriamento Remoto 2003. CD-ROM (2003)
9. Austin, P.M.: Relation between measured radar reflectivity and surface rainfall. Monthly Weather Review **115**, 1053–1070 (1987)
10. Batlan, L.J.: Radar Observations of the Atmosphere, p. 324. The University of Chicago Press, Chicago (1973)
11. Doviak, R.J., Zrnic, D.S.: Doppler radar and weather observations. Dover Publications (1993)
12. INMET. Brazilian National Weather Institute. http://www.inmet.gov.br/portal/index.php?r=estacoes/mapaEstacoes
13. Redemet. http://www.redemet.aer.mil.br/
14. Garton, L., Haythornthwaite, C., Wellman, B.: Studying Online Social Networks. Journal of Computer-Mediated Communication. http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1997.tb00062.x/full
15. Recuero, R.: Redes sociais na internet, Sulina (2009)
16. de Souza Seron, W.F.M.: Análise de Redes Sociais - Um Estudo do Twitter. São Paulo. Dissertação. Instituto de Ciência e Tecnologia. Universidade Federal de São Paulo (2015)
17. Twitter. Uso do Twitter / Fatos Sobre a Empresa. https://about.twitter.com/pt/company
18. AREPPIM. http://stats.areppim.com/stats/stats_socmediaxsnapshot.htm
19. Matheus, R.F., de Oliveira Silva, A.B.: Análise de redes sociais como método para a Ciência da Informação. Revista da Ciência da Informação, **7**(2) (2006). http://www.dgz.org.br/abr06/Art_03.htm
20. Twitter. Error Codes & Responses. https://dev.twitter.com/overview/api/response-codes
21. OAuth. Disponível em: http://oauth.net/
22. IBGE. Brazilian Institute of Geography and Statistics. http://www.ibge.gov.br/home/geociencias/geodesia/bdgpesq_googlemaps.php

Algorithms Performance Evaluation in Hybrid Systems

Rafael Manochio, David Buzatto, Paulo Muniz de Ávila and Rodrigo Palucci Pantoni

Abstract Over recent years a new branch of parallel computing has emerged: hybrid architectures composed of CPUs (Central Processing Unit) and GPU (Graphics Processing Unit). The objective of this work is to quantify the dimension of the performance gains comparing sequential processing of known algorithms with parallelized processing offered by hybrid architecture. It was used OpenCL (Open Computing Language) generated by the AMD Aparapi Framework to implement popular computational algorithms. Results show the performance using Aparapi Framework and the conclusions details the expected scale of algorithm speed up comparing with other works found in the literature using pure OpenCL library.

Keywords Parallelism · GPGPU · CPU · OpenCL · Aparapi · Hybrid systems

1 Introduction

There are many processors developed to perform a specific function with higher performance than would be obtained by a general-purpose processor, known as CPU (Central Processing Unit). Graphics processors, for example, are designed to

R. Manochio · D. Buzatto
Department of Informatics, Federal Institute of Education,
Science and Technology of São Paulo (IFSP), São João da Boa Vista, SP, Brazil
e-mail: {rafael.manochio,davidbuzatto}@ifsp.edu.br

P.M. de Ávila
Department of Informatics, Federal Institute of Education, Science and Technology
of South of Minas Gerais (IFSULDEMINAS), Poços de Caldas, MG, Brazil
e-mail: paulo.avila@ifsuldeminas.edu.br

R.P. Pantoni(✉)
Department of Informatics, Federal Institute of Education,
Science and Technology of São Paulo (IFSP), Sertãozinho, SP, Brazil
e-mail: rpantoni@ifsp.edu.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_101

calculate huge amounts of data in parallel and display their results (in the form of images) on the monitors.

As a result, using all the power of parallel computing contained in graphics processors has become an object of study. The concept is called GPGPU (General Purpose Graphics Processing Unit).

In this context, this paper aims to implement popular computational algorithms in order to collect quantitative information about the scale of algorithm speed up using GPGPU with OpenCL (Open Computing Language) generated by the Aparapi Framework.

Section 2 presents the related works and Section 3 details the methodology and tools applied, as OpenCL and Aparapi Framework, and popular computational algorithms applied. Section 4 presents the results. In the last section, the discussion and conclusions are delineated.

2 Related Works

Papers related to GPU performance testing are strictly focused on analyze the performance, for those who are affectionate to video games, through measurement data such as FPS[1] (Frames per second). Such measurements are usually performed with specific software like RightMark and 3DMark, which are specifically designed to provide performance measurement of GPUs when it executes a code in DirectX[2].

Furthermore, most related work focus on measuring performance of GPUs with parallel code in general purpose GPUs in Hybrid systems, aims the NVIDIA[3] architecture and its proprietary CUDA (Compute Unified Device Architecture) language [2]. Another system called Gbench (NCBI Genome Workbench), is used for comparison of performance between CPU and GPU in common operations from the well-known engineering software MatLab.

Parboil and Rodinia [1] are benchmark systems focused on behavior analysis in the execution of scientific applications and kernels on GPUs, however, those systems were also written using the CUDA language from NVIDIA. Another system called SHOC (Scalable Heterogeneous Computing Benchmark Suite), works in a way slightly different from the others, according Danalis: "The SHOC is distinct from other systems because it was designed to be a truly scalable set of performance tests, capable of testing large clusters and a huge number of different devices" [1]. The SHOC includes support for OpenCL language; it also distinguishes of this study because it does not operate specifically with the Aparapi framework.

---

[1] Frames per second: The number of frames per second that the GPU is capable of rendering in a given game.

[2] DirectX: Microsoft DirectX is a collection of application programming interfaces (APIs) for handling tasks related to multimedia, especially game programming and video.

[3] NVIDIA: (NVIDIA Corporation, Santa Clara, CA, www.nvidia.com) A leading designer of graphics processors for personal computers.

It is possible to find on the web many unspecific scientific papers related to codes using Aparapi framework, which it is surprising, when it comes to a Framework as specific and recent, which it shows the strength this new computing proposal as a Hybrid computing has.

There are works like the Hung-Shuen Chen research [3], which currently works with the integration of frameworks like Aparapi on mobile devices that use the Android operating system, bringing in this way, the benefits of heterogeneous computing to these devices whose main have Java as programming language.

Other studies however, are focused on performance analysis with Aparapi in very specific applications like the Krishan Gopal Gupta research [4], where it made performance tests with Aparapi framework, through the implementation of an algorithm called Sobel Edge Detection, which is an image processing for detecting edges of shapes obtained from any image. The results obtained by Kristian were enlightening, according to the results, since only in large workloads is that the Aparapi framework shown compensatory, as its execution needs to create the threads and OpenCL kernels don't compensate for small workloads when faced with the performance of a current multicore CPU.

Finally, the paper by Shrinivas Joshi [5] uses the Aparapi to accelerate the performance of financial applications developed in Java. Their conclusions were very similar to the Kristian Gupta [4].

The Aparapi is a simple and effective mechanism to bring the computational power of GPUs for applications in Java. If there is a Java application with data parallelism in its nature, it is highly recommended the application of Aparapi framework.

All the papers that were used give a clear and concise view of the points where the Aparapi framework could prove a breakthrough performance. The following section shows different tests with Aparapi framework, using simple and routinely algorithms in order to prove whether those statements are true for the chosen cases in this study.

3 Methodology

This section describes the methodology used in this work, including the programming language and the specific algorithm implemented.

3.1 Processing Forms

In this study, it was used three distinct ways of implementation to measure the variation in code execution performance.

The first way is sequential execution on the CPU, using only one of the cores available in the processor to perform routines. The second way is the parallel processing in CPU using JTP (Java Thread Pool). The JTP manages a group of ready threads to run and maintains a queue of tasks awaiting execution. This method is able to use all available CPU cores so as to optimize use of processing time and

speed up application execution. The third way is the GPGPU where the code runs on the graphics card (GPU) in parallel using its many processing cores to perform the requested task.

3.2 Used Framework

OpenCL is a framework for parallel programming in heterogeneous systems. This framework includes the C language variation for OpenCL, where compiler and development environment needed to run using code written in OpenCL C [6].

Despite all the benefits, program and optimize a code in OpenCL is extremely labor intensive, due to the use of languages that are far from the high-level (e.g. C/C ++) and also the fact that the programmer must deal with, inside the code, data transfer between the memory of the host and the device (CPU to GPU).

The Aparapi (at runtime) identifies whether a system supports code execution in OpenCL and converts the bytecode of the "*run()*" method in Java to OpenCL. If the code, for any reason, cannot be converted, the API executes the code using JTP, which allows the execution of multiple threads in Java concurrently.

The Aparapi framework was used in all the tests conducted in this work and the results collected provide a more detailed view at the performance gains achieved.

3.3 Algorithms

To measure the variation in performance related to the execution time of algorithms in the three ways of processing that were described, the following algorithms were chosen.

An array multiplication algorithm is a variation of the code provided by Vasanth Raja Chittampally [7] that performs a simple multiplication of the number positions of a two-dimensional array and returns the elapsed time of this task in different methods, as well as makes a full comparison of the multiplications results to ensure that the calculations are correct. This algorithm is able to show the power of parallel calculation of the GPUs compared to the power of crude mathematical calculations of a common multicore CPU.

Three popular sorting algorithms are chosen: BubbleSort, SelectionSort and QuickSort; these sorting algorithms have a code structure rather sequential and/or recursive. The decision to test and verify sorting algorithms and the selection of the three algorithms mentioned above not have been by chance. According to Robert Sedgewick [8], at the beginning of computing, it was believed that 30 percent of all processing cycles from one computer were spent performing some kind of sorting algorithm. If this fraction is lower today, one reason is that the sorting algorithms are relatively more efficient, not the need or the importance of conducting the ordination has decreased, in fact, the ubiquity of computing put us in a sea information the first step to be organized is usually ordering. Therefore, all computer systems have sorting algorithms implementations.

Sorting algorithms have a huge importance for the processing of commercial and scientific data in modern computing. Applications related to transaction processing, combinatorics, astrophysics, molecular dynamics, linguistic, genomic, weather and many other fields.

One of sorting algorithms included in this work, QuickSort, is considered one of the ten most commonly used algorithms for science and engineering at 20 century [8].

With selected test algorithms, the next step is to transcribe the code for parallel execution. Dependencies between data structures and processes should be analyzed to select the areas of the code that will be executed in parallel. This step is very important because is quite clear in asserting that the maximum performance gain to parallelize the code is 1/y with "*y*" representing a percentage of code that cannot be ported to parallelism.

The BubbleSort code structure, in the implementation used in this paper uses only two "*for'*s " loops to sort an input array with "*n"* positions.

The code below, written in Java, presents BubbleSort algorithm applied in this work.

By analyzing the code, it is noticed that it is possible, in this particular algorithm, to parallelize almost all of it, since there are no dependencies between processes.

According to the Amdhal's Law the performance gain expected of BubbleSort algorithm when it is executed under a graphics accelerator with 1.024 cores, assuming a value of "*y*" as 1%, will be 504 times. As will be evidenced in the results, this value cannot be proved due to factors not considered by the Amdhal's Law which is not designed to take hybrid architecture.

```
public class CpuBubbleSort {
  public float[] BubbleSort( float array[] ) {
    for ( int i = 0; i < array.length - 1; i++ ) {
      for ( int j = 0; j < array.length - 1 - i; j++ ){
        if ( array[j] > array[j + 1] ) {
          float temp = array[j];
          array[j] = array[j+1];
          array[j+1] = temp;
        }
      }
    }
    return array;
  }
}
```

4 Results

The tests were performed on a desktop computer with AMD Quad-core processor. The full details of the hardware configuration are described in Table 1.

Table 1 Hardware Configuration

Hardware	Description
Processor	AMD Phenom II X4 810 (2,6 GHz, 4x 512KB L2 e 4MB L3)
RAM	8GB DDR3 1333 MHz
Disk	1TB Sata 3Gbs 7200 RPM
GPU	AMD Radeon HD 7850 (860MHz, 2GB/256Bit 1200Mhz 4,8Gbps GDDR5)
Motherboard	Gigabyte GA-78LMT-S2P (Chipset AMD 760G, South-Bridge AMD SB710).
OS:	Windows 8 Pro X64
Drivers	AMD Catalyst 13.1 (9.012-121219a-152192C-ATI)
Framework	Aparapi version 23-01-2013 X64

All tests were performed with a varied number of inputs, using 256, 1.048.576 and 4.194.304 positions of floats, and those arrays were randomly arranged in order to make a fair comparison for all algorithms, avoiding the risk of change in the results by running a worst-case performance for each one of them, because, test the efficiency of these algorithms is not the goal of this work as these are already well known and justified.

4.1 Multiplication of Array Elements

This test consists in performing a multiplication algorithm of an array A and B and as result, the attribution to a C array. The calculations will be done independently, comparing the performance of the algorithm execution between sequential execution (CPU), JTP (multithreaded CPU using all cores) and GPU.

Different sizes of array of floats are used to examine the variation in runtime according to the data increase: 256, 1.048.576 and 4.194.304 positions were used. Before the execution of each test, some computer processing load measurements were made in order to avoid the interference from the other software loads in the results.

Fig. 1 presents the compiled results including the execution times using the hardware configuration listed with 256, 1.048.576 and 4.194.304 positions.

First, the Fig. 1 shows the result for a multiplication with 256 positions, since the overhead of the threads generation and the start time of the framework took at least 0.5 seconds, making this CPU the most quick, taking less than one millisecond to perform the calculations. This is due to the fact that, beyond the framework delays that was already mentioned, the data still must pass from the computer's main memory to the frame buffer of the GPU, which causes a higher delay. There was no noticeable change in the CPU load while performing this test, which shows that the CPU was not required satisfactorily.

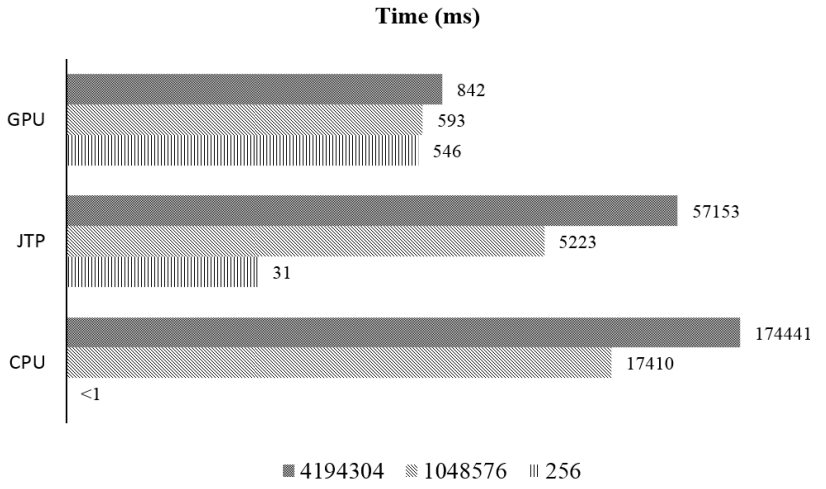


Fig. 1 Results of Multiplication of Array Elements

With a significantly higher array of 1.048.576 positions, the positive results of the ported code begin to emerge. Fig. 1 shows the huge advantage gained by the GPU over the code execution in the JTP mode and the CPU. It can be seen that the GPU had no significant gain in execution time compared to the 256 positions, which allows us to say, with certainty, that the times achieved in the previous figure are actually relate to the overhead of the framework.

During code execution in CPU, only one of the processor cores was used, which shows the serial nature of code execution on CPU; in JTP the four processor cores are used which is reflected in a shorter execution time.

With a significantly higher array of 4.194.304 positions, the GPU took 842 ms to process, while JTP took 67 times longer to run the code and, the sequential code in CPU took 207 times longer to complete than GPU.

The CPU load on JTP stood at about 100% in the four cores all the time during the execution.

The CPU was not affected by the execution on the GPU, of which, only one of the cores was approximately 40% load during the test.

These high differences in execution time, among the various processing forms, especially in the array of 4.194.304 positions, shows the power of GPUs when the requirement is data parallelism.

The architecture of GPUs in general is optimized to perform SIMD (Single Instruction, Multiple Data) operations required to execute the most stringent modern computer graphics, this same architecture has proved extremely efficient, as expected, for mathematical calculations.

One of the promises when AMD announced its GPUs GCN architecture (used in the test) was just superior performance in general purpose calculations on the GPU; however, their performance in front of your NVIDIA GPU competitor in this sense is not the purpose of this paper.

4.2 Sorting Arrays

In this test were executed array sorting algorithms in the above mentioned hardware configurations. These tests differ totally from previously conducted, because it does not focus on GPU data parallelism capability (which is known to be the strong point of the GPUs), but to the ability to perform parallel instruction with a high data movement load and access to memory and disk, which is the strong point of common processors.

As in the previous test, current computer processing load measurements were made in order to avoid interference from other software load on the results.

In Fig. 2 are compiled the execution results of the sort algorithms tested, BubbleSort, SelectionSort and QuickSort with a 256 positions array filled with random float type numbers using sequential execution in CPU, parallel CPU - JTP, and parallel GPU.

With the exception of QuickSort algorithm, the results did not differ much from results of the array multiplication algorithm of 256 positions; the creation of threads and the initialization of the Aparapi framework still take about 500 ms processing time to be initialized. As can be seen in the Fig. 2, there is no measurement time for QuickSort algorithm in the GPU, this is due to the fact that, in the implementation used in this work, the Aparapi framework refused to execute the code because the use of unsupported instructions by the GPU in the code. It was not found in the available literature a specific reason for this, but it is known that during the conversion of Java bytecode to OpenCL, some not found instruction was generated in the error report which made it impossible the execution by the GPU.

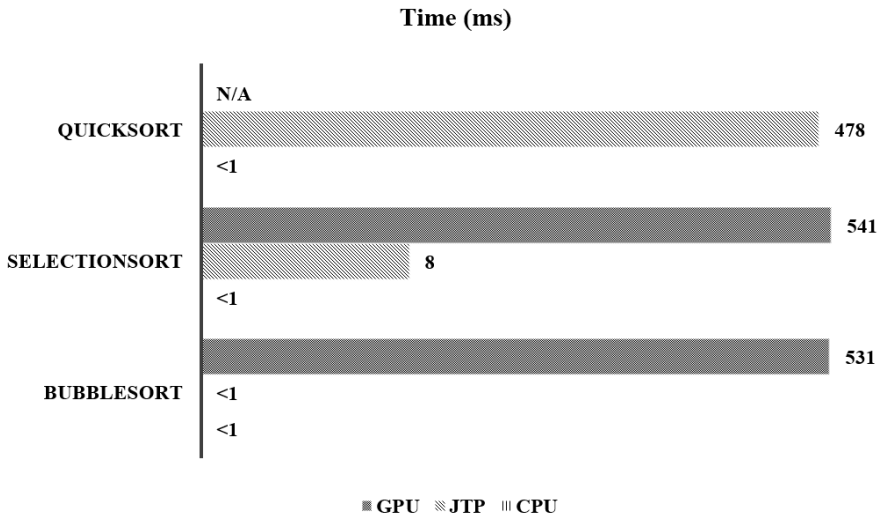


Fig. 2 Results of Sorting Algorithms in an array of 256 positions

Following up the tests, the Fig. 3 compiles the results obtained by the same algorithms and the same hardware configuration, but with an array of 1.048.576 positions.

The SelectionSort algorithm, that has a $O(n^2)$ order of growth, in relation to execution time, achieved a slightly lower performance than the BubbleSort which also is classified as a $O(n^2)$ algorithm; due to their order of growth, both algorithms did not had a good performance to sort large lists, which is impacted in a negative outcome regarding QuickSort which has linear-logarithmic (O(n log n)) order of growth and therefore was able to sort the array in less than 1 ms in CPU. On the other hand, in GPU and parallel processing in CPU, all the algorithms had lower results with those obtained in the CPU for this array size.

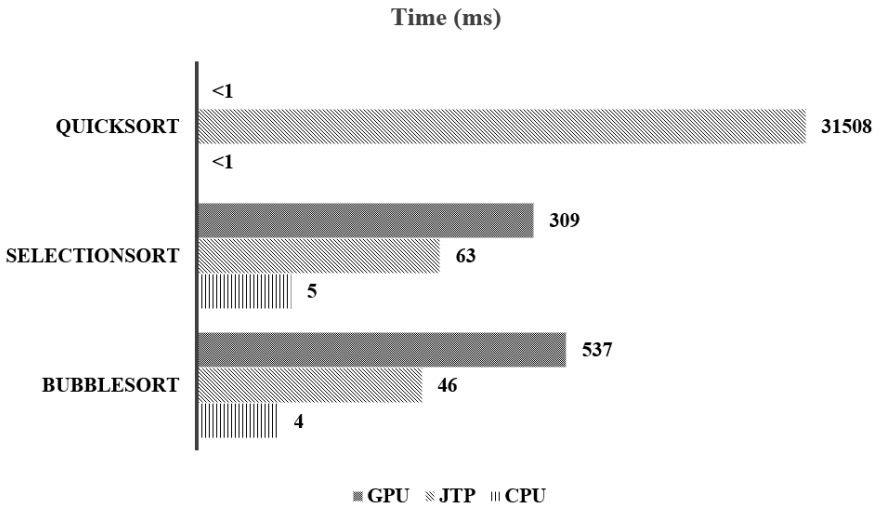


Fig. 3 Results of Sorting Algorithms in an array of 1.048.576 positions

Therefore, the Fig. 4 compiles the results obtained by the same sorting algorithms, but now, with an array of 4.194.304 positions.

An array, twice the previous size, it was not enough to change the runtime of the QuickSort algorithm in the CPU, which is still able to order the array in less than 1ms; however, the BubbleSort and SelectionSort algorithms remained virtually the same proportion of performance of the previous test, differing only the result in the JTP (parallel CPU), where the implementation of BubbleSort was 8 times more efficient than the SelectionSort, but still, all lower than the serialized processing in CPU. In the GPU column results are consistent with the proportion of running on CPU but much higher at runtime than the options in CPU; this highlights the latency that exists in communication with the GPU, as well as the time required to prepare the threads for parallel execution.

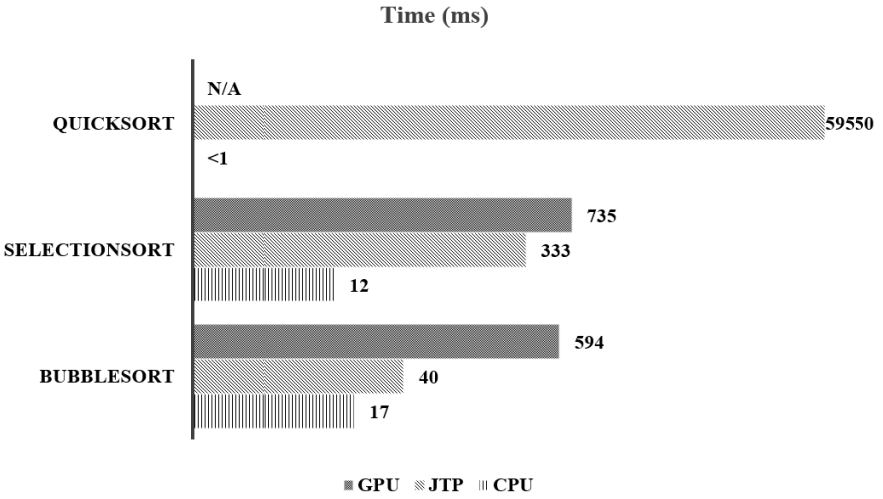


Fig. 4 Results of Sorting Algorithms in an array of 4.194.304 positions

When performing an analysis on the results obtained, there is minimal variation in execution time of the BubbleSort algorithm, regardless of the input value. This algorithm was the most simple to implement and the easiest to port to the parallelism because of the extremely simple features of the code.

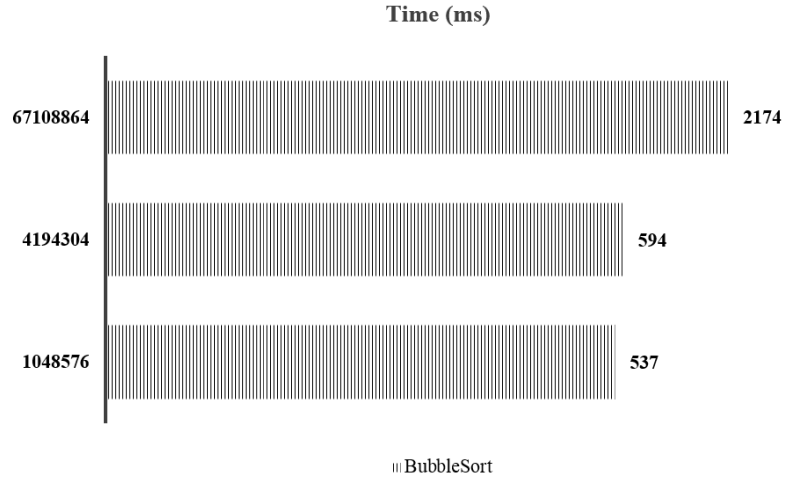


Fig. 5. BubbleSort algorithm execution time comparison

In order to explain this behavior, a final test was performed with BubbleSort algorithm using an array with 67.108.864 positions, as showed in Fig. 5. With this array, there was a considerable variation in the BubbleSort in GPU runtime. This

result shows that the array size is determining factor for the result obtained. In theory, an input with 4.194.304 positions was not large enough to stress the graphical accelerating processor during the execution of this particular algorithm, in essence this code has more mathematical operations that other sorting algorithms and, for this reason, the values are similar.

As shown in Fig. 5, the performance improvement of up to 504 times according to Amdhal's Law was not observed.

5 Discussion and Conclusions

In the tests conducted, both in array multiplication as in sorting algorithms, the Aparapi API took about 500ms to be initiated and convert the Java bytecode into the OpenCL code and start the execution in the hardware configuration used in the tests, as predicted by Gupta in his work [4].

In conclusion, the implementation of sorting algorithms performed in Java using the API Aparapi used in this work have not gained any performance relative to its execution in CPU, on the contrary, there is significant loss of performance with increased time required to run regardless of the input used, as measured in the last test with BubbleSort algorithm. The same thing did not occur when Algebra operations were used, in this case, the gain of performance becomes increasingly significant as the input increases and, the CPU usage, was minimal during the tests, causing the computer as a whole can perform other tasks while the graphics accelerator is in charge of running the heavy calculations.

The test results show that the developer should test the performance of different areas of their code in parallel execution to optimize the running time, since some functions will get much underperformed when executed in parallel, but the gains obtained by some types of algorithms when executed in parallel on the GPU cannot be overlooked, especially those that use a lot of mathematical data and calculations.

The performance gain should not be measured only by the runtime of a particular piece of code, but also by the decreasing of the CPU workload.

During the execution of instructions by the GPGPU, the CPU has remained practically idle, thus the processing power can be used to process other areas of code that cannot or should not be processed by the GPU.

Professor Wu Feng in his work entitled "The 13 (computational) dwarfs of OpenCL" [9] deepened in the algorithms that best fit for vector processors, a very important document for users of the Massively Parallel Computer.

In this document it is stated that the 13 best type algorithm to be used in vector processors are:

- Algebra: array multiplication (used in the test 1);
- Sparse Linear Algebra: Multiplication of matrices consisting mainly zeros;
- Spectral Methods: solving certain differential equations, such as the Fourier transform;
- N-Body type Methods: dynamic particles;

- Structured Grids: image processing;
- Unstructured Grids: fluid dynamics;
- Map-Reduce and Monte Carlo: pi calculus, collision simulation, processes that do not need to communicate often;
- Graph Traversal: sorting algorithms like QuickSort (used in the test 2);
- Dynamic Programming: problems with Graphs;
- Backtracking: puzzle games as the chess game;
- Graphic-Probabilistic Models;
- Finite State Machines: Data Mining.

Based on results, it can be assumed that something is wrong with the tests; Wu Feng clearly states in his text that GPUs are excellent for array multiplication (which was attested in the tests, with the graphics accelerator and up to 200 times faster than serial processors - depending on the size of the input) and Graphs Traversal as used in sorting algorithm, QuickSort. However, it became apparent in tests with sorting algorithms in this paper, the performance of these algorithms on GPUs, and even using the various cores of sequential processors fell below the performance achieved by the same algorithm and the same input using sequential processing.

The comparison with Wu Feng work cannot be done directly because it refers to pure OpenCL and does not take into account the API Aparapi; pure OpenCL uses code written in C ++, and Aparapi uses Java and have to convert Java bytecode in OpenCL code before start executing.

It's also important to emphasize that, even using parallelism; it does not change the order of growth of any algorithm presented in this work. To do so, it would be necessary lots of threads, preventing a real world solution, since many threads would cause an extreme overhead, making its use not worth. Here it is used parallelism to improve performance, does not to change or improve the asymptotical order of growth of each algorithm.

Acknowledgment The authors gratefully acknowledge the academic support and research structure from the IFSP and IFSULDEMINAS. Furthermore, the financial support of IFSP, FAPESP (São Paulo Research Foundation) and CNPq (National Council for Scientific and Technological Development).

References

1. Danalis, A., Marin, G., McCurdy, C., Meredith, J.S., Roth, P.C., Spafford, K., Tipparaju, V., Vetter, J.S.: The scalable heterogeneous computing (SHOC) benchmark suite. In: ACM 3rd Workshop on General-Purpose Computation on Graphics Processing Units, pp. 63–74. ACM, New York (2010)
2. Nvidia Corporation 2015 About Cuda. https://developer.nvidia.com/about
3. Chen, H.S., Chiou, J.Y, Yang, C.Y.: Design and implementation of high-level compute on android systems. In: 2013 IEEE 11th Symposium on Embedded Systems for Real-time Multimedia (ESTIMedia), pp. 96–104. IEEE Press, New York (2013)

4. Gupta, K.G., Agrawal, N., Maity, S.K.: Performance analysis between aparapi (a parallel API) and JAVA by implementing sobel edge detection algorithm. In: 2013 IEEE National Conference on Parallel Computing Technologies (PARCOMPTECH), pp. 1–5. IEEE Press, New York (2013)
5. Joshi, S.: Leveraging Aparapi to Help Improve Financial Java Application Performance. http://developer.amd.com/wordpress/media/2012/10/Leveraging_Aparapi_to_Improve_Java_Financial_Application_Performance_after_legal_review.pdf
6. Tsuchiyama, R., Nakamura, T., Iizuka, T., Asahara, A., Miki, S.: The OpenCL Programming Book. Fixstars Corporation, California (2012)
7. Chittampally, V.R.: Aparapi Java Matrix Multiplication Example. http://vasanth-experiments.wordpress.com/2011/11/20/aparapi-java-matrix-multiplication-example
8. Sedgewick, R., Wayne, K.: Algorithms. Addison-Wesley Professional, Boston (2011)
9. Feng, W.C., Lin, H., Scogland, T., Zhang, J.: OpenCL and the 13 dwarfs: a work in progress. In: ICPE 2012 Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering, pp. 291–294. ACM, New York (2012)

An Efficient Method for the Open-Shop Scheduling Problem Using Simulated Annealing

Haidar M. Harmanani and Steve Bou Ghosn

Abstract This paper presents a simulated annealing algorithm in order to solve the nonpreemptive *open-shop scheduling problem* with the objective of minimizing the *makespan*. The method is based on a simulated annealing algorithm that efficiently explores the solution space. The method was implemented and tested on various benchmark problems in the literature. Experimental results show that the algorithm performs well on the benchmarks. The algorithm was able to find an optimum solution in many cases.

1 Introduction

Scheduling is the allocation of shared resources over time to competing activities [16]. The $n \times m$ minimum makespan job shop scheduling problem consists of n jobs that are to be processed on m machines. A job is composed of a sequence of operations that are to be processed on a particular machine in a fixed order specified by this sequence. Each machine can perform a specific operation during a fixed processing time without interruptions. A schedule for a given machine specifies the exact time interval, start time and finish time, during which each job in the shop is to be processed on that particular machine. A schedule for the shop is a set of machine schedules, one for each machine and each schedule satisfies resources constraints.

The open shop scheduling problem (OSSP) is a variation of the general job scheduling problem. The open-shop scheduling problem consists of a set of n jobs, $\mathcal{J} = \{J_1, J_2, ..., J_n\}$, that are to be processed on a set of m machines, $\mathcal{M} = \{M_1, M_2, ..., M_m\}$. Each job $J_j \in \mathcal{J}$ consists of m operations $o_{j,i}$ that have

H.M. Harmanani(✉)
Department of Computer Science and Mathematics, Lebanese American University, Byblos 1401 2010, Lebanon
e-mail: haidar@lau.edu.lb

S.B. Ghosn
Department of Computer Science, Westfield State University, Westfield, MA 01085, USA

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_102

Table 1 A 4x4 *Taillard Benchmark* instance for the OSSP

Jobs	*(Processing Time, Machine)*			
Job 1	(54, M_3)	(34, M_1)	(61, M_4)	(2, M_2)
Job 2	(9, M_4)	(15, M_1)	(89, M_2)	(70, M_3)
Job 3	(38, M_1)	(19, M_2)	(28, M_3)	(87, M_4)
Job 4	(95, M_1)	(34, M_3)	(7, M_2)	(29, M_4)

processing time $p_{j,i}$ and must be processed on M_i. At any given time, each machine can process a single operation while two operations that belong to the same job cannot be processed on different machines simultaneously. The problem is to find a schedule for the operations on the machines that minimizes the makespan, C_{max}, that is, the time from the beginning of the first operation until the end of the last operation [8]. An optimal finish time schedule is one that has the least finish time among all schedules. A schedule is said to be *non-preemptive* if for each individual machine there is at most one tuple $<i, s(i), f(i)>$ for each job i to be scheduled. A preemptive schedule is a schedule in which no restrictions are placed on the number of tuples per job per machine. The *non-preemptive* open-shop scheduling problem has been shown to be $\mathcal{NP}$-hard through a transformation to the *partition* problem [4]. It has been shown that one can establish the following lower bound for the optimal finish schedule [13]:

$$\mathcal{L} = \max \left\{ \max_i \sum_{j=1}^{m} p_{ij}, \max_j \sum_{i=1}^{n} p_{ij} \right\} \tag{1}$$

Where $1 \leq i \leq n$ *and* $1 \leq j \leq m$. The first part in equation 1 deals with the maximum completion time among all jobs that are to be processed, and simply states that the processing time for every job must necessarily take at least the sum of the processing times required to perform each of its tasks. The second part refers to the maximum completion time for the jobs allocated to a given machine. Thus, every machine must necessarily take a processing time which is at least equal to the sum of the times required to perform each of the jobs allocated to it. The optimal makespan for the open-shop scheduling problem can never be less than $\mathcal{L}$, but it is not necessarily equal to $\mathcal{L}$.

We illustrate the OSSP using the 4×4 *Taillard Benchmark* shown in Table 1. The benchmark instance consists of 4 jobs and 4 machines. Figure 1 presents a possible schedule with an optimal makespan of 193.

2 Related Work

One of the earliest and most significant efforts related to the open-shop scheduling problem was reported by Gonzalez et al. [4] who showed that the problem is

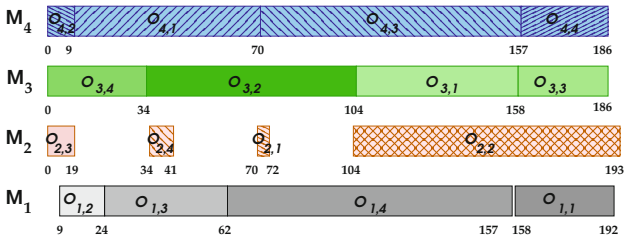


Fig. 1 Optimal schedule for problem in Table 1 with makespan = 193

$\mathcal{NP}$-complete by reducing it to the partition problem. The authors proposed a linear-time algorithm for the open-shop scheduling problem with $m = 2$, and a polynomial time algorithm to find the optimal schedule for the preemptive open-shop scheduling problem. Some attempts were made at solving the open-shop scheduling problem using branch and bound techniques [5]. Approximate methods were also reported. For example, Davis [3] proposed a genetic algorithm method that uses a memory intensive chromosome representation and very simple genetic operators. Fang et al. [6] proposed a genetic algorithm with an ordinal chromosomal representation, and adapted the approach to solve the open shop scheduling problem [15]. The algorithm used a representation that has the advantage of only producing valid schedules when altered by the genetic operators; thus allowing more accurate solutions in considerably less time. Prins [14] proposed a competitive genetic algorithm for solving the open-shop scheduling problem. Other researchers proposed hybrid genetic algorithms techniques that involve combining genetic algorithms with other optimization techniques. Khuri et al. [7] presented three different genetic algorithms approaches. One of these approaches is based on a genetic algorithm algorithm that uses a selfish gene. Another implementation was based on a hybrid GA implementation that proved to be better than the first two proposed approaches. Liaw [9] proposed an iterative improvement method for the open shop scheduling problem. The author, proposed later other efficient solution approaches based on a hybrid GA implementation [12], a tabu search technique [11], and a simulated annealing algorithm [10]. Blum [1] proposed a method based on a hybrid approach that combines ant colony optimization with beam search. Colak et al. [2] formulated the problem using an augmented neural network approach.

3 Simulated Annealing

Simulated Annealing is a global stochastic method that is used to generate approximate solutions to very large combinatorial problems. The annealing algorithm begins with an initial feasible configuration and proceeds to generate a neighboring solution by perturbing the current solution. If the cost of the neighboring solution is less than that of the current solution, the neighboring solution is accepted; otherwise,

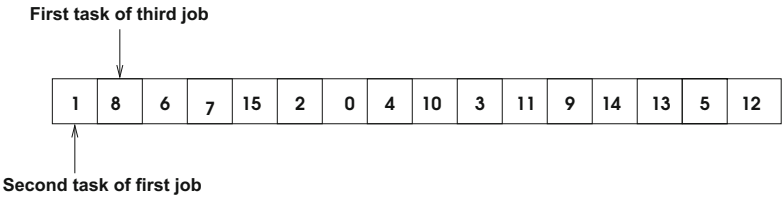


Fig. 2 Configuration representation and interpretation

it is accepted or rejected with probability $p = e^{-\frac{\Delta_C}{T}}$. The probability of accepting inferior solutions is a function of the *temperature*, T, and the change in cost between the neighboring solution and the current solution, Δ_C. The temperature is decreased during the optimization process and thus the probability of accepting a worse solution decreases as well. The set of parameters controlling the initial temperature, stopping criterion, temperature decrement between successive stages, and the number of iterations for each temperature is called the *cooling schedule*. Typically, at the beginning of the algorithm, the temperature T is large and an inferior solution has a high probability of being accepted. During this period, the algorithm acts as a random search to find a promising region in the solution space. As the optimization process progresses, the temperature decreases and there is a lower probability of accepting an inferior solution. The algorithm behaves like a down hill algorithm for finding the local optimum of the current region.

4 Simulated Annealing OSSP

The proposed algorithm starts with a number of machines, and a set of jobs and corresponding tasks. The algorithm generates, through a sequence of transformations, a set of compact schedules. The key elements in implementing the annealing test scheduling algorithm are: 1) the definition of the initial configuration, 2) the definition of a neighborhood on the configuration space and a perturbation operator exploring it; 3) the choice of the cost function; and 4) a cooling schedule. In what follows, we describe the simulated annealing algorithm with reference to the benchmark in Table 1.

4.1 Configuration Representation

In order to solve the OSSP, we propose the configuration shown in Figure 2. The representation is based on a vector where every cell corresponds to an operation in the schedule. The values are a permutation of integer values between 0 and $(m \times n) - 1$,

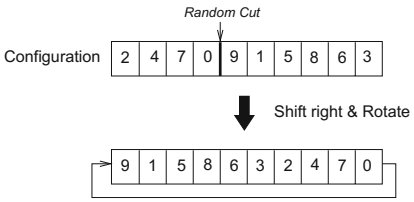


Fig. 3 Shift and Rotate Transformation

and where m is the number of machines and n the number of jobs. A value V is interpreted as the $(V \ mod \ m)$ operation of the $(V \ div \ m)$ job.

4.2 Initial Configuration

The initial configuration is created randomly by selecting the values to be unique permutations between 0 and $(n \times m) - 1$, where n is the number of jobs and m is the number of machines. The algorithm ensures that all $n \times m$ elements in the initial configuration have different values.

4.3 Objective Function

The objective is to find a schedule for the operations on the machines that minimizes the makespan, C_{max}. That is, the time from the beginning of the first operation until the end of the last operation. The objective function is given by the following:

$$F = \frac{1}{C_{max}} \tag{2}$$

4.4 Neighborhood Transformation

The algorithm uses two neighborhood transformations in order to explore the solution space. However, it should be noted that when scheduling operations, the algorithm applies an as soon as possible scheduling strategy in order to minimize idle gaps in the middle of the schedule. Thus, the algorithm iterates over the neighborhood solution and finds the earliest idle time slot that is large enough to fit the operation to

Algorithm 1. Annealing Algorithm

Input: (S_0, T_0, α, β, M, MaxTime)

```
1: T = T0                                   ▷ T0 is the initial temperature
2: CurSol = S0
3: CurCost = Cost(CurSol)
4: BestCost = CurCost
5: Time = 0
6: do
7:     Metropolis(CurSol, CurCost, BestSol, BestCost, T, M)
8:     Time = Time + M
9:     T = α × T
10:    M = β × M
11: while (Time ≥ maxTime || T > 0.001)
```

Algorithm 2. Metropolis

Input: (CurSol, CurCost, BestSol, BestCost, T, M)

```
1: do
2:     NewSol = Neighborhood(CurSol)
3:     NewCost = Cost(NewSol)
4:     Δ_Cost = NewCost - CurCost
5:     if Δ_Cost < 0 then
6:         CurSol = NewSol
7:         if NewCost < BestCost then BestSol = NewSol
8:         end if
9:     else
10:        if Random < e^(-Δ_Cost / T) then CurSol = NewSol
11:        end if
12:    end if
13:    M = M - 1
14: while M ≠ 0
```

be scheduled. If no available time slot is large enough then the operation is scheduled after the last operation that is scheduled on that machine.

4.4.1 Swap

The swap operator selects two *random* positions between 0 and $(m \times n) - 1$, and swaps their positions. Since a solution is valid to the extent that it represents a feasible schedule, swapping two operations will alter the order of the execution of operations in the schedule resulting in a new feasible schedule that is 'similar' to the original schedule. The number of swaps we perform in each call to the neighborhood function is controlled by a tuning parameter that is kept very small in order not to drastically alter the solution in a single annealing iteration.

4.4.2 Shift and Rotate Transformation

The shift and rotate transformation is used in order to perturb the solution while maintaining some of the attributes of the original solution while guaranteeing that the resulting solution is feasible. The transformation selects a random position between 0 and $(m \times n) - 1$. The neighborhood solution is next shifted and rotated either right or left as determined by a random variable. Figure 3 illustrates a shift right and rotate by seven.

4.5 Cooling Schedule

The cooling schedule is the set of parameters that control the initial temperature, the stopping criterion, the temperature decrement between successive stages, and the number of iterations for each temperature. The idea is to initially set the initial temperature, T_0, to a high number, which causes the algorithm to perform a random walk in the solution space by moving to every newly-created solution regardless of how good it is. This implies that the probability of accepting a solution should be 1 or close to 1:

$$P(T_0) = \frac{\text{Number of moves accepted}}{\text{Total Number of Moves Attempted}} \tag{3}$$

We determine the initial temperature T_0 by initially setting it to a small value and computing $P(T_0)$, if the value isn't close to 1 then we keep gradually incrementing the temperature by multiplying it by a constant K $(K > 1)$ and repeating this procedure until we reach a proportion which is 1 or very close to 1. The temperature we use to obtain that proportion becomes T_0. This procedure models the process of heating the material until all its atoms are completely free. After performing this process we empirically determined the optimal initial temperature (T_0) to be 400.

Another very important parameter we need to tune for our cooling schedule is the rate at which we decrease the temperature. The cooling rate is determined by α which is the multiplier that is used to decrease the temperature. Ideally, we would like to decrease the temperature as slowly as possible until to eventually reach 0, at which point the algorithm is doing nothing more than plain Hill-Climbing. It follows that we should select a value of α that is very close to 1. In our empirical testing we have determined that the optimal value range is $0.99 \leq \alpha \leq 0.999$, and for most samples setting α to 0.999 seems to give us the best results. Since the approach we have followed for our annealing implementation is a metropolis approach, the algorithm will stop either when the temperature reaches 0.001 or until the *MaxTime* is reached.

The *MaxTime* parameter is problem dependent as bigger problems require more running time in order to achieve good results. We have also empirically determined the metropolis parameters, M and β to be 5 and 1.05, respectively.

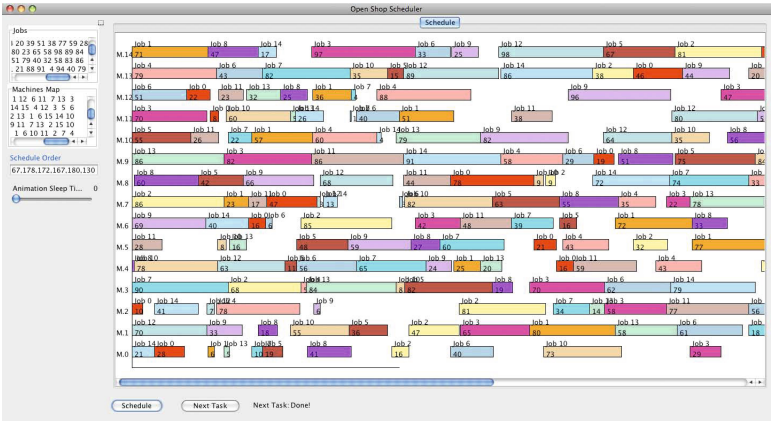


Fig. 4 Schedule for an instance of problem 20 × 20 with a makespan of 1292

5 Results

The algorithm was implemented using C++ and tested on the *Taillard Benchmark* problem instances. Figures 5(a) and 5(b) show plots representing the results we obtained for a small 4 × 4 problem instance and a large 15 × 15 problem instance. Tables 2, and 3 show the results obtained after running each of the *Taillard Benchmark* instances using the determined parameters. It can be observed that the algorithm was able to find the optimum solutions in many benchmark instances. The execution time was less than five minutes for all benchmark instances. The neighborhood

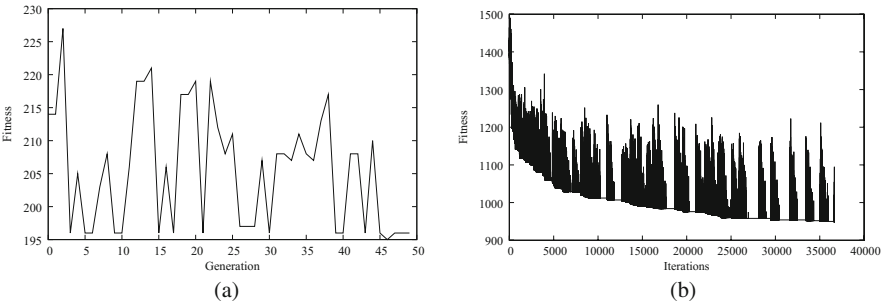


Fig. 5 (a) Plot for Taillard 4 x 4 - 0; b) Plot for Taillard 15 x 15 - 0

Table 2 Results Comparison for Taillard's instances 4x4, 5 x5, 7x7, and 10x10

Problem Instance	Optimal	Ours	Khoury [7]	Hybrid GA [7]	Fang [15]	Prins [14]
4 x 4 - 1	193	193	193	213	193	193
4 x 4 - 2	236	236	236	240	236	239
4 x 4 - 3	271	271	271	293	271	271
4 x 4 - 4	250	250	250	253	250	250
4 x 4 - 5	295	295	295	303	295	295
4 x 4 - 6	189	189	189	209	189	189
4 x 4 - 7	201	201	201	203	201	201
4 x 4 - 8	217	217	217	224	217	217
4 x 4 - 9	261	261	261	281	261	261
4 x 4 - 10	217	217	217	230	217	221
5 x 5 - 1	300	300	301	323	300	301
5 x 5 - 2	262	262	262	269	262	263
5 x 5 - 3	323	323	331	353	323	335
5 x 5 - 4	310	310	N/A	N/A	310	316
5 x 5 - 5	326	326	N/A	N/A	326	330
5 x 5 - 6	312	312	312	327	312	312
5 x 5 - 7	303	303	N/A	N/A	303	308
5 x 5 - 8	300	300	N/A	N/A	300	304
5 x 5 - 9	353	353	353	373	353	358
5 x 5 - 10	326	326	326	341	326	328
7 x 7 - 1	435	435	438	447	435	436
7 x 7 - 2	443	443	455	454	443	447
7 x 7 - 3	468	468	N/A	N/A	468	472
7 x 7 - 4	463	463	N/A	N/A	463	463
7 x 7 - 5	416	416	N/A	N/A	416	417
7 x 7 - 6	451	451	N/A	N/A	451	455
7 x 7 - 7	422	422	443	450	422	426
7 x 7 - 8	424	424	N/A	N/A	424	424
7 x 7 - 9	458	458	465	467	458	458
7 x 7 - 10	398	398	405	406	398	398
10 x 10 - 1	637	637	667	655	637	637
10 x 10 - 2	588	588	N/A	N/A	588	588
10 x 10 - 3	598	598	N/A	N/A	598	598
10 x 10 - 4	577	577	586	581	577	577
10 x 10 - 5	640	640	N/A	N/A	640	640
10 x 10 - 6	538	538	555	541	538	538
10 x 10 - 7	616	616	N/A	N/A	616	616
10 x 10 - 8	595	595	N/A	N/A	595	595
10 x 10 - 9	595	595	627	598	595	595
10 x 10 - 10	596	596	623	605	596	596

transformations were applied in every iteration with a probability 0.75 for the swap operator and 0.25 for the shift and rotate transformation. We show the solution for an instance of the 20×20 benchmark in Figure 4 using our GUI interface.

Table 3 Results Comparison for Taillard's instances 15x15 and 20x20

Problem Instance	Optimal	Ours	Khoury [7]	Hybrid GA [7]	Fang [15]	Colak [2]	Prins [14]
15 x 15 - 1	937	937	967	937	937	937	937
15 x 15 - 2	918	918	N/A	N/A	918	918	918
15 x 15 - 3	871	871	904	871	871	871	871
15 x 15 - 4	934	934	969	934	934	934	934
15 x 15 - 5	946	946	N/A	N/A	946	946	946
15 x 15 - 6	933	933	N/A	N/A	933	933	933
15 x 15 - 7	891	891	N/A	N/A	891	891	891
15 x 15 - 8	893	893	928	893	893	893	893
15 x 15 - 9	899	899	N/A	N/A	899	901	899
15 x 15 - 10	902	902	N/A	N/A	902	902	902
20 x 20 - 1	1155	1155	1230	1165	1155	1115	1115
20 x 20 - 2	1241	1241	N/A	N/A	1241	1242	1241
20 x 20 - 3	1257	1282	1292	1257	1257	1173	1257
20 x 20 - 4	1248	1274	N/A	N/A	1248	1248	1248
20 x 20 - 5	1256	1289	1315	1256	1256	1256	1256
20 x 20 - 6	1204	1204	1266	1207	1204	1204	1204
20 x 20 - 7	1294	1294	N/A	N/A	1294	1294	1294
20 x 20 - 8	1169	1169	N/A	N/A	1169	1173	1169
20 x 20 - 9	1289	1307	1339	1289	1289	1289	1289

6 Conclusion

In this paper, we presented a simulated annealing algorithm for minimizing the makespan in the nonpreemptive open shop scheduling problem. The method was implemented and tested on various benchmark problems in the literature. The algorithm was able to find an optimum solution in many cases. The computing times did not exceed five minutes for the largest benchmark example.

References

1. Blum, C.: Beam-aco-hybridizing ant colony optimization with beam search: an application to open shop scheduling. Computers and Operations Research **32**(6), 1565–1591 (2005)
2. Colak, S., Agarwal, A.: Non-greedy heuristics and augmented neural networks for the open-shop scheduling problem. Naval Research Logistics **52**(7), 631–644 (2005)
3. Davis, L.: Job shop scheduling with genetic algorithms. In: Proceedings of the International Conference on Genetic Algorithms and their Applications, pp. 136–140 (1985)
4. Gonzalez, T., Sahni, S.: Open shop scheduling to minimize finish time. Journal of the Association for Computing Machinery **23**(4), 665–679 (1976)
5. Gueret, C., Prins, C.: Classical and new heuristics for the open-shop problem: a computational evaluation. European Journal of Operational Research **107**, 306–314 (1998)
6. Fang, H.-L., Ross, P., Corne, D.: A promising genetic algorithm approach to job shop scheduling, rescheduling and open-shop scheduling problems. In: Proceedings of the Fifth International Conference in Genetic Algorithms, pp. 375–382 (1993)

7. Khuri, S., Miryala, S.R.: Genetic algorithms for solving open shop scheduling problems. In: Proceedings of the 9th Portuguese Conference on Artificial Intelligence. Lecture Notes in Computer Science, pp. 357–368 (1999)
8. Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G., Shmoys, D.B.: Sequencing and scheduling: Algorithms and complexity. Handbooks in Operations Research and Management Science: Logistics of Production and Recovery **4**, 445–522 (1993)
9. Liaw, C.-F.: An iterative improvement approach for the nonpreemptive open shop scheduling problem. European Journal of Operational Research **111**, 509–517 (1998)
10. Liaw, C.-F.: Applying simulated annealing to the open shop scheduling problem. IIE Transactions, Scheduling and Logistics **31**(5), 457–465 (1999)
11. Liaw, C.-F.: A tabu search algorithm for the open shop scheduling problem. Computers and Operations Research **26**(2), 109–126 (1999)
12. Liaw, C.-F.: A hybrid genetic algorithm for the open shop scheduling problem. European Journal of Operational Research **124**, 28–42 (2000)
13. Pinedo, M.: Scheduling: Theory, Algorithms, and Systems. Prentice Hall (1995)
14. Prins, C.: Competitive Genetic Algorithms for the Open-Shop Scheduling Problem. Mathematical Methods of Operations Research **52**(3), 389–411 (2000)
15. Fang, H.-L., Ross, P., Corne, D.: A promising hybrid GA/heuristic approach for open-shop scheduling problems. In: Proceedings of the 11th European Conference on Artificial Intelligence, pp. 590–594 (1994)
16. Yamada, T., Nakano, R.: Job-Shop Scheduling. Genetic algorithms in Engineering Systems, chap. 7, pp. 134–160. IEE (1997)

Schematizing Heidegger

Sabah Al-Fedaghi

Abstract Philosophy is typically discussed and taught by abstract discussion of texts; however, a considerable body of research suggests that this method is unsuited to professionals and students who struggle to navigate philosophical writings. Diagrams are used in diverse areas of study to depict knowledge and to assist in understanding of problems. This paper aims to utilize schematic representation to facilitate understanding of certain philosophical works. The paper employs schematization as an apparatus of specification for clarifying philosophical language by describing ideas in a form familiar to computer science. Specifically, it is an attempt, albeit tentative, to schematize Martin Heidegger's philosophical approach. His writings are notoriously difficult to understand, and the high-level representation described here seems a viable tool for enhancing the relationship between philosophy and computer science, especially in computer science education.

Keywords Conceptual model · Philosophy · Computer science and engineering · Diagrams

1 Introduction

According to Farrow [1],

> Philosophical pedagogies are typically based on abstract discussion of texts, and have remained largely unchanged throughout the history of the subject. However, there is a considerable body of research which suggests that this is unsuited to some learning styles and may discourage some students from prolonged study. Many prefer to learn through visual cues and models alongside engaging with literature resources.

S. Al-Fedaghi(✉)
Computer Engineering Department, Kuwait University,
P.O. Box 5969, Safat 13060, Kuwait
e-mail: sabah@alfedaghi.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_103

This paper aims to utilize schematic representation to facilitate understanding of certain philosophical works. The paper employs *schematization* as an apparatus of specification for clarifying philosophical language by describing philosophical ideas in a form familiar to computer science. Specifically, it is an attempt, albeit tentative, to schematize Heidegger's ontological approach. It targets professionals and students in fields outside of philosophy such as computer science and engineering, who often look to sources in philosophy for design ideas or for a critical framework for practice. According to Schwill [2],

> It is necessary that students obtain a sketch of the fundamental ideas, principles, methods and ways of thinking of computer science. Only these fundamentals seem to remain valid in the long term and enable students to acquire new concepts successfully during their professional career.

Without loss of the general applicability of the proposed diagrammatic method, the paper focuses on Martin Heidegger, generally recognized as one of the most important philosophers of the twentieth century. "Heidegger was one of the last philosophers, that is, one of the last pure philosophers" [3]. His thinking has contributed to such diverse fields as phenomenology, existentialism, hermeneutics, political theory, psychology, and theology [4]. Heidegger is one of few well-known philosophers who have written specifically for an audience of architects [5], and his works have been included in some computer science courses [6]. In engineering science, Heidegger's thought introduces ways in which engineers should question technology and highlights some of the hazards and injustices associated with technology and its subtle sociological and psychological influences. These issues are important to engineering ethics and design and in developing ways to confront "a narrowly rationalistic technological" [7] world.

However, "Heidegger is one of the most controversial thinkers of the twentieth century. His writings are notoriously difficult; they both require and reward careful reading" [8]. "For the student, Martin Heidegger's *Sein und Zeit (Being and Time)* stands as one of the most difficult texts in Western philosophy in large part because of its unusual and idiosyncratic language" [9].

In the next section, the paper begins by reviewing the modeling tool, called the Flowthing Model (FM) [10,11], to be used in interpreting Heidegger's notions through schematic representations. *Schematization* is one of the main tools used in computer science to "read" a system, e.g., flowcharts, UML, and SysML, and it lends itself to a type of flowcharting of Heidegger' s philosophical concepts in a systematic way. The result is expression of philosophical thought in computer science language.

Given the number of papers produced over the years on Heidegger's philosophy, a separate review of literature is not necessary. Instead, quotations about notions under discussion will be interwoven into the appropriate text in the paper. We assume that the reader is at least slightly familiar with Heidegger's concepts and knows basic philosophical terminology such as *ontology*. Additionally, because of space limitation, only a few of Heidegger's ideas are discussed to demonstrate the viability of the diagramming method.

2 Flowthing Model

The Flowthing Model (FM) was inspired by the many types of *flows* that are found in diverse fields, including information flows, signal flows, and data flows in communication models. This notion has nothing to do with a psychological concept of flow, of being totally immersed. FM produces a diagrammatic schema that uses *flowthings* to represent a range of items, for example, electrical, mechanical, chemical and thermal signals, blood, food, concepts, pieces of data, and so on. Yet, *flow* in FM designates not only mobility, but also creation and transformation.

Flowthings are defined as *what can* be created, released, transferred, processed, and/or received (see Fig. 1). The (abstract) machine shown in Fig. 1 is a generalization of the typical input-process-output model used in many scientific fields (Fig. 2).

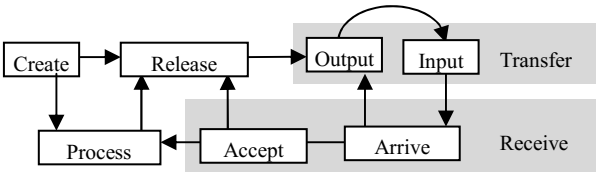


Fig. 1 Flow (abstract) machine

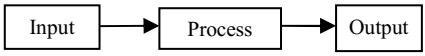


Fig. 2 Input-process-output model

FM depicts *flow by using flow machines* (Fig. 1) comprising up to six stages (states). The term machine is used here in the sense of *system or organism*. The machine is the conceptual fiber used to handle flowthings (to change them through stages) from inception or arrival to de-creation or transmission to outside the system. The flowthing is ready to be handled (released, processed, …); the flow machine has the capability to do the handling.

Machines can be embedded in a network of assemblies and hierarchies called *spheres*.

The stages in Fig. 1 can be described as follows:

Arrive: A flowthing reaches a new machine (curved arrow in Fig. 1).

Accepted: A flowthing is permitted to enter the machine. If arriving things are always accepted, *Arrive* and *Accept* can be combined as a *Received* stage.

Processed (changed): The flowthing goes through some kind of transformation that changes it without creating a *new* thing, but the change might trigger the creation of new flowthings.

Released: A flowthing is marked as ready to be transferred outside the machine.

Transferred: The flowthing is transported somewhere from/to outside the machine.
Created: A new flowthing is born (created) in a machine and its sphere. It is the *becoming* of that which has no prior being (appearance of a new thing in the sphere), e.g., a new actor appears in the context of a scene, not as a person coming from outside, but by suddenly being in the spotlight on a previously dark place on the stage.

In general, a flow machine is thought to be an abstract machine that receives, processes, creates, releases, and/or transfers things. The stages in this machine are mutually exclusive for atomic flowthings; that is, they are indivisible, nor do they spread over two stages. Suppose that a *car* is being created (manufactured); it cannot be released from the assembly line before the end (e.g., say at the stage where it is just a body with some electrical wiring). It must become *a car* and fulfill certain conditions before it can be released.

An additional stage of *Storage* can also be added to any machine to represent the storage of things (memory); however, storage is not an exclusive stage because there can be *stored processed* things, *stored created* things, etc.

FM also uses the notions of *spheres and subspheres*. These are the network environments and relationships of machines and submachines. Multiple machines can exist in a sphere if needed. A sphere can be a person, an organ, an entity (e.g., a company, a customer), a location (a laboratory, a waiting room), a communication medium (a channel, a wire). A flow machine is a subsphere that embodies the flow; it itself has no subspheres.

FM also utilizes the notion of *triggering*. Triggering is the activation of a flow, denoted in FM diagrams by a *dashed arrow*. It is a (causative) dependency among flows and parts of flows. A flow is said to be triggered if it is activated by another flow (e.g., a flow of electricity triggers a flow of heat), or activated by another point in the flow. Triggering can also be used to initiate events such as starting up a machine (e.g., by remote signal). Multiple machines captured by FM can interact by triggering events related to other machines in those machines' spheres and stages.

Example: According to Harman [12, p. 13], "The tool's meaning is the visible term termination of its underground action. Just as the meaning of:

- a signal-arrow is the region to which it alerts us by pointing elsewhere,
- a word evokes its meaning by deflecting attention away from itself,
- cable and pillars "mean" the bridge-system by "vanishing into it."

An FM representation of a signal arrow is shown in Fig. 3.

The real arrow appears (is created) in the sphere under consideration (circle 1) to trigger (2) the creation of an arrow sign (3). The sign flows to the sphere of an observer (4) to make him notice (process, 5) what's pointed to (6), the referent, which appears in the situation for the first time (7). The referent creates a sign (8) (e.g., perceptual data) that flows to the observer sphere (9).

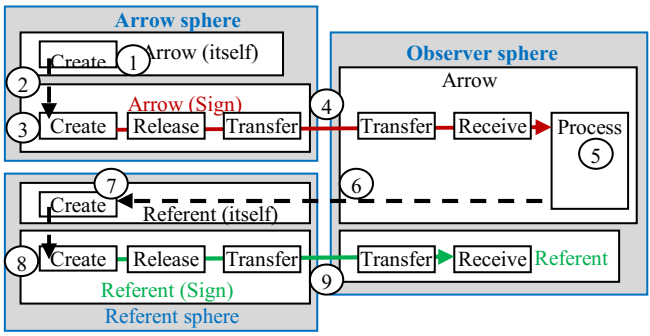


Fig. 3 The arrow as a tool

Note how the schema completely describes *a signal-arrow in the region where it alerts us by pointing elsewhere*. Similar diagrams can be drawn to represent the other two cases described by Harman [12].

3 Present-at-Hand versus Ready-to-Hand

Contrasting *present-at-hand* with *ready-to-hand* was one of Heidegger's major concepts in the analysis of being. Systems engineers "seemed to regard the distinction as important for understanding of the system specification and design phase of system development… this issue should be investigated and may well prove to very significant" [13]. It seems that present-at-hand and ready-to-hand provide "an insight" that "may prove valuable in understanding the design of systems" [13].

Adopting Harman's [12] interpretation, *present-at-hand* and *ready-to-hand* can be regarded as signifying modes of being and marking two irreducible aspects of every object. Present-at-hand refers to conceiving *being* in terms of tangible "objects of some sort of discussion or perception, and is recognized by a specific shape and color and texture" [12]. "Ready-to-hand is the mode of beings in themselves: an invisible thing radiating its being is dissolved into the world. This is a departure from the Aristotelian view of the *thing* as a single distinct concrete thing, a single intersection of form and matter" [12].

Harman [12] uses Richard Rorty, Professor of Philosophy at Stanford University, as an example to illustrate the difference between the two modes.

To emphasize the point further, it might be noted that even Richard Rorty himself is simultaneously *present-at-hand* and *ready-to-hand*.

- We all know certain merely "extant," present-at-hand facts about him: that he wrote such and such books, that he has a particular hair and eye color, that he has recently left Virginia for Stanford, and so forth. These are nothing but present-at-hand properties, even if it can justly be claimed that Rorty can never be present-at-hand in the same way as a broken window or a piece of paper.

- But by the same token, Rorty is also no less *ready-to-hand*... it means that prior to any list of properties that can be drawn up, Richard Rorty is *what he is*. The world would be a different place if he did not *exist*: different for his vast network of readers, colleagues, family, friends,... [12]

Fig. 4 illustrates these two modes of Richard Rorty. The upper diagram shows a real *object* characterized by a specific shape and color and the type of object described by means of *categories*. "A present-at-hand object is simply there, situated in space/time, bearing the sorts of properties that make it a suitable object of the mathematical-physical sciences [and object oriented programming]" [14].

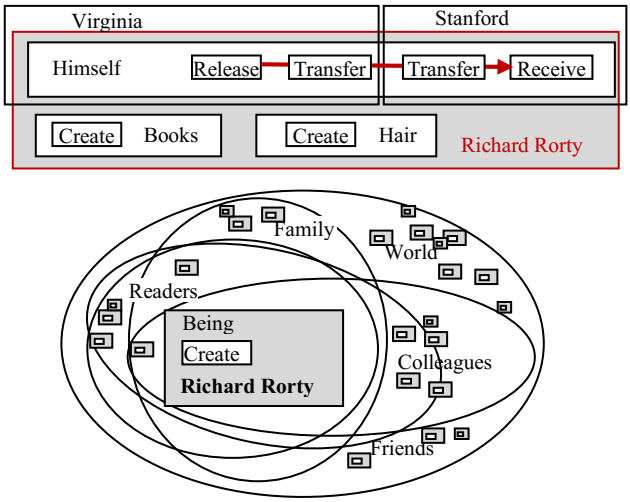


Fig. 4 Present-at-hand (upper) and ready-to-hand (lower)

In the lower diagram, Richard Rorty is not approached categorically; rather his being is described existentially. "The world would be a different place if he did not exist" [12]. "Ready-to-hand entities are constitutive of our ontological structure as being-in-the-world. Only based on this structure, therefore, can we go on to reveal objects as distinct from our practices and concern; that is, only on account of readiness-to-hand is there presence-at-hand." [15].

4 The Thing Jug

The term thing is used as a counterpoint to the term object. For Heidegger, thing is a good term and object a bad one [16]. Take for example, the essence of a thing such as a jug,

> When we fill the jug, the pouring that fills it flows into the empty jug. The emptiness, the void, is what does the vessel's holding. The empty space, this

> nothing of the jug, is what the jug is as the holding vessel.... But if the holding is done by the jug's void, then the potter who forms sides and bottom on his wheel does not, strictly speaking, make the jug. He only shapes the clay. No—he shapes the void ... The vessel's thingness does not lie at all in the material of which it consists, hut in the void that it holds. [17]

Accordingly, Fig. 5 diagrams the jug as a thing as described by Heidegger. All the following sentences are taken from Heidegger [17]; the numbers in parentheses refer to numbered circles in the figure.

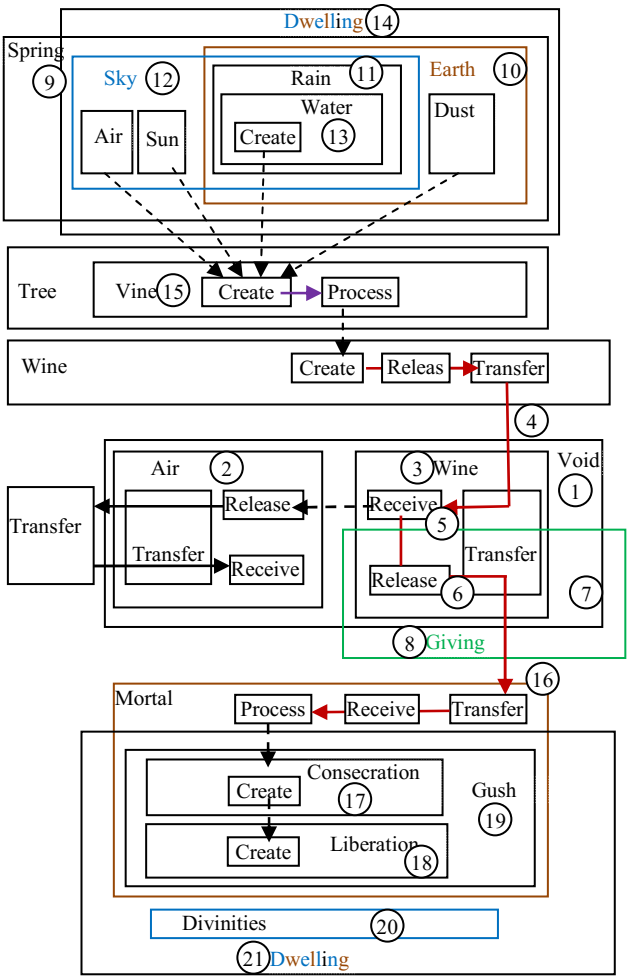


Fig. 5 The *thing* jug

When we fill the jug, the pouring that fills it flows into the empty jug. The emptiness, the void (1), is what does the vessel's holding.
The jug is filled with air (2).
When we pour wine (3) into the jug, the air that already fills the jug is simply displaced by a liquid.
How does the jug's void hold? It holds by taking what is poured in (4 and 5).
To pour from the jug is to give (6 and 7). Outpouring…, first constitutes the full presence of giving (8): the poured gift.
The spring (9) stays on in the water of the gift. [Next, Heidegger explains how the spring stays in the wine]
In the spring the rock dwells, and in the rock dwells the dark slumber of the earth (10), which receives the rain (11) and dew of the sky (12).
In the water (13) of the spring dwells (14) the marriage of sky and earth. It stays in the wine given by the fruit of the vine (15). In the jugness of the jug, sky and earth dwell.
The gift of the pouring out is drink for mortals (16). But, the jug's gift is at times also given for consecration (17).
The gift of the pouring now is neither given in an inn nor is the poured gift a drink for mortals. The gift of the outpouring as libation (18)...
The consecrated libation is what our word for a strong outpouring flow, "gush," (19) really designates: gift and sacrifice.
In the gift of the outpouring, mortals and divinities (20) each dwell (21) in their different ways.

Wow, amazing! What a jug! The diagram follows Heidegger's words, producing an astonishing structure that covers everything in the world. This schematization of the text seems to amplify its underlying structure.

5 The Hammer

Analysis of *tools* and reflections on *technology* are topics on which Heidegger has elaborated. "When he speaks of 'tools', his analysis holds for trees and monkeys no less than for hammers" [18]. In this context, *Ready-to-hand* refers to tools that remain concealed from view insofar as they function effectively. *Present-at-hand* refers to objects like broken tools that become obtrusive once they no longer function effectively [18]. "Objects can withdraw into their hidden underground action or they can become objects of explicit awareness" [18].

In *Being and Time*, Heidegger uses the example of a *hammer* to describe its functions in these two distinct ways [19]:

- When we use the hammer, it becomes *ready-to-hand*, the hammer is ready to be put to work. In this case, when we are deploying it as a tool it begins to *disappear* from view. It *withdraws* from our conscious perception and our attention moves from the hammer towards the nails being driven into the wood.

- When something goes wrong with the process of hammering and the nails stop disappearing into the wood, the hammer *re-appears* in our conscious perception when it refuses to do its job.

The FM representation is shown in Fig. 6. In the sphere of Attention 1, the hammer is brought in (circle 1), then used (2) to trigger (3) activity on the task (4), which, in turn, in the context of Attention 2, involves the presence (5) of a nail that is being driven into wood (6). Circle 7 indicates a return to Attention 1 because of a problem in the task.

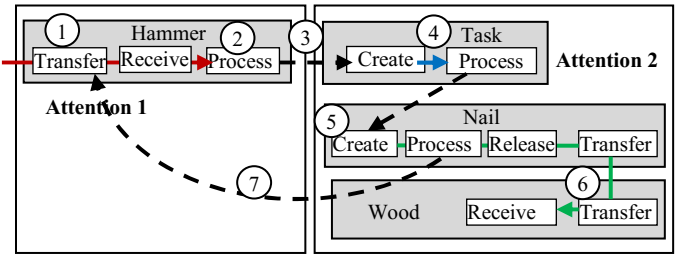


Fig. 6 Attentions to hammer and task

Hale [19] explains these Heidegger notions as follows:

> The bodily grasp of a task provides a kind of persistent *background awareness*, a process of monitoring that allows us to make sure the tool is still 'on track' – to make continual minute adjustments to… keep the hammer-head hitting the nail, instead of crashing down onto my thumb… the tool becoming incorporated (literally) into an extended 'body-schema', as a kind of prosthetic bodily extension that allows me to experience the world through it. [Italics added].

Hale [19] argues that, in the case of tools and equipment, there is no such thing as total withdrawal. The *ready-to-hand* of Heidegger's tool analysis is a continual monitoring through a kind of habitual or background awareness.

This concept can be represented in FM by adding the new sphere of Awareness with two subspheres: Background Awareness and Foreground Awareness, as shown in Fig. 7. The figure can be used to illustrate this disagreement over the meaning of *withdraw* from conscious perception.

Additionally, according to Hale [19], Heidegger describes the tool as being *un-ready-to-hand* in certain cases where the situation goes wrong:

- The case Heidegger calls *Obstinacy*, when things get in the way of doing the job, e.g., space must be cleared on the workbench before we can even begin the new activity.
- The case Heidegger calls *Obtrusiveness*, when the needed tool is absent.

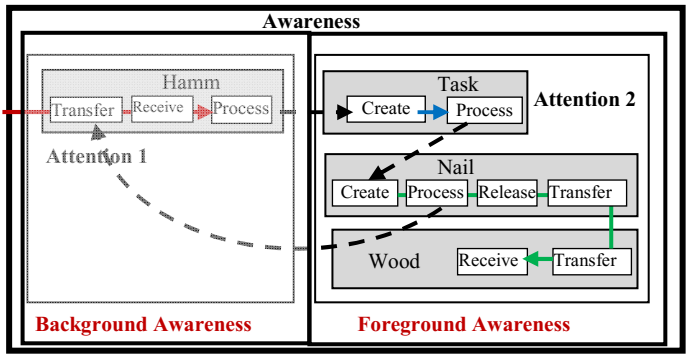


Fig. 7 Background and foreground

Assuming the hammer needs to be located and the workbench needs to cleared, Fig. 8 shows the resultant representation. At Attention 1, a search is initiated for the hammer (circle 1); at the end, the hammer is found on the pegboard (2) and flows to the Attention 3 (3). Then clearing of the workbench is started (4), followed by performing the task (5). Many scenarios can be drawn, including searching for a nail and dropping a nail while hammering.

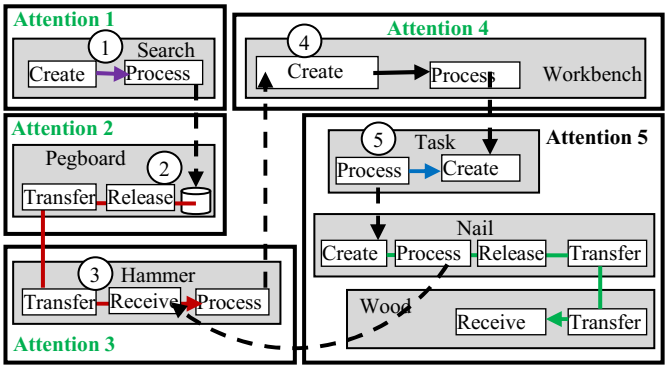


Fig. 8 Attentions for finding the hammer, clearing the workbench, then doing the task

6 The Dasein

The following is a paraphrasing of Harman's [12] description of Dasein as the *tool-being* of any entity. Note that Harman's aim in describing Dasein is different from our aim, which is merely to focus on the issue of explaining Dasein and its being.

Dasein is a being (human) that is irreducible to anything present-at-hand in a psyche. Dasein has special properties that are not shared by machines or animals.

Human Dasein (circle 1 in Fig. 9) is a being that deals with its existential status with respect to its own *being* (2 in Fig. 9). It is the entity posing (creating) the *question of being* (circle 3) [quoting Heidegger]. Dasein is an entity concerned with its own being, the entity that has an understanding of being. But to be is not the same as to understand being.

The crucial factor here is *the being of the question* (see Fig. 10). The *question of being* cannot be elucidated until the *meaning of being* itself has somehow been clarified, prior to any special description of Dasein.

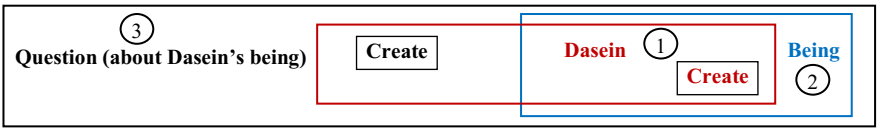


Fig. 9 Dasein and question about own being

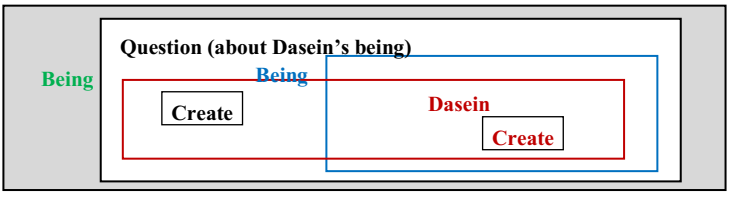


Fig. 10 The being of the question

Dasein unifies the past, present, and future, referred to by Heidegger as the "ecstases of temporality." Dasein is a temporal mode of being, and its authentic temporality is the being [20]. Harman [12], "considering the human observer as the subject matter of the temporal analysis," gives the following example:

> Imagine that an unintelligent but exceedingly tranquil person is chained to a pillar somewhere on the earth, motionless. Placed before him is an immense but well camouflaged machine, a device that controls all aspects of his environment— the unchanging temperature and scent of the air, the uniform amount of light that shines on him from a fixed angle, the constant sonorous drone in the air, the steady infusions of intravenous liquid that provide his nourishment. Furthermore, we can assume that a second blend of chemicals introduced into his veins keeps his moods at an optimum level of stability

This situation can be applied to Dasein as a *being*. Instead of defining Dasein as an entity that poses the *question of being* shown (Fig. 9), we can define it as an entity that processes its being (Fig. 11), since questioning something is a type of processing of it. Accordingly, the chained person described by Harman [12] is an entity that "receives" and consumes its being without processing (e.g., questioning) it (Fig. 12).

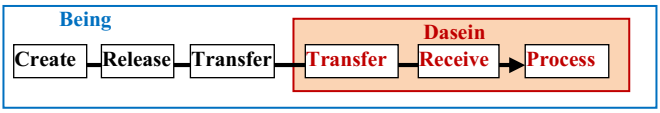


Fig. 11 Dasein and processing about own being

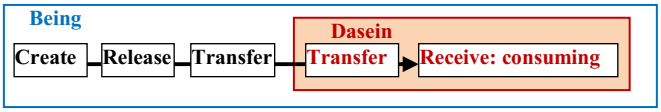


Fig. 12 Dasein that receives being

Harman [12] continues the scenario as follows:

> Imagine that it goes on in this way for years, until suddenly; the machine enters an era of gradual decay. Each day, two or three of its thousands of functions cease to operate, leading to various failures in its workings. Only now will the drugged man begin to *notice* the "temporal" structure of the machine and of his world.
>
> My question about this scenario is as follows: would it ever occur to such a person to consider himself as the source of the ecstatic-temporal nature of this environment?
>
> Obviously, faced with the sudden decay of his environment, he would have to regard the machine's own degradation as the source of all temporal ecstases. (Italic added)

It seems that this "noticing" implies processing of his being (Fig. 11).

Alas, it seems that the diagramming scenario has lured this author, a nonphilosopher, into getting involved in a philosophical analysis. This is a good indicator of the potential worth in pursuing the FM representational method.

7 Conclusion

This paper has attempted to employ schematization *to understand* philosophical concepts. While the method is applicable to several philosophical works, here it focuses specifically on a portion of Heidegger's philosophy. The resulting representation seems to introduce new ways to discuss the meanings embedded in philosophy. This initial attempt points to the viability of this method in this context and is worthy of pursuit. Further work would expand the investigations to represent additional philosophical notions to explore the potential of this diagrammatic method.

References

1. Farrow, R.: Visual and philosophical pedagogies. In: 3rd Visual Learning Conference, Budapest (2012)
2. Schwill, A.: Fundamental Ideas of Computer Science. Bull. Eur. Assoc. Theor. Comput. Sci. **53** (1994)
3. Tonkin, C.C.: The Institutional Essence of Learning as an Anthropocentric Praxis Following Heidegger. Ph.D. dissertation, Univ. Sydney (1999)
4. Korab-Karpowicz, Q.J.: Martin Heidegger (1889–1976). Internet Encyclopedia of Philosophy (IEP) (2015). (ISSN 2161-0002) (access)
5. Sharr, A.: Heidegger for Architects. Routledge (2007). http://m.friendfeed-media.com/e4821608683fc2184a083b3814f288f60584b202
6. Dreyfus, H.L.: Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian. Philos. Psychol. **20**(2), 247–268 (2007)
7. Dias, W.P.: Heidegger's Relevance for Engineering: Questioning Technology. Sci. Eng. Ethics **9**(3), 389–396 (2003)
8. Mulhall, S.: Heidegger and Being and Time, 2nd edn. Routledge (2005)
9. Tietz, J.: An Outline and Study Guide to Martin Heidegger's Being and Time. Humanities Online (2001). ISBN 978-3-934157-08-8
10. Al-Fedaghi, S.: Personal information flow model for P3P. In: W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, Ispra, Italy, October 17–18, 2006
11. Al-Fedaghi, S.: Crossing privacy, information, and ethics. In: 17th International Conference, Information Resources Management Association (IRMA 2006), Washington, DC, May 21–24, 2006. Republished in Emerging Trends and Challenges in Information Technology Management, Khosrow-Pour, M. (ed.), IGI, Hershey, PA (2006)
12. Harman, G.: Tool-being: Elements in a Theory of Objects. Ph.D. dissertation, Dept. Philosophy, College of Liberal Arts and Sciences, DePaul Univ., Chicago (1999)
13. Ferris, T.L.J.: Exploration of the application of 'ready-to-hand' and 'present-at-hand' in the design of systems. In: 9th ANZSYS Conference, Melbourne, November 18–20, (2003)
14. McDaniel, K.: Heidegger's Metaphysics of Material Beings. Philos. Phenomenological Res. **87**(2) (2013). doi:10.1111/phpr.12000
15. Stepanid, L.V.: Heidegger: between idealism and realism. In: Harvard Rev. Philos., Spring, pp. 20–28 (1991)
16. Harman, G.: Technology, Objects and Things in Heidegger. Cambridge J. Econ. **34**, 17–25 (2010)
17. Heidegger, M.: 'The Thing' in Poetry, Language, Thought, trans. A. Hofstadter (1971)
18. Harman, G.: Objets et architecture/objects and architecture. In: Brayer, M.-A., Migayrou, F. (eds.), Naturaliser l'Architecture/Naturalizing Architecture. Editions HYX (2013)
19. Hale, J.: Harman on Heidegger: 'Buildings as Tool-Beings', posted by body of theory, May 29, 2013, extract from a talk at The Swedenborg Society, London, May 2013. http://bodyoftheory.com/2013/05/29/harman-on-heidegger-buildings-as-tool-beings/
20. Scott, A.: Heidegger's Being and Time (2015). http://www.angelfire.com/md2/timewarp/heidegger.html (access)

Constrained Triangulation of 2D Shapes

Laxmi P. Gewali and Roshan Gyawali

Abstract Algorithms for triangulating two dimensional shapes have been used as sub-problems in many application areas that include finite element analysis, geographic information systems, and geometric compression. We consider a constrained version of triangulation problem in which the objective is to increase the proportion of even degree vertices. We present an effective approach for generating triangulated polygons with increased number of even degree vertices. The proposed approach is based on the convex decomposition of polygon followed by 'diagonal flipping' operation.

Keywords Constrained triangulation · Mesh generation · Polygon decomposition

1 Introduction

Triangulation is the problem partitioning a two dimensional domain into triangles. The input is usually in the form of polygons and/or point sites and the output is a triangulated mesh. Triangle meshes have been extensively used in finite element analysis for obtaining approximate solution for fluid flow problem [6]. The quality of the solution depends on the quality of the generated mesh. A triangulated mesh with large proportion of non-skinny triangles is highly desired. In recent years it has been found that a triangulated polygon with increased number of even-degree vertices can be applied to tackle illumination problems [15].

In this paper we consider the problem of triangulating polygons with increased number of even degree vertices. We present how two well-known techniques of computational geometry can be used to obtain a triangulated polygon with high proportion of even-degree vertices. Specifically, we combine 'convex decomposition' and 'diagonal flipping' techniques to generate a triangulated mesh with high number of even-degree vertices. In Section 2, we present preliminaries and a brief

L.P. Gewali(✉) · R. Gyawali
Department of Computer Science, University of Nevada, Las Vegas, USA
e-mail: laxmigewali@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_104

review of recent progress made in understanding degree constrained triangulation problems. In Section 3, we present two approaches for triangulating simple polygons which can lead to increased number of even degree vertices. Finally, in Section 4, we discuss extensions of the proposed algorithms and direction for further investigation.

2 Preliminaries

A simple polygon is triangulated if its interior is partitioned into triangles by adding diagonals which do not interest on their interior. The problem of triangulating simple polygon is a well investigated problem and several algorithms for its solution have been reported [2,3,4], [7], [8], [13]. A linear time algorithm for triangulating a simple polygon is found in [3] which is rather complex for implementation. For practical implementation, randomized triangulation algorithms are preferred [7]. In recent years triangulation that satisfy additional properties have been considered [15]. One such property is the vertex degree requirement as defined below.

Definition 1: A simple polygon P is said to admit *even-degree triangulation* if every vertex in a triangulated graph of P is of even degree. Fig. 1 shows an example of even degree triangulation. The notion of odd-edge triangulation can be defined similarly.

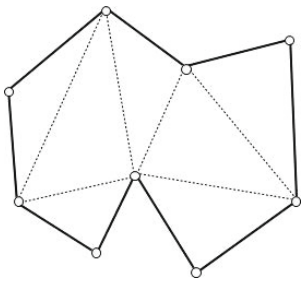


Fig. 1 Illustrating even-degree triangulation

An interesting issue in this regard is the possibility of triangulating a given polygon to make all vertices of even degree. This is stated in the following lemma.

Lemma 1: Not every polygon admits even degree triangulation.

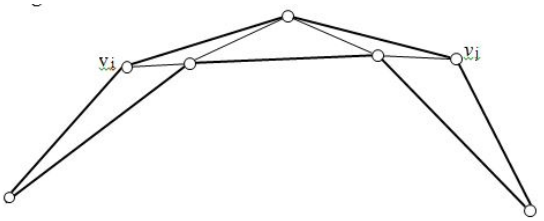


Fig. 2 Some polygons do not admits even-degree triangulation

Fig. 2 illustrates Lemma 1 for a polygon with 7 vertices. Since this polygon can be triangulated only in one way as shown, it trivially does not admit even degree triangulation. We can construct a similar spiral polygon with any given number of vertices that do not admit even degree triangulation.

So, it is interesting to develop efficient algorithms to triangulate a given polygon with large proportion of even degree vertices. This problem was first introduced in [15] and can be stated as follows.

Degree Constrained Triangulation (DCT) Problem: Given a polygon P, the DCT problem asks to triangulate P with increased number of even-degree vertices. DCT problem for a set of nodes in two dimensions can be defined similarly.

Some of the proofs of the lemmas stated in this section are omitted due to lack of space. Complete proofs are available in [5]. We can first consider the case for convex polygons. It turns out that any convex polygon with number of vertices that are multiples of 3 indeed admit even degree triangulation. This is stated in the following lemma.

Lemma 2: Any convex polygon with $n = 3k$ (k>1) vertices admits even degree triangulation.

Proof: (Omitted)

Convex polygons not having number of vertices a multiple of 3 can be triangulated so that most of the vertices except two are of even degrees. Algorithm that produces such triangulation is named alternate quadrangulation triangulation algorithm (**AQT algorithm** in short) in [5].

We now give a very brief sketch of AQT algorithm. Details and complete analysis can be found in [5]. The approach is to first partition the polygon into alternate quadrilaterals and triangles by constructing diagonals originating from a single vertex called **source vertex**. Note that due to convexity, such diagonals can always be drawn from any vertex. Let TQ denote the polygon whose interior is partitioned into such alternate triangles and quadrilaterals. An example of TQ is shown in Fig. 3, where v_0 is the source vertex. The quadrilaterals in TQ are partitioned into two triangles each by using the diagonals connecting vertices other than the source vertex. The resulting triangulation has at least $n-2$ vertices of even degree as shown in Fig. 4.

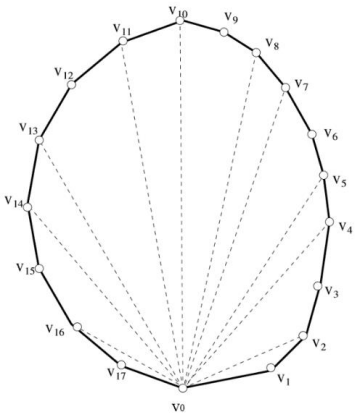


Fig. 3 Partitioning into alternate quadrilaterals and triangles

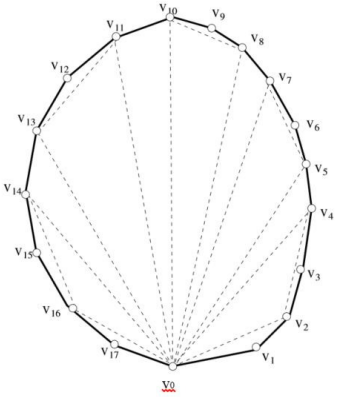


Fig. 4 Triangulation produced by AQT Algorithm.

This is stated in the following lemma.

Lemma 3: Any convex polygon with n = 3k + 1 or 3k + 2 (k>0) vertices can be triangulated to have at least n − 2 vertices of even degree.

Proof: (Omitted)

Theorem 1: DCT problem for convex polygon can be solved in *O(n)* time by using AQT Algorithm.

Proof: (Omitted)

Lemma 4: No simple polygon admits odd-degree triangulation.

Proof: (Omitted)

3 Algorithms Development

We now present the development of algorithms to solve DCT problems for simple polygons, not necessarily convex. The first algorithm we present works by applying flipping operations after the polygon is triangulated by using the well-known Delaunay triangulation [4]. The second algorithm applies flipping operations by first partitioning the input polygon into convex components. The flipping is done on each triangulated convex component separately.

Development of Scan/Flip Algorithm
The input is a simple polygon *P* which is given by listing the coordinates of its vertices in counterclockwise order. The vertices in this order are denoted as v_0, v_1, v_2, ... , v_{n-1}. The polygon *P* is first triangulated by using the Fortunes' Delaunay triangulation algorithm [4]. We use doubly connected edge list data structures (DCEL) [2] to store the resulting triangulated polygon *T(P)*. The algorithm proceeds by processing diagonals incident on a vertex. The vertices are visited by traversing them along the boundary of the polygon. It is noted that each diagonal *d* of the triangulated polygon *T(P)* corresponds to a unique quadrilateral *Q(d)* formed by combining two triangles incident on it. The algorithm processes *Q(d)* to check if its diagonal can be flipped to increase the number of even-degree vertices. The two rules for checking valid flipping conditions for quadrilateral *Q(d)* can be listed as follows.

Flipping Rules
Rule 1: *Q(d)* must be convex.
Rule 2: *Q(d)* must have more than two odd-degree vertices.

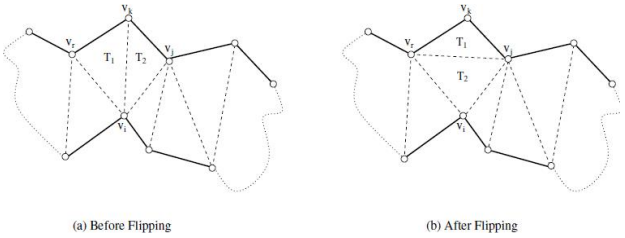


Fig. 5 Illustrating flipping operation

A flipping operation is illustrated in Fig. 5. The left part of Fig. 5 displays a portion of a triangulated polygon, where two adjacent triangles are marked as T_1 and T_2. The union of these triangles forms a convex quadrilateral and hence satisfies Rule 1. This quadrilateral has three odd degree vertices and hence satisfies Rule 2. Thus T_1 and T_2 can be flipped to increase the number of even degree vertices. The flip operation is applied we obtain the triangulation shown in the right part of Fig. 5. Our algorithm is based on applying such flipping operation on carefully selected adjacent triangles.

We can illustrate the progress of the algorithm with a running example. Fig. 6 (top) shows a triangulated simple polygon *T(P)* with twenty four vertices v_0, v_1 ..., v_{23}. The triangulation dual (Fig. 6 - bottom) is drawn by thick segments connecting black dots. Note that the dual of a triangulated polygon *T(P)* is a graph whose nodes are the triangles and edges are formed by connecting adjacent triangles. It is easily seen that the dual of a triangulated simple polygon is a tree.

In order to traverse diagonals of *T(P)* , we perform a vertex scan starting with any vertex say, v_0. For each vertex v_i, we find the diagonals incident on it. In Fig. 6 (top), the incident diagonals for vertex v_0 are (v_0,v_2), (v_0,v_3), (v_0,v_4). These diagonals are checked for possible flipping. In order to apply the flipping operation, the diagonal must satisfy the two rules stated above. Furthermore, there are two additional conditions stated next that the candidate diagonal *d* must satisfy.

Condition 1: Diagonal *d* was not processed before.

Condition 2: Diagonal *d* cannot be a newly created diagonal, obtained by applying flipping operation.

Theorem 2: Given a triangulated polygon, DCT problem for simple polygons can be solved in O(n) time.

Proof: (Omitted)

When these rules and conditions are applied to the triangulated polygon shown in Fig. 6 (top), two flipping operations are triggered. In order to emphasize the occurrences of flipping operations, we have drawn the dual of the triangulation with thicker edges. The triangulation after the flipping is shown in Fig. 6 (bottom). The two diagonals formed after the flip operation (v_{12}, v_{23}) and (v_1, v_3) are drawn dotted in Fig. 6 (bottom). Since the initial triangulation is available in DCEL data structure [2], the triangulated polygon can be scanned and checked for possible flipping by traversing the polygon along its boundary. A formal sketch of the Flip/Scan algorithm is listed as Algorithm 1.

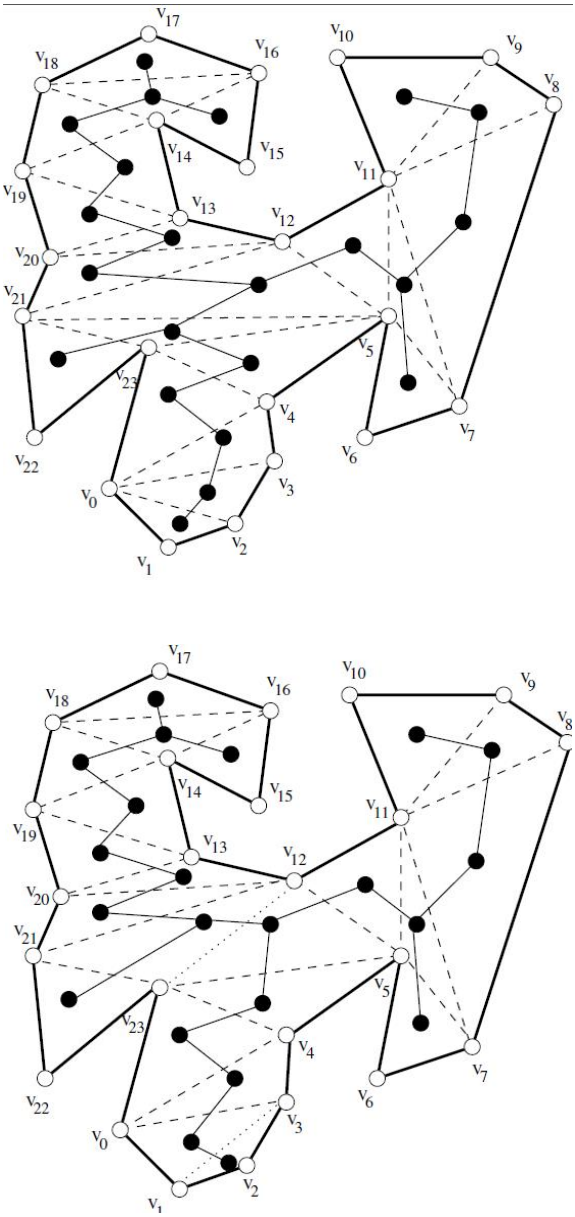


Fig. 6 Triangulated simple polygon *T(p)* and its dual

Algorithm 1: Scan-Flip Algorithm

```
INPUT:   A simple polygon P with vertices v0,v1,...,vn-1
OUTPUT:  Triangulated polygon T(P) with increased
                  number of even-degree vertices
Step 1:   Obtain a triangulated polygon T(P) of P
Step 2:   i = 0;
Step 3:   repeat
Step 4:      for each diagonal d incident on vi do
Step 5:           if (Q(d) satisfies flippability rules
                                 and Conditions 1 and 2)
Step 6:              T(P) = flip(T(P),d)
Step 7:           end if
Step 8:      end for
Step 9:      i = (i + 1) mod n
Step 10:  until (i ≠ 0)
```

Development of Partitioning-Flip Algorithm

The quality of the solution obtained by applying Scan-Flip Algorithm (Algorithm 1) may not always yield high quality solutions. We observed in Section 2 that convex polygons can be triangulated to have almost all vertices with even degrees. This gives us a hint that we can break the polygon first into convex components and triangulate them separately to increase the number of even degree vertices. Breaking a simple polygon into convex pieces is itself a very difficult problem. Some of the well-known algorithms for breaking a polygon into convex pieces are reported in [13]. For our investigation, we pick Hertel-Melhorn's algorithm [8] (HM-algorithm, for short) for convex decomposition. HM algorithm is simple to understand and implement.

An Additional merit of HM algorithm is the fact that it has a guaranteed bound for the quality of the resulting solution. In fact, HM algorithm obtains a solution which is no more than four times the optimal solution. The algorithm we propose that makes use of convex partitioning is called the partitioning-AQT algorithm. Let c_0, c_1, ..., c_k be the k convex components when the input polygon P is partitioned by applying HM algorithm. Each convex component c_i is triangulated to make almost all vertices of even degrees. Then these triangulated convex components are merged together to obtain the overall triangulation of the whole polygon. If a small number of convex pieces are generated in the convex decomposition then the quality of the solution obtained by applying partitioning-flip algorithm improves significantly. A formal sketch of Partitioning-Flip Algorithm in reported in [5].

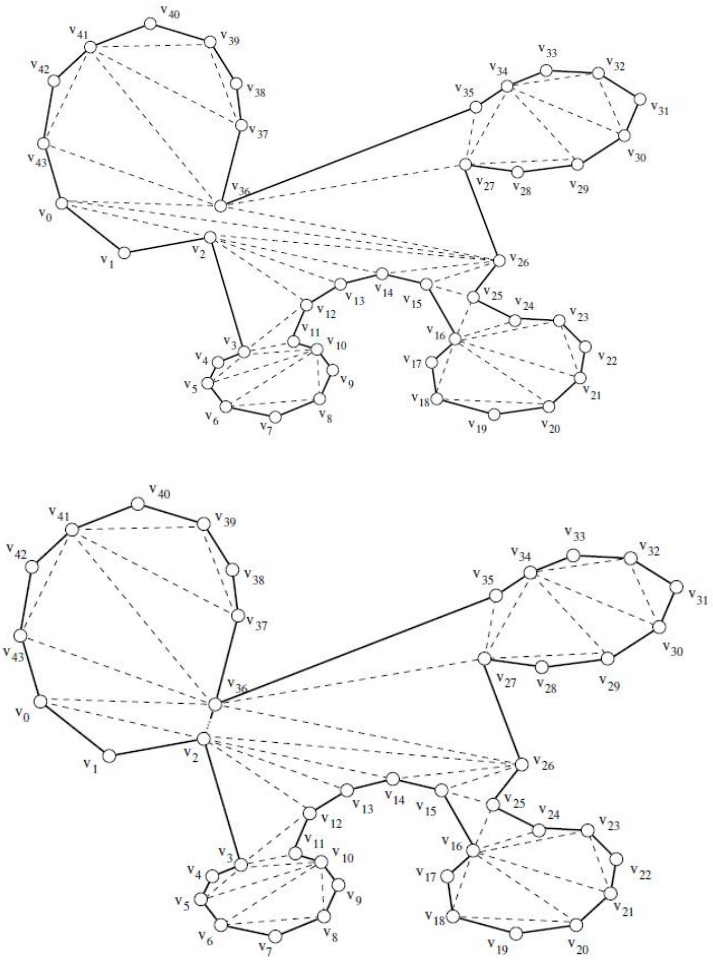


Fig. 7 Progress of Partitioning-Flip Algorithm

4 Discussion

We presented a review of algorithms for DCT problem for convex polygons. We showed that convex polygons can be triangulated to have almost all vertices of even degrees. We presented two algorithms to solve DCT problem for simple polygon. In order to get further insight on the performance of the proposed algorithm, it would be interesting to perform experimental investigation by implementing them. One approach could be to generate random simple polygons [1] of various sizes (number of vertices) and apply the proposed algorithms on them.

For some polygonal shapes any triangulation would have very few number of even degree vertices. This happens when the boundary consists of many consecutive reflex vertices. In such polygons, it can be verified easily that no matter how the polygon is triangulated it will have high proportion of odd degree vertices.

References

1. Auer, T., Held, M.: Heuristics for the generation of random polygons. In: Fiala, F., Kranakis, E., Sack, J.-R. (eds.) Proceedings of the 8th Canadian Conference on Computational Geometry, Carleton University, Ottawa, Canada, August 12–15, 1996, pp. 38–43. Carleton University Press (1996)
2. de Berg, M., Cheong, O., van Kreveld, M., Overmars, M.: Computational Geometry: Theory and Applications, 3rd edn. Springer (2008)
3. Chazelle, B.: Triangulating a simple polygon in linear time. Discrete Comput. Geom. **6**, 485–524 (1991)
4. Fortune, S.: A sweepline algorithm for voronoi diagrams. Algorithmica **2**, 153–174 (1987)
5. Gyawali, R.: MS Thesis, Department of Computer Science, University of Nevada, Las Vegas, August 2012
6. David, J.F., George, P.-L.: Mesh-Generation: Application to Finite Elements, 2nd edn. Wilet (2010)
7. Guibas, L.J., Knuth, D.E., Sharir, M.: Randomized incremental construction of delaunay and voronoi diagrams. Algorithmica **7**(4), 381–413 (1992)
8. Hertel, S., Mehlhorn, K.: Fast triangulation of simple polygons. In: Proceedings of the 1983 International FCT-Conference on Fundamentals of Computation Theory, pp. 207–218. Springer-Verlag, London (1983)
9. Hurtado, F., Noy, M., Urrutia, J.: Flipping edges in triangulations. In: Proceedings of the Twelfth Annual Symposium on Computational Geometry, SCG 1996, pp. 214–223. ACM, New York (1996)
10. Lawson, C.L.: Software for c1 surface interpolation. Mathematical Software III, pp. 161–194 (1977)
11. Meisters, G.H.: Polygons have ears. American Mathematical Monthly **82**, 648651 (1975)
12. O'Rourke, J.: Art gallery theorems and algorithms. Oxford University Press Inc., New York (1987)
13. O'Rourke, J.: Computational Geometry in C, 2nd edn. Cambridge University Press (1998)
14. Osherovich, E., Bruckstein, A.M.: All triangulations are reachable via sequences of edge-flips: an elementary proof. Comput. Aided Geom. Des. **25**(3), 157–161 (2008)
15. Urrutia, J., Pelez, C., Ramrez-Viguer, A.: Triangulations with many points of even degree. In: Proceedings of the 22nd Annual Canadian Conference on Computational Geometry, Winnipeg, Manitoba, Canada, August 9–11, 2010, pp. 103–106 (2010)

Software Project and Analysis of a Training Screen Based System for Healthcare Professionals Working in the NICU in Brazil

Daniel Rocha Gualberto, Renata Aparecida Ribeiro Custódio, Alessandro Rodrigo Pereira Dias, Gabriel Bueno da Silva, Clarissa Gonçalves Eboli and Alexandre Carlos Brandão Ramos

Abstract *Objective:* To develop a screen-based training system (SBTS) for training of healthcare professionals working in NICU in Brazil. *Materials and Methods*: Using of Intelligent Tutoring Systems (ITS) and Production Rules Based System (PRBS) for the training of these professionals to learn and recognize the various normal and abnormal situations of the incubator interface as well as the associated alarms, what can significantly contribute to the increasing of patient safety. *Results and Discussion*: The system developed by applying an ITS assists a teacher in the training task of a new user or even experienced users with the purpose of training and update. The developed tool allows modeling of new cases in the system. *Conclusion*: The systematic adopted for the development of training systems allows that new ITS are developed from the reuse of classes already developed, easing the implementation process and increasing the effectiveness for new training.

Keywords Component · Software engineering · Newborn intensive care units · Infant incubator · Clinical alarms · Screen-based training system

1 Background

The relevance of this project is the growing concern in achieving high safety and quality indices in the care of patients, which is closely linked to technology industry associated with this care, for treatment as well for diagnostics and personnel training.

D.R. Gualberto(✉) · R.A.R. Custódio · A.R.P. Dias · G.B. da Silva · C.G. Eboli · A.C.B. Ramos
Institute of Mathematics and Computer Science,
Federal University of Itajubá, Itajubá, MG, Brazil
e-mail: rocha.gualberto@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_105

The technology applied to medicine has become increasingly important in patient care. This growth is linked to changes, updates and new projects of medical equipment [1]. These technologies increase the possibilities for diagnosis, treatment and improve the quality of life of the patient. However, such advances can greatly enhance the complexity of the available systems for healthcare and compromise safety, especially when we talk about alarms that are extremely common throughout this technology [2]. Every year the Emergency Care Research Institute (ECRI), a non-governmental organization USA that conducts healthcare researches, has its main approach healthcare technologies and risk management, quality and environment in healthcare field. In an annual report, this institute publishes the top 10 hazards to health technology. The reports made in 2012, 2013, 2014 and 2015 point the hazards of Alarms as the most important problem [3].

Recent evidence the US Food and Drug Administration (FDA) and The Joint Commission has shown a link between alarm fatigue and death of patients. High false alarm rates have the potential to lead to fatigue of it, leading nurses to delay their responses to them, ignore them or turn them off entirely [4]. To reduce problems with alarms, the Health Technology Foundation recommends initial and continuing training on alarm based medical devices that people are expected to operate. To reduce alarms, the Healthcare Technology Foundation recommends initial and ongoing training on alarm-based medical devices que staff are expected to operate [5].

ECRI's reported experience, problems often stem from alarms being improperly configured or inadvertently defeated by staff. These are impacted both by the human factors design of the device's alarm system, as well as the nursing staff's level of proficiency with configuring and managing alarms. Thus, effective initial and ongoing training is still of vital importance. While on the surface, many alarm systems seem straightforward, the intricacies are often not well understood by staff [6]. Thus, the training becomes more important as more complex the health environment is and where on-site trainings are not allowed because they are related with patient safety, as the Newborn Intensive Care Unit (NICU) for example, where the number of alarms per hour per patient is higher than the adult ICU [7]. A research realized shows that most NICU alarms had pulse oximetry (68%) the majority of our alarms in NICU [8]. This equipment is usually associated with neonatal incubator, which in this life-support context becomes a protagonist equipment by meeting the basic health needs of the newborn. It is also important to know how to associate other equipment (physiological monitors, mechanical ventilators, infusion pumps) or even equipment that are used during certain procedures and perform procedures in the newborn interfering as little as possible in ideal conditions for the baby. This has been a challenge for the health team. To assist the professional, the incubator also has alarms that indicate if the equipment is out of the chosen parameters or if there is a system failure. Therefore, knowing well the neonatal incubator and learn to fix possible negative situations that may occur in this technological ensemble becomes, in many cases, crucial to proper development and survival of the newborn. Thus, in order to contribute to the reduction of hazards and damage to the health of newborn this work aims to develop a screen-based training system (SBTS) for training of healthcare professionals working in the NICU in Brazil. By using Intelligent Tutoring Systems (ITS) and

Production Rules Based System (PRBS) for the training of these professionals, it is possible to learn and recognize the various normal and abnormal situations of the interface of the incubator and the associated alarms, which can significantly contribute to increasing of patient safety.

2 Materials And Methods

This research is exploratory and qualitative. The SBTS was developed with the structure of Media and Interactivity Lab and the Usability and Human Factors Lab.

Thus, this paper provides an approach based on the formalization of the Unified Modeling Language - UML and software engineering, for the analysis and design of a SBTS for healthcare professionals working in the NICU. It shows the circumstances found in the training process of these professionals and allows learning and recognition of various normal and abnormal occasions as well as the application of problem-solving strategies. It presents the methodological aspects of a case study that aims to provide the climate proposed by the project, applying state of the art in terms of existing technologies in Software Engineering.

The training was conducted by ITS with the test results in SBRP. The methodology chosen for the development of the project was the systems prototyping methodology, due to the ease of prototyping using graphical and object-oriented interface in order to train the NICU professional to normal and adverse event occurrences related to neonatal incubator and its alarms. Typically, a device such as neonatal incubator has a rich alarm system that, through a beep, gives science to the healthcare provider of which parameters are beyond the ideal or even if system failures are occurring. There are incubator models, for example, that has 30 alarms.

3 Results And Discussion

A. Intelligent Tutoring Systems – ITS

To be considered an intelligent system, the system must have three characteristics that shows intelligence:

- The subject should be well known enough in order for the system to make inferences about the domain or solve problems that are in their scope of action;
- Must be able to evaluate the student;
- It must provide teaching strategies that minimize the difference between the apprentice and the expert.[9]

An ITS has the domain knowledge narrowly and clearly articulated, using the student's knowledge to adapt education, the sequence of teaching is not predetermined, perform diagnostic procedures best suited to the student and allow the tutor-student communication [10]. It has in its traditional architecture four functional components: the subject that teaches (domain module) the ones that characterize the student (apprentice module) to provide individualized instruction, the strategies for teaching (tutor or teaching module) and the communication module, which should be well planned and easy manipulated to favor the tutor-student communication (Fig. 1).

This system dynamically combines component information to make appropriate decisions in specific situations in a tutoring section [11].

Student Module

The student module represents the knowledge and skills of the student at the time of interaction with the ITS. The fundamental characteristics of research in ITS is the meticulous attention to the student, since this information is used for decision making, consisting of static and dynamic data that is critical to the tutor prove hypotheses about the student [12].

There are different methodologies that are employed for the construction of the student module, for example: comparing the conduct of a student with a specialist, inclusion of preferences of the student, indicating of particular goals, and inclusion of pattern recognition applied to history of the answers provided by the student, the last one is adopted in this work.

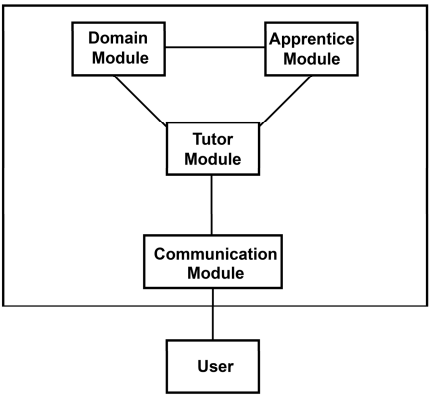


Fig. 1 Model of an Intelligent Tutor System – STI

The student module has three tasks:

- Collect data about and of the apprentice, which can be explicit (ordering any response) or implicit (some kind of interaction);
- Use the data to represent the student's knowledge and learning processes;
- Represent the data making some kind of diagnosis, both in the student's knowledge state and in terms of optimal selection of teaching strategies to present the domain information to the student.[13]

Domain Module

The domain module is the specialist element of the tutor and contains the information about the domain that he wants to teach. It is an organized database into declarative and procedural knowledge in a specific domain [14]. The purpose of this module would reproduce these knowledge structures in the apprentice's mind.

Module Tutor or Educational

The tutor module has the strategies and tactics of teaching. Strategies compose the knowledge of how to generate, from the diagnostic information, monitoring and analysis, a sequence of teaching tactics to present a topic to a particular student. A teaching strategy should define:

- Information on when to stop the course of reasoning or student learning;
- Information on which topics present the order of presentation;
- Information on how to present content.

A widely used method is the so-called Socratic method, where the tutor teaches through questions and dialogue, leaving the students draw their own conclusions. Another theoretical model used is the training model (coaching) that uses entertainment activities to show related concepts where learning is an indirect consequence. A third model is the one of hypertexts, where the student browses a hypertext structure and explores the content according to their interests working in a more participatory and dynamic way.

Communication Model

During the interaction, the tutor system presents the teaching material and monitor student progress by receiving their answers. Therefore, the communication model is essential for the success of the system. It should offer a number of features in the presentation of the contents to prevent the student to bore, such as: hypertext and hypermedia systems that have been used in this interface model. It should allow ease for swapping the dialogue initiative where students can intervene at the tutor's speech and vice versa (the response time should be within acceptable limits and monitoring should be conducted so as not to charge the student with excessive questionnaires).

In face of the exposed models, the high degree of interconnection between the four components of an ITS architecture makes that techniques used in a model can be applied in others. The use of a cognitive model to check apprentice errors (student module) may present knowledge to him (domain module) and also communicate real knowledge used for the solution of relevant problems, rather than abstract situations (teaching module). The possibilities of communicating knowledge (communication board) also provide opportunities to implement educational strategies, providing basis for the apprentice during the early stages that are disappearing as he is learning.

B. Training of Health Professionals Using STI

This system consists of an application of an ITS for training of healthcare professionals working in the NICU. Its purpose is to assist a teacher in the training task of a new user or even experienced users for the purpose of training and update in relation to the neonatal incubator use mode in common and adverse conditions and its alarms. The tool developed seeks to facilitate the modeling of new cases, as well as the introduction of new episodes in the system. The prototype includes the following features:

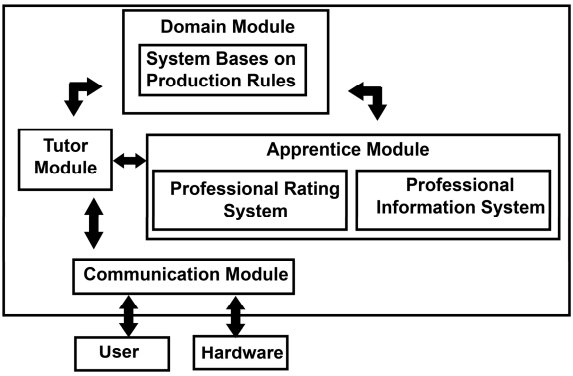


Fig. 2 Diagram of ITS blocks

- User friendly interface for modeling new cases: the developer provides an alternative to the case definition models, enabling the creation of models that describe a script for an episode.
- User-friendly interface for entering cases: use the data defined in the schedules to offer different combinations of attribute values in the insertion of new cases.
- Similarity measures: it calculates similarity that has a given case for all other cases that follow the same model;
- Model views and Cases: looks the cases that were modeled in the system and also the cases already entered on it.

The Figure 2 shows a block diagram of the ITS for NICU, where it can observe its relationship with the many modules:

C. System Based on Production Rules (SBPR)

Using the SBPR is possible to set up a knowledge base (specifications and functions of the neonatal incubator), based on interviews with experts in the problem domain. The SBPR is integrated with a database of historical data values corresponding to the interaction with users and witch serve it to the representation of the instabilities in the process. Therefore it is possible that the inferences caused by various interactions are identified and explained with absolute precision, allowing the use of artificial intelligence techniques to infer the occurrences arising from interactions and their operating procedures.

The operating cycle of the SBPR are the facts (alarms and events) that are collected from the process through the communication program developed in Java that soon after, are analyzed and executed with the premises of the rules in OPS5, enabling them and then the conclusions of the rules that are presented to the user by the system.

D. System Procedures and Process Simulation

This system's main purpose is to be a graphical and user-friendly interface with the healthcare professional. Its most important activities are:

- Provide to health professionals all operational procedures of neonatal incubator (specifications, functions, operations, alarms, etc.) from a hypermedia and interactive interface;
- Present real situations of instability using the information contained in the database provided by the Data Capturing System of the process;
- Present diagnosis of adverse situations, as well as possible actions to be taken by health professionals in order to make the appropriate action in order to correct it;
- Allow to the healthcare professional interaction with the system in order to facilitate the training of standardized procedures to correct adverse situations;
- Evaluate the healthcare professional learning in order to discover their weaknesses and facilitate their training at these points.

E. Systematic Adopted for the Acquisition System

To obtain the system and solutions were identified the problems and solutions through the general context following the heuristics to specify requirements, model, implement and test their applications. The prototype development process of a Software System (as product) must follow the methodology of the Object Modeling Techniques - OMT, adapted to the Standard Unified Modeling Language - UML and the Rational Unified Methodology Process - RUP. The latter suggests the development of a project in four phases:

- Inception - Define project scope: At this stage the business case of the system is established, delimiting the project scope and identified all external entities (actors) with which the system will interact identifying all use cases and all relevant documentation to the project as: overview of the main requirements, key features, main constraints, glossary, initial risk estimation and project plan showing the phases and iterations. At the end of the inception phase is reached the first major milestone of progress: Lifecycle Objective.
- Development - Plan, specify the features and design the architecture: In this preparation phase, the architectural foundations are established for the application and developed a project plan, eliminating high-risk elements for the project and obtaining a superficial understanding of the entire system. As a result you get the software architecture description, an executable prototype, a list of revised risks, a development plan for the entire project and a preliminary user manual, reaching the second major milestone of progress: Lifecycle Architecture.
- Construction - Build the product: At this stage, the project aims to develop the remaining components and application features and integrate them into the final product. Test all the features and optimize cost and quality, obtaining with it an integrated product in a suitable platform, user manuals and description of the current version, reaching the third major milestone of progress witch is the Initial Operational Capability.
- Transition - Transfer the product into the user community: Its purpose is to transfer the system to users; develop new versions of the system (in the event of errors or features that were delayed), make sure that an acceptable level of quality has been reached and the user documentation is available, getting to reach the user self-support; achieve development baselines according to the system view, achieving the base product lines quickly and in accordance with the estimated cost, concluding the fourth major milestone of progress: product release.

F. Scenario

According to the requirements specification provided by the user, initially identified the scenario: It focuses on the need for a computerized solution for employee training. Usually when a new employee is hired, he needs training in order to obtain the required familiarity with the control systems of the NICU model. This training, currently, as a rule, is held from conventional lessons in class, where an instructor, usually an experienced employee, with the support of a poster with the photograph (or drawing) of the NICU model instruments presenting their routine operating and emergency procedures.

These conventional classes are costly to the hospital, because during this phase, an experienced employee (the instructor) is prevented from acting in the NICU. The creation of a training system for healthcare staff offering a rich and interactive environment where information covering all operating and emergency procedures, could become a viable solution to the difficulties mentioned above, since there are enough technologies for the completion of such a project.

An ITS reduces the costs involved at this stage, as it provides the employee training from tutorials supported in the use of hypermedia developing tools, database management systems and a monitoring tool of the evolution of the employee's learning process.

G. Requirements Analysis

It is understood by software requirements, sentences that express customer requirements, which affect the quality of software. Because of this, we classify as functional requirements (functions that the system should provide, how it should react to particular inputs and how to behave) and non-functional (are restrictions on the service). Below are the functional requirements for the its shall, these user functional requirements define specific resources to be provided by the system:

- Allow the administrator to select the registration option, where he can include and exclude information from apprentices; The inclusion and exclusion of information in the system's database; Offer to the correlation between the information contained in the database and the evaluation of information from the apprentice emitted;
- Allow the apprentice to view the tutorials and tests on the various procedures of the NICU;
- Allow the apprentice and administrator to view the evaluation of the many subsystems of the incubator.

H. Use Case

The purpose of the Use Case is to describe the system functional requirements; it provides a consistent and clear description of the responsibilities that must be met by the system and offer possible real-world situations to test the system. The overview of SBTS functions are shown in Figure 3 and in Table 1 is described the use case.

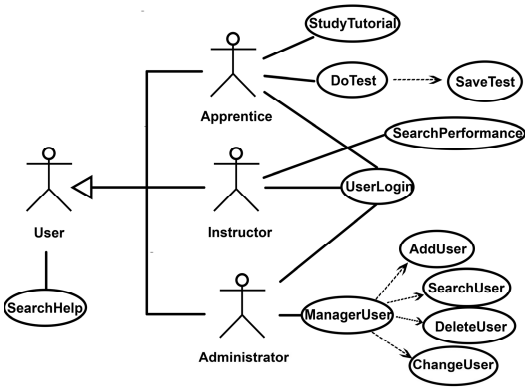


Fig. 3 General vision of the SBTS functions

Table 1 Use Cases

Use Cases: "UserLogin"; "AddUser"; "ChangeUser"; "SearchUser"; "DeleteUser"; "SearchTutorial"; "SearchTest"; "SearchTutorial"; "SearchHelp"; "DoTest"; "SearchPerformance"
Primary Actor: User
Description of the actors: **Administrator:** Responsible for carrying out managerial system operations. Arrange for the correct operation of the system and its connection to the database. **Apprentice:** Performs its training with the help of tutorials, as well as making assessments to value its learning. **Instructor:** Its most important function is to track the performance of the operator in the use of tutorials.
Description: SBTS presents tutoring systems, simulations, registration and evaluation. Through these items the student can begin their studies where, through an interactive multimedia environment, can get all the knowledge related to NICU model as well as be evaluated for the information submitted to it.
Preconditions: The user must be properly registered in the system to be able to make use of it.
Main stream: **1.** When triggered, the system presents to the user the system's home screen; **2.** The user enters valid login and password; **3.** The system presents with all the procedures that composes its permission; **4.** User Apprentice with permission (A1); **5.** User Instructor permission (A2); **6.** User with Administrator permission (A3); **7.** The system terminates the use case.

Table 1 (*Continued*)

Alternate flow: **A1: [Apprentice permission with User]** **1.** It is presented to the apprentice the theory of all procedures that make up a model of NICU and small tests that will consolidate what has been presented. **2.** The learner selects the tutorial that wants to study **3.** The system displays tutorial information **4.** The apprentice goes to the evaluation stage, where it is evaluated for his degree of learning; **5.** The evaluation system provides the apprentice the learning level and shows results of evaluation. **A2: [User Instructor permission]** **1.** The instructor seeks the apprentice from whom wants to search performance; **2.** The system shows the performance of the selected apprentice. **A3: [User with Administrator permission]** **1.** It is presented with the administrator the essential functions to the user control (Add, Change, Search and Delete a user);
Post-Conditions Apprentice goes through all the important steps to learn and consolidate the knowledge of a NICU model and, when obtain a result of satisfactory assessment, can begin to operate in the NICU sector
Functional Requirements Satisfied: The system offers, through an interactive environment, all operating and emergency procedures. The system evaluates the apprentice's learning.

Sequence Diagram

In order to represent the Apprentice behavior, our aim actor, with the system, the sequence diagram is used. Inside a context, objects are shown participating in interactions according to their lifelines and the messages they exchange, in order to establish the objects that interact and their relationships within a context (use case). Figure 4 shows the Sequence Diagram for apprentice's interaction with the tutorials.

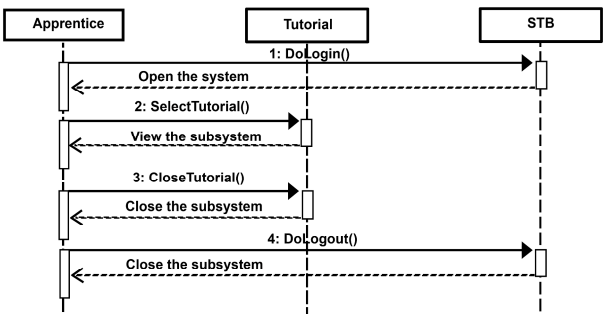


Fig. 4 Sequence diagram for apprentice's interaction

The Figure 5 shows the sequence diagram for apprentice's interaction with the evaluations.

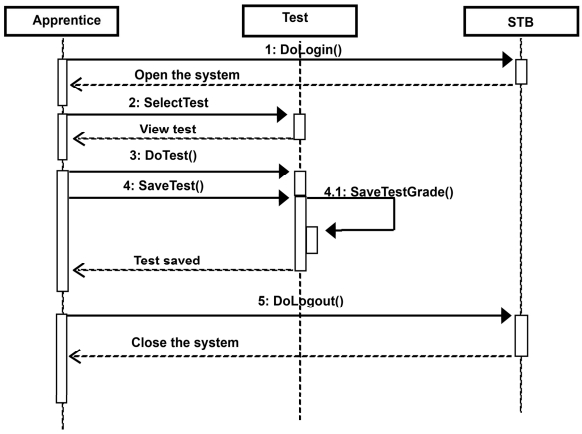


Fig. 5 Sequence diagram for operator interaction with the evaluations

Class Diagram

The class diagram is intended to describe the various types of objects in the system and the relationship between them. This offers the prospect of implementation, addressing the implementation details, such as navigability, type of attributes, etc. For the SBTS is represented in Figure 6.

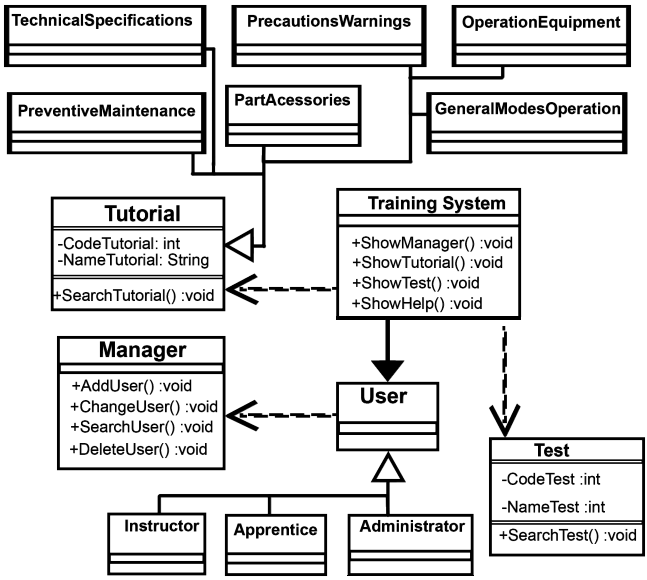


Fig. 6 Class diagrams for SBTS

4 Conclusion

This article aims to present the project of an SBTS for healthcare professionals working in the NICU, employing technologies that represent the state of the art in the field of Software Engineering through UML use for system modeling. The systematic adopted for the development of training systems allows that new ITS are developed from the reuse of classes already developed, easing the implementation process and increasing their productivity. It is hoped that this SBTS collaborate in training and updating of healthcare professionals working in the NICU and reducing dangers with alarms.

Acknowledgments The authors would like to thank the funding institutions: CAPES, CNPq, FAPEMIG and Ministério da Saúde.

References

1. Liljegren, E.: Usability in a medical technology context assessment of methods for usability evaluation of medical equipment. Int. J. Ind. Ergon. **6**, 345–352 (2006)
2. Reason, J.: Human error: models and management. BMJ **320**, 768–770 (2000)
3. ECRI Institute. https://www.ecri.org/press/Pages/ECRI-Institute-Announces-Top-10-Health-Technology-Hazards-for-2015.aspx (accessed September, 2015)
4. Bonafide, C.P., Zander, M., Graham, C.S., et al.: Video Methods for Evaluating Physiologic Monitor Alarms And Alarm Responses **48**, 220–230 (2014)
5. ACCE Health Technology Foundation. 2006 Impact of Clinical Alarms on Patient Safety. http://thehtf.org/documents/White%20Paper.pdf (accessed March 30, 2015)
6. Borowski, M., Gorges, M., Fried, G., et al.: Medical device alarms. Biomed. Tech. **56**, 73–83 (2011)
7. Pul, C.V., Dijkman, W., Mortel, H., et al.: Alarm management in an ICU environment (2014) (retrieved March 10)
8. Brewer, M., Diaz-Arrastia, T., Teklits, S.: Management of Alarm Fatigue in the Neonatal Intensive Care Unit (NICU) (accessed August 27, 2015)
9. Jonassen, D.H., Wang, S.: The physics tutor: Integrating hypertext and expert systems. Journal of Educational Technology Systems **22**, 19–28 (1993)
10. Urretavizcaya, L.M.: Sistemas inteligentes em el âmbito de la educacion. Revista Iberoamericana de Inteligência Artificial **12**, 5–12 (2001)
11. Akhras, F., Self, J.: Beyond intelligent tutoring systems: situations, interactions, processes and affordances, pp. 1–30. Kluwer Academic Publishers, Netherlands (2002)
12. Self, J.: Bypassing the intractable problem of student modeling. In: Frasson, C., Gauthier, G. (eds.) Intelligent Tutoring Systems: at the crossroads of artificial Intelligence and education, pp. 107–123. Ablex, Norwood (1990)
13. Wenger, E.: Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communications of Knowledge. Morgan Kaufmann Publishers, Los Altos (1987)
14. Mctaggart, J.: Intelligent tutoring system and education for the future. CI 512X Literature Review (2001)

Part XII
Short Papers

A Self-configuration Web-API for the Internet of Things

Eric Bernardes C. Barros, Admilson de Ribamar L. Ribeiro, Edward David Moreno and Luiz Eduardo C. Neri

Abstract Due to the fast growth of the Internet of Things (IoT), many concepts of distributed systems need to be adapted to this new paradigm. Due to the large size of these networks, it is important that such systems present autonomic characteristics. In this context, this paper performs the implementation of an autonomic system emphasizing the self-configuration of devices, aided by cloud computing through a Web-API that was developed for this purpose. The initial results were obtained from tests performed on the Z1 device with the help of the ContikiOS, running inside the Cooja simulator. The results revealed that there is an overflow of the code developed for the tested devices, but after deep analysis, and more development, we can get a good level of energy consumption and it is possible to reach the self-configuration of devices on the IoT environment.

Keywords Internet of Things (IoT) · Web of Things (WoT) · Self-configuration · Motes · Web-APIs · Architecture · Framework · Californium · Proxy · CoAP · Cooja · Simulation · Rest · Cloud computing

1 Introduction

The Internet is the greatest way for world's communication. It is possible to search information and communicate with all people who are connected. These features meant that several people wanted to get connected all the time not only through computers, but also using their smartphones, making the Internet into something mobile.

Currently, the Internet communication forms are expanding and making possible searching information in real time within environments that were not possible before, such as streets, houses, refrigerators and everywhere imaginable. These new possibilities are creating a new paradigm, called the Internet of Things (IoT).

E.B.C. Barros(✉) · A. de R.L. Ribeiro · E.D. Moreno · L.E.C. Neri
Computer Department, Universidade Federal de Sergipe, UFS, Aracaju, Brazil
e-mail: {eric,admilson}@ufs.br, edwdavid@gmail.com, edu_neri@hotmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_106

However, as well as on the Internet it is necessary the use of robust and easy tools, the IoT also needs applications that make it possible not only communication between people and objects, but also the interaction among them.

Faced with this reality, there is the emergence of the IoT Web-APIs, which make it possible the interaction between people and objects to facilitate the treatment of the data obtained. Making not only the IoT possible, but also possible for all people and not just for experienced programmers.

However, only interaction and communication with the devices are not enough for the IoT, because the environment is always changing and this requires that the devices and Web-APIs themselves should be configured and reconfigured to meet the needs of local in where they are.

In this context, we present the development of a Web-API which provides a solution that makes possible the self-configuration of devices within the IoT through a self-configuration policy. So our Web-API uses cloud computing that works with HTTP (Hypertext Transfer Protocol) and communicates with devices running on CoAP (Constrained Application Protocol) in the ContikiOS operating system, allowing to modify the devices' program, readjusting the Web-API to interact with this new programming.

The remainder of this paper is organized as follows. In Section 2, we describe the related works. In Section 3, we present the architecture used in the Web-API. In Section 4, we present the implementation of Web-API and explain how they were made the Server side and Client side. In Section 5, we do the evaluation of the proposed Web-API. Finally, we conclude the paper in Section 6.

2 Related Work

Due to the increasing importance of the IoT and facilities it can offer to the user, the communication between people and devices involved in this paradigm must become easier. After all, not all users are programmers. Faced with this reality several studies and tools began to be developed for this purpose.

Thus, as can be seen in [1] the Web-APIs studied as Cosm, ThingSpeak, Nimbits, SensorCloud, EvryThng, iDigi, GroverStream and Open.sen are APIs that use REST (Representional State Transfer) and which are focused on the development of friendly user interfaces focused on the receipt and processing of data, as well as receiving many data simultaneously and use multiple motes at the same time.

The **COSM tool** (formerly called Pachube) was developed to be a platform as a service (PaaS) for the IoT. With it, you can manage multiple devices through the RESTful resources, thus it is possible to deal with all the components of the API (Feeds, triggers, datastreams and datapoint) using commands via HTTP requests, such as PUT, GET, DELETE and POST [1]. Other tool is the **ThingSpeak** which is an open-source API for IoT that stores and retrieves data from devices using the HTTP over the Internet or simply over a LAN (Local Area Network) [1].

Nimbits is a collection of software components designed to record data of time series, such as, for example, the changes in temperature read by a given sensor [1]. This API is event driven (triggers) in order to record data. In this way, it is possible to perform calculations or trigger alerts along with your receipt [1].

The **SensorCloud** is a tool storage for sensors "things". SensorCloud provides a REST API to allow the upload of data to the server. The API implementation is based on patterns of HTTP commands. So, it is easily adapted to any platform [1].

Evrythng is a platform for powering applications or services directed by dynamic information about physical objects. Its goal is that all things must be connected, thus sets a world where all 'Thng' have a digital presence of assets on the Internet, even in social networks if desired, allowing the rapid development of Web applications using real-time information flowing from, any object in the world [1]. **IDigi** is a platform in the cloud for managing network devices. It offers management gateways and endpoints on the network. It presents security policies of leaders in the industry, and great scalability for the exponential growth of devices on the network [1].

On the other side, **GroveStreams** is one of the most powerful platforms in clouds capable of providing real-time decision making for millions of users and devices. Among several of its qualities is the code generation per device. With this API it is possible that when you choose your device and the function that it will run a code that can be used to synchronize the device with the API is generated, thus there is only a need to copy this code and paste in device's compiler, after sending to mote for which the code was generated [1].

Finally, **Open.sen** is a tool which it allows a rich visualization of results, so that by SenseBoard, you can see the incoming data in real time. The SenseBoard is powered by applications that are developed and installed within the API itself. These applications are independent but can be easily integrated with feeder (feeds) devices [6]. These feeders communicate with the API through channels that are connected to devices. Even so, it is possible to capture information from other applications, it is not necessarily a direct contact with the device [1].

In the aspect of security these Web-APIs use both HTTPS (Hypertext Transfer Protocol Secure) as an API-keys that serve to allow only authorized persons have access to data. Another important point is interoperability between the APIs, so that everything on the web can be used regardless of platform. However, these APIs have not strengthened the automation device configuration, which is the more complex side of working.

3 Comparison of Web-APIs

In table 1 and 2, we present the main characteristics of the Web-APIs and we note that there are one gap that is still not handled by any Web-API. This characteristic is the self-configuration [7]. Although, many of the studied Web-APIs provide examples of codes for configuring devices, none of them provides a side focused for the weak link in the IoT, the motes.

However, studies by Dunkels [2] addressed the transmission of data and reprogramming of devices using Deluge protocol for data dispersion and execution binary files using "elf" files respectively, using the embedded operating system ContikiOS. Nevertheless, this concept of reprogramming uses the code's distribution made between devices and not possess any connection with the Web.

In this way, our Web-API proposes combining the concepts found in Web-APIs exist in [2], the reprogramming devices that use the operating system ContikiOS found in [1] and the concepts that are used for a self-configuring environment, for example, creating a self-configuration policy. This policy is basically a CSV (Comma-Separated Values) that tells the devices that need to be reconfigured, the binary files that will be performed by the devices and the time (in minutes) that each program will be running on the device. Therefore, users need only enter the binary files on the Server, inform motes (with API-KEY) being controlled by it and its self-configuration policy. With this, the Web-API reads the self-configuration file and automatically keep changing the each device settings according to its policy, moreover, based on the binary file sent, the Web-API can get in time-to-time data that was captured by the motes and available so that the user can view on the web.

Table 1 Characteristics of Web-APIs [1].

	Open-Source	REST	Markup language data	Centralized Architecture	Security	Self-Configuration	Providers' Code for the device configuration	Cloud Storage
ThingSpeak	x	x	x	x	x		x	x
Nimbits	x	x	x	x	x		x	x
Open.sen		x	x	x	x		x	x
Cosm		x	x	x	x		x	x
SensorCloud		x		x	x		x	x
Evrythng		x	x	x	x		x	x
iDigi		x	x	x	x		x	x
Grovestreams		x	x	x	x		x	x

4 The Self–configuration Web-API Architecture

The idea of this architecture came from the perception that existing Web-APIs for the IoT are trying to facilitate the interaction between users and the devices that constitute the IoT. These features are clearly perceived at the application level, where Web-APIs begin to provide an interactive interface of the simplest devices and easy to use. However, the motes' configuration remains complex and the configuration and installation of devices that integrates large and complex systems is a challenge that takes time and is prone to error even for great experts [3].

The proposed Web-API was divided into two parts: Client and Server as is shown in the Fig. 1. The CLIENT's side Web-API is native into motes and was developed using techniques for low power consumption and memory. Thus, the Client provides resources so that the Server can communicate with it and the self-configuration can be performed.

In this way, upon receiving the self-configuration policy the Server looks for devices that are connected to it and configures them accordingly as required by policy or by interaction of the user. Before configuring the device, the Server also adapts with the code that will be sent to the mote, also the Server prepares itself to receive the data that can be sent by the mote. More details on this architecture can be seen in [1].

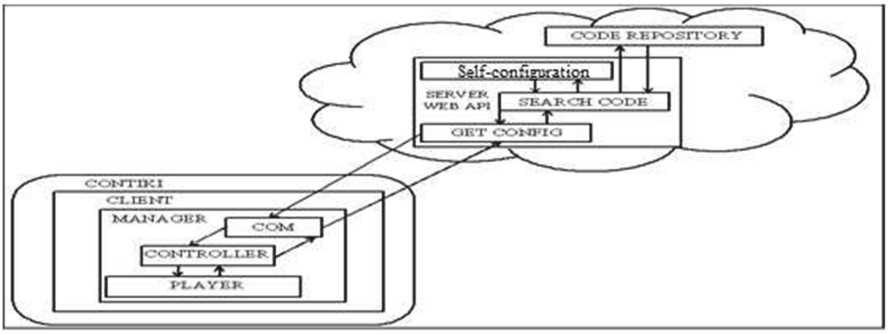


Fig. 1 Overview Web-API Architecture [1].

5 The Self-configuration web-API Implementation

The implementation of our architecture was performed using in Server side: the Apache Web server, the programming language PHP and the MySQL database. On the Client, we used the ContikiOS operating system and the C programming language, and also some applications the Contiki were used, such as Erbium with CoAP-13, the cfs-coffe and elfload.

5.1 Server Side

Thus, using the self-configuration policy file and the binary files that are stored in the WEB-API database, users can register the motes that will be managed and easily change their settings. An example of the CSV file used in Web-API can be seen in the Fig. 2.

```
mote1,blink.ce,60
mote2,humiditysensor.ce,30
mote1,temperaturesensor.ce,120
```

Fig. 2 Self-configuration policy.

It is possible to observe that there is a "name" from the mote that is being managed by the API, and the files that it will send to were being executed by the device for as long as it was determined by the user.

If there is more than one configuration to the same mote, it will enter in the queue setup to that mote, and when the time of the first configuration finish then the next item in the queue it will be sent and executed. For example, the mote 1 blink.ce program runs for 60 minutes and then temperaturesensor.ce program it is executed in 120 minutes. After this, the mote runs indefinitely lightsensor.ce program.

Due to the storage capacity of some motes is possible that some binary files are already in the mote, so before sending any information the API checks the file list that is stored in the mote, if the file to be sent is already stored in the mote the Web-API only sends the execution request of this file, and if there isn't enough space available on the device, the WEB-API orders the mote to delete the first program from ROM memory.

5.2 Client Side

To perform these procedures is necessary that the mote is running Web-API Client, which is the junction of erbium implementing the CoAP-13 designed by Kovatsch [4], the file system Contiki CFS-COFFEE and the macro elfloader that is responsible for executing the elf binary files. An overview of the Web-API Client can be seen in the Fig. 3. Thus, the devices provide four CoAP resources to be managed: file_receiver, file_exec, file_list and file_delete.

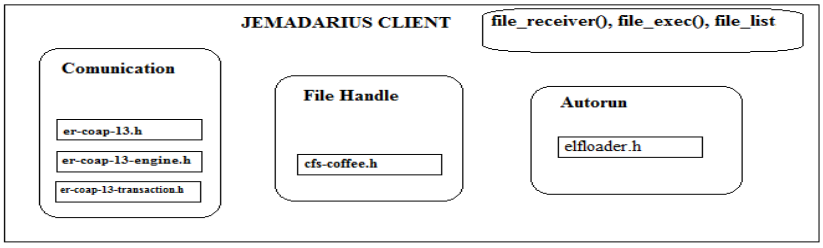


Fig. 3 General Structure of Client.

The file_receiver is responsible for receiving the resource files to be executed by the Client. It receives the data through the chunk variable that is manipulated by methods POSTs sent by the Server.

The file_exec is the resource responsible for the execution of the file that was sent or is already in the device memory, the file name is passed as parameter.

The file_list is the resource that allows the listing of all the files that are stored on the device.

The file_delete is the resource that allows to erase some binary file if necessary, the file name is passed as parameter.

5.3 Proxy

As seen in the last section, the implementation of the Server was developed using REST based on the HTTP Web platform as also seen that the Client Web-API was made using the resources created with CoAP-13, then to perform communication between these Web-API's two components it is necessary to use a translator between HTTP and CoAP and vice versa.

The research carried out showed that the Proxy developed with Californium framework, could perform this translation without large losses and with only minor adjustments to the proposed architecture as will be seen in section V.

Californium is a framework implementation in Java that performs CoAP. With it is possible to abstract the user of CoAP details and provide a framework that interacts with the CoAP endpoints to provide specific services. The Californium is modular and extensible, with this, it is possible to provide a level of abstraction that allows free use of your library. Based on this, users do not need to deal with internal issues of the CoAP as relay messages, block-wise transfer of control and manipulation of observers [4] [5].

Based on this, the Californium [8] was used to develop a proxy that can enable communication between the motes and Web-API. To develop this proxy was extended framework of the EndPoint class and with it created a client and a server to make it possible to send and receive data CoAP and other client and server to send the data via HTTP.

6 Evaluation

In the tests made in this paper, the Client program ran on the Zolertia Z1 device, although the run-time execution is limited to the symbol table.

Even with these limitations, it was possible to play automatically some programs using the self-configuration of Web-API architecture, such as LED manipulation of devices or writing data to a file sent by the program (useful for sending data back to the Server). Therefore, tests were done in order to analyze the use of CPU and memory devices, and power consumption of the whole process of self-configuration by powertrace.

6.1 Experimental Setup

Until this moment we saw the techniques used to build the Web-API architecture and the way they were implemented. Now we will validate the proposed structure. In this section, we describe the results and the costs obtained with the execution of the implementation.

The tests conducted in this experiment were performed using the Cooja simulator, which has simulated five Z1 motes that each one them has 8KB RAM and 92KB of ROM. One of these devices was used as border router, which is required for IP distribution and it allows communication with the external network. In the other four devices were placed the proposed Client.

6.2 Memory Footprint

The size of the implementation made can be found in Table III, where it is possible to observe the memory RAM and ROM sizes that were used in the simulated device in the Cooja simulator.

Table 2 Memory footprint of Z1

	RAM (KB)	ROM (KB)
Client Code	**7,4**	**61**
Full symbol table	**-**	**32**
Symbol table used	**-**	**1.4**

During client implementation some obstacles were found as the code size to be inserted into the device and the symbol table that is necessary for the execution of the code sent.

To resolve these issues it was necessary to change the device. Initially, it was planned to use the Tmotesky mote (10KB of RAM and 48KB of ROM), but we had to replace it with the Z1 mote (8KB of RAM and 92KB of ROM).

Still, it was not possible to perform the whole implementation due to the symbol table. For this reason, there was made a decrease in the symbol table, where only the symbols required to code execution were kept.

6.3 Energy Consumption

Now, we present the graphs representing the energy consumption during the execution time of receipt and storage activities.

In Fig. 4 can be seen the mote graphic on stationary processes. It can be observed that the RDC (Radio Duty Cycling) to become operational raise the energy consumption for approximately 10 mA. This figure has been inserted to make a comparison with the execution of processes involving Web-API architecture.

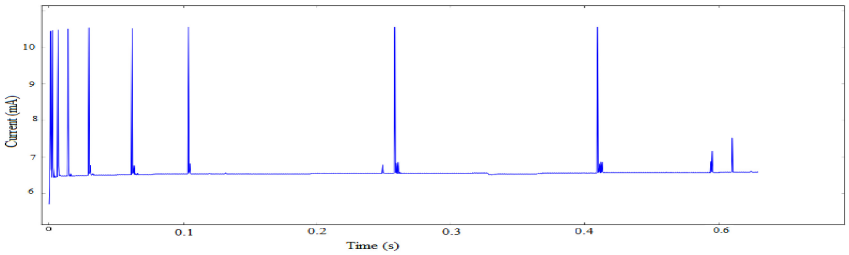


Fig. 4 Stopped Web-API processes.

In Fig. 5 can be seen the implementation of the process of Communication Layer and File Handle. On this chart beyond the usual power consumption peaks

that occurs when the RDC happens, there is CPU usage for the storage of data, as the data is sent every 64 bytes sequence increases of up to 6.5 mA are due to receiving and storing such data. In this instance due to the loss of an ACK acknowledgment packet receipt process the total file took a little longer to complete.

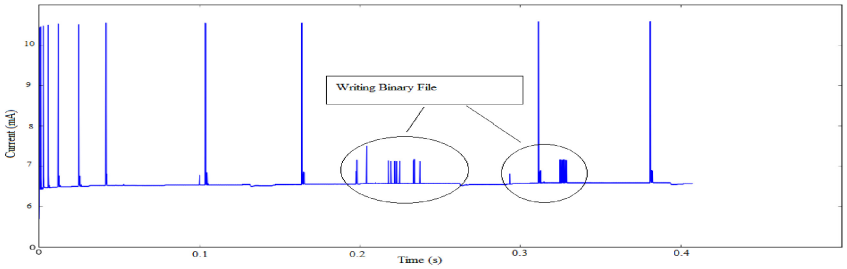


Fig. 5 Writing and receiving data in ROM, 64 bytes at a time.

Finally, it was analyzed that the AutoRun Layer, Fig. 6, where it can be seen that the energy consumption exceeds 20 mA at the time of the execution of the program received the call. After running the binary execution file that works with handling device LEDs stabilizes slightly above 5 mA per second.

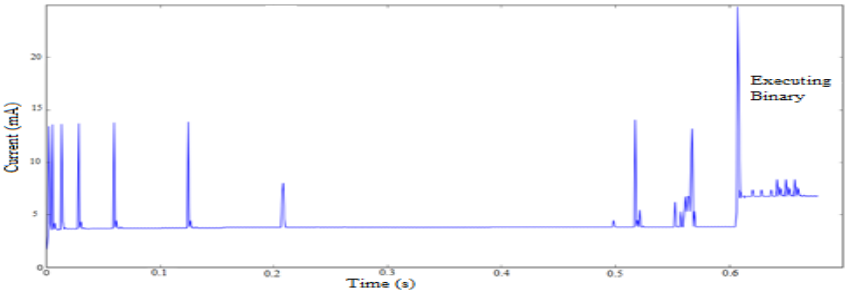


Fig. 6 Running binary file.

Therefore, we can state that the results obtained and displayed in Fig.4, Fig. 5 and Fig. 6 are very similar to those obtained from [2], wherein the devices' reprogramming was performed using the Deluge protocol.

7 Conclusion

In this research work, we implement a Web-API for IoT that performs the self-configuration of devices through a self-configuration policy.

Our Web-API provides two benefits: the first is the development of techniques such as transmission of binary file through the CoAP, which can be useful for systems that want to implement self-configuration. The second is the actual

development of Web-API used the proposed architecture, as once implemented this mechanism can assist in the network configuration of computers that work with wireless sensors and IoT. Thus, if the focus is the analysis of new sensor system configuration may no longer be a complicated step in the completion of the studies, all the attention can be directed to the main objective of the research.

Once running device configuration, power consumption tests were performed using the PowerTrace which is a software that allows measurement with 94% accuracy and memory through the resources of Cooja simulator that allowed validate the proposal. Although the device requested (TMotesky) could not work with the proposed architecture due to lack of resources, it is noted that other motes (Z1, Wismote) performed or could run code designed to validate this work, but were necessary additional implementations, as the implementation of separate response in the Client, and these tests (energy consumption during data transmission and during the execution of binary files) could only be made in Cooja simulator.

The tests in this article were made in 5 motes that sent and received 64-byte packets over the network through the Cooja simulator. Thus, was measured the power consumption of the sending and receiving of these 64-byte packets.

Thus, for future work is important to take this experiment to the real world using embedded devices and performing tests on local interference with communication signals and problems that cannot be simulated. Another future work would be the adaptation of the CoAP-13 to only serve for the purpose of self-configuration, this adaptation could effectively decrease the size of the communication layer code and make the use of Web-API architecture in smaller devices. Finally, introduce the concepts of context awareness and adaptation to Web-API architecture so that it can do self-repair, self-optimizing and self-manage. Making intelligent devices, self-sufficient in the world of pervasive computing.

References

1. Barros, E.B.C., Ribeiro, A.R.L.: A self-configuration architecture for web-API of internet of things. In: Proc. of the 10th International Conference on Web Information Systems and Technologies, Barcelona, pp. 328–334 (2014)
2. Dunkels, A., Ericsson, J., Voigt, T.: Run-Time Dynamic Linking For Reprogramming Wireless Sensor Network. ACM SenSys (2006)
3. Kephart, J.O.: The vision of Autonomic Computing. IEEE Computer Society (2003)
4. Kovatsch, M., Duquennoy, S., Dunkels, A.: A low-power CoAP for Contiki. ACM SenSys (2011)
5. Kovatsch, M., Lanter, M., Shelby, Z.: Californium: Scalable Cloud Services for the Internet of Things with CoAP (2011)
6. Zeng, D., Guo, S, Cheng, Z.: The Web of Things: A Survey. Journal of Communications **6**, September 2011
7. Parachar, M., Hariri, S.: Autonomic Computing: An Overview. Spring (2005)
8. Pauli, D., Obersteg, D.: Californium. Institute for Pervasive Computing. Departament of Computer Science. ETH Zurich. Lab Project (2011)

Automated Behavioral Malware Analysis System

Saja Alqurashi and Omar Batarfi

Abstract Nowadays, with the spread of internet and network-based services, malware has become a major threat to computers and information systems. Actually, different malware share similar behaviours, also they have different syntactic structures due to the incorporation of obfuscation techniques such as polymorphism, Oligomorphic and meta-morphism. The different structure of same behavioral malware poses a serious problem to signature-based detection techniques. In this paper we propose an automated prevention system based on malware behaviours. Our system has the ability to collect suspicious software from client computers, then to automatically analyses the behaviour of detected malware. Then agent then sends an alarm to all network clients. The results from an implementation of the proposed system show that our approach is effective in analysing detected malware in automated security systems.

Keywords Malware · Automated security · Behaviour -based method · Behaviour analysis · Agent

1 Introduction

Malware has become a major security threat and the underlying root cause of most internet security problems. With the increasing spread of internet and network-based services, the malware threat is increasing. This speedy growth of malware has resulted in an increased demand for an automated security system [1]. Malware is a general term used to refer to a malicious code such as a virus, worms, a Trojan, spyware, adware, or any malicious program [2]. Since 1960, malware has evolved into the most significant threat in the computer network

S. Alqurashi(✉) · O. Batarfi
Information Technology Department, King Abdul Aziz University,
KSA, Jeddah, Saudi Arabia
e-mail: {ssmalqurashi,obatarfi}@kau.edu.sa

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_107

domain especially in the last three decades [1]. The complexity of modern malware is making this problem more difficult. The rapid malware generation requires automated and robust security techniques to increase the protection of computer systems [3].

Various approaches have been proposed for malware detection: Signature-based methods, behaviour-based methods, and heuristic methods.

1. Signature-based method: This is the most popular method for malware detection. Signature is a unique feature for each malware. Signature-based methods extract signatures from various malwares to identify them.
2. Behavior-based method: This involves observing the behavior of a software to conclude whether it is malicious or not The main disadvantages of this method are the non-availability of promising False Positive Ratio (FPR) and a longer time in the scanning process.
3. Heuristic method: This method uses data mining and machine learning techniques to learn the behavior of an executable file [4].

The malware analysis knows how malware functions allow for better defenses to protect resources from any attack of malware. There are two types of malware analysis: static analysis or dynamic analysis. Static analysis is the actual viewing of code to get a better understanding of the malware and what it is [5]. Dynamic analysis is how the malware behaves when executed [5].

To overcome this limitation in traditional protection mechanisms based on static analysis, we are proposing a framework for behaviour-based detection and classification of malware in a local network environment. We propose using an agent based system to monitor each computer in the network; if there is any suspicious software the system will analyse this software dynamically based on its behaviour, then trigger the system if there is a malware attack and then send an alarm with information about the malware to other computers in the network to protect them from the possible harmful effects of the malware.

2 Related Work

There are currently many studies that discuss an automated malware analysis in different ways. In this section we review some studies that discuss malware analysis.

Peidai et al. [6] proposed a new framework for malware analysis. They perform analysis the two type of analysis.Static analysis to obtain basic information of malware; and the dynamic analysis to record all execution activities which are monitored when it is running in a restricted environment. I Panad consists of an SA Module that is responsible for static analysis and the main analyzer responsible for dynamic analysis includes information flow tracking, system call monitoring and analysis of network behavior. AsisModule is an assisting module for connecting between Main and SA Module. This work has a high ability to analyze the malware statically to get all paths in software. Furthermore it uses dynamic analysis to capture malware activities that may evade itself during static analysis [6].

Yosha et al.[7] proposed new schema for behavior analysis for malware. In [7] their schema the focus of analysis has three positions: system Folder , window Folder and temporary folder. Behavior analysis occurs during API hooking technique. Hooking API calls allows us to intercept an API call and redirect it to another function. By hooking API calls we obtain useful information about them. In this work [7] has high rate True positive rate and low False Positive rate but their experiment applied on 83 malware that small sample to prove the validity of proposed schema.

Cesar et al. [8] proposed a solution for creating a technique that allows automatic malicious code analysis with a high level of accuracy by using sandbox and machine learning to automate malicious code identification. Firstly, it collects malicious and non-malicious code and analyses these samples using sandboxing to generate activity reports. After that, it uses data mining to classify the reports. Also, it applies machine learning for automation where no longer suffer human intervention and will be automated by machine learning techniques. The following figure gives the scenario of the proposed solution [8]. The con of the proposed system if we implement it is that it is not fully automatic since it requires malware analyst support.

All previous proposed solutions show us different ways to analyze the behavior of malware. Our work shows how to protect other computers in the network by analyzing the suspicious software based on its behavior, if it is malware we have to alert other users in the network.

3 Methodology

Our system comprises of agent software that acts as a monitor for the Registry Window. We chose Registry because it is at the heart of the Windows Operating System in that any change that is made in the system appears in the Registry. In our proposed system, agents scan the Registry periodically. Each scan compares the current state image of Registry with a clean state image. If there is any suspicious software found in the Windows Registry, the agent sends this software into our system. Then, the system analyses the behaviour of the suspicious software using an automated behaviour analysis tool. If the behaviour is found to be malicious, the system considers it as malware, then sends the information about it to all computers in the network, in order to alert other users about the malware and to seek advice.

4 Experiment

To create a dataset, we chose large malware datasets with obfuscated and unknown malware. Those used in this research study consisted of 500 pieces of malware collected from Heavens [9].

We tested our proposed system in a local network consisting of 20 computers using the Windows Operating system. We implemented agents to monitor each computer in the network and to check for any change in the computer, and to list such changes as suspicious software. During this stage we used the Regshot tool that has the ability to list changes in the Registry [10]. These agents then sent reports of the suspicious software to our system in order to analyze their behavior.

In the analysis phase we ran these suspicious pieces of software in separate machine and analyzed dynamically. We used separate machines to protect our machines from any harm that might have occurred from the use of suspicious software. In this phase we used the Anubis tool in order to analyze the suspicious software dynamically [17]. The execution of the Anubis resulted in the generation of a report file that contained enough information to give the user a very good impression of the purpose and the actions of the analyzed binary [11]. Furthermore our proposed system deployed alarm messages which were sent to all computers in the local network by the agents. This alarm contained information about the detected malware as shown in Fig. 1.

Fig. 1 Warning message sent to all computers in the network

5 Results and Discussion

In this section we set out how our proposed system can analyze the malware which it detected in the clients' computers. It then shares the malware information with other computers by providing an alarm as shown in Fig. 1.

In this work we used a Heavens database [9]. We tested our system for 500 executable files in the form of malware and 100 as clean files that existed in a client's computer connected to the main computer contained in our system. We evaluated the proposed system based on the True Positive rate and the False Positive rate. The test results are as shown in Table 1.

Table 1 Result of True Positive an False Positive Rates

	Sample size	TP	FP
malware	500	70%	30%
benign	100	97%	3%

Our experimental results show that the true positive rate is 70%. In our experiment, we used Anubis to analyse the behaviour of the software. So we presume that this low true positive rate is due to the Anubis tool. However, we presume that the use of the Anubis tool could lead to the high False Positive rate (30%). In fact ,malware can be stop its malicious activities, there is malware that attempts to evade an analyser tool when they detect the analyser tool in action in the running environment.

Our study showed the possibility of building an automated security system for local networks. Our study showed how our proposed system analyzes detected malware based on the behavior method and how it shares its information with the automated security system. This sharing provides the ability to develop an alarm to protect other computers in the network.

We summarize our findings from the proposed system as follows:

Analyze the behavior of the detected malware.

Develop an alarm to protect other computers in the network.

Share the detected malware information with network partners and peers.

The limitation with regard to our proposed system is that it has a high False Positive rate and considers some clean software as malware - we will work on this in future. Also, our system doesn't have the ability to classify malware according to family. This too will require work in future.

In conclusion, we confirm that automated security systems have become increasingly essential for any network environment. In terms of our system, we provide users with the ability to install an automated prevention system.

6 Conclusion

Nowadays, malware has become a major threat to computers and information systems on the internet. Malware has different structures but with the same behavior which poses a serious problem to signature-based detection techniques. In this paper, we discuss how the proposed system develops and disseminates warnings to all users in the network with regard to malware that affects networked computers. Our system automatically detects malware in each computer. The agent collects suspicious software and stores it in a system database, after which the proposed system dynamically analyses the behavior of the suspicious software. The proposed system then sends a warning to all computers in the network about this malware. We tested our proposed approach with a Heavens dataset and found our system to be effective in analyzing the malware. This then allowed us to share the malware information as a means of developing a warning system to protect other computers. However, our system has a limitation in that it cannot classify malware from similar families. Our proposed system offers any organization a means of preventing malware attacks and protecting users' computers from any malware by developing a warning system that allows other users to protect their computers. In future, we will work on malware classification based on the Heuristic method in order to mitigate this limitation.

References

1. Kumar, S., Rama Krishna, C., Aggarwal, N., Sehgal, R., Chamotra, S.: Malicious data classification using structural information and behavioral specifications in executables. 2014 Recent Adv. Eng. Comput. Sci., pp. 1–6, 2014 (October 2015)
2. Santos, I., Penya, Y., Devesa, J., Bringas, P.: N-grams-based File Signatures for Malware Detection. Iceis **2**, 317–320 (2009)
3. Bailey, M., Oberheide, J., Andersen, J., Mao, Z.M., Jahanian, F., Nazario, J.: Automated Classification and Analysis of Internet Malware. Recent Adv. Intrusion Detect. **4637**, 178–197 (2007)
4. Bazrafshan, Z., Hashemi, H., Fard, S.M.H., Hamzeh, A.: A survey on heuristic malware detection techniques. In: The 5th Conference on Information and Knowledge Technology, pp. 113–120 (2013)
5. Egele, M., Scholte, T., Kirda, E., Barbara, S.: A survey on automated dynamic malware analysis techniques and tools. ACM Comput. Surv. **V**, 1–49 (2011)
6. Xie, P., Lu, X., Su, J., Wang, Y., Li, M., Xie, P., Lu, X., Su, J., Wang, Y., Li, M., Xie, P., Lu, X., Su, J., Wang, Y., Li, M.: iPanda: a comprehensive malware analysis tool. In: The International Conference on Information Networking 2013 (ICOIN), pp. 481–486 (2013)
7. Yoshiro Fukushima, K.S., Sakai, A., Hori, Y.: A behavior based malware detection scheme for avoiding false positive. In: 2010 6th IEEE Work. Secur. Netw. Protoc., pp. 79–84 (2010)
8. de Andrade, C.A.B., de Mello, C.G., Duarte, J.C.: Malware automatic analysis. In: 2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence, pp. 681–686 (2013)
9. Virus collection (VX heaven). http://vxheaven.org/vl.php (accessed: October 25, 2015)
10. Ravula, R.R.: Classification of Malware using Reverse Engineering and Data Mining Techniques (2011)
11. Anubis–Malware Analysis for Unknown Binaries. https://anubis.iseclab.org/ (accessed: October 25, 2015)

A Message Efficient Group Membership Protocol in Synchronous Distributed Systems

SungHoon Park, SuChang Yoo, YeongMok Kim, SangGwon Lee and DoWon Kim

Abstract In distributed systems, a group of computer should continue to do cooperation in order to finish some jobs. In such a system, a group membership protocol is especially practical and important elements to provide processes in a group with a consistent common knowledge about the membership of the group. Whenever a membership change occurs, processes should agree on which of them should do to accomplish an unfinished job or begin a new job. The problem of knowing a stable membership view is very same with the agreeing common predicate in a distributed system such as the consensus problem. Based on the termination detection protocol that is traditional one in asynchronous distributed systems, we present the new group membership protocol in arbitrary wired networks.

Keywords Synchronous distributed systems · Group membership · Fault tolerance · Wired arbitrary network environment

1 Introduction

In distributed systems, a group of computer should continue to do cooperation in order to finish some jobs. A group membership protocol is especially helpful tools to allocate processes in a same group with a same view of the membership of the group. Whenever a membership change occurs, processes can consent to which of them should do to finish a waiting job or begin a new job. The problem of getting a stable membership view is very same with the one of getting common knowledge in a synchronous distributed system such as the consensus problem [1].

S. Park(✉) · S. Yoo · Y. Kim · S. Lee · D. Kim
School of Electrical and Computer Engineering,
Chungbuk National University, Cheongju, Korea
e-mail: spark@chungbuk.ac.kr

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_108

The Group membership protocol [2] is that every process connected in a network requires getting a stable same group membership view if all connected process are belong to just one group. The problem was widely discussed at the study community. The reason for this great study is that many distributed systems need a group membership protocol [3,4,5,6,7]. In spite of such practically usefulness, to our knowledge there is only a few research that have been committed to this problem in a wired arbitrary connected computing environment.

Depending on process failure and recover, network topology is changed and process may dynamically connect and disconnect over a wired network. In such wired networks, group membership can be changed so much, making it a special critical module of system software part. In wired arbitrary network systems, a lot of environmental adversities are more common than the static wired network systems such as that can cause loss of messages or data [8]. In particular, a process can easily get to fault by hardware or software problem and disconnect from the wired network. Implementing fault-tolerant distributed applications in such an environment is a complex and difficult behavior [9,10].

In this paper, we propose a new protocol to the group membership protocol in a specific wired distributed computing system. Based on the termination detection protocol that is traditional one in asynchronous distributed systems, we address the new group membership protocol. We make up of the rest of this paper as follows. In Section 2 we address the system model we use. In Section 3, we describe a specification to the group membership problem in a traditional synchronous distributed system. We also address a new protocol to solve the group membership problem in a wired arbitrary computing system in Section 4. In Section 5, we address conclude.

2 Computing System Model, Definition and Assumption

In this section, we describe our models for capturing behavior of distributed systems. We use these models foe reasoning about correctness of our protocol as well as for analysis of distributed computations. Our model for distributed systems is based on message passing, and all of protocol is around that concept. Many of these kinds of protocol have analogs in the shared memory computing system but will not be addressed in this paper.

First, we define our system model based on some assumptions and after that we address our goals. We model a distributed system as a loosely coupled messing-passing system without shared memory and a global clock. Our distributed computation model for a wired network is made up of as an undirected graph. That is, the undirected graph is described as G = (V, E), in which vertices V facing each other with set of process {1, 2,….,n} (n >1) with unique identifiers and edges E between a pair of process correspond the fact that the two process are in each other's transmission radii. Hence, our distributed system has a channel to directly communicate with each other which changes over time when processes move.

Every process i has a variable Ni, which denotes the neighboring processes, with that i can directly communicate the neighboring processes. Every process communicates with a channel that is bidirectional; $j \in Ni$ iff $i \in Nj$. More accurately, in the network G = (V, E), we decide E such that for all $i \in V$, $(i, j) \in E$ if and only if $i \in Nj$. Depending on process's movement, the graph could be disconnected that means that the network is partitioned. Because the processes may alternate their position, Ni position would be unexpectedly changed and therefore G also may be changed accordingly. The assumptions about the processes, wired network and system architecture are followings.

Without network partition, the sender and the receiver do successful message delivery that means the message would be successfully delivered only when the two processes remain connected for the all period of message transfer. Every process has a big receiving buffer enough to avoid buffer overflow all the time in its lifetime. Even though a finite number of topology changes, every process i eventually has a same view of group membership of the group to which i belongs.

3 Group Membership Specification

We assume that our specification is as followings, it is consist of four properties for a group membership protocol.

Safety(1) : At any time, all processes in the group have a stable consistent view.

Progress : If there are no more changes in the each views of the processes in one group, they eventually getting to their stable consistent views.

Validity : If all processes in a event know a view as their local view and they have eventually reached their stable states, then the last process of their sequences of global views are all at same position and must be equal to each other.

Safety(2) : When a view is committed as a global view, it cannot be changed.

The first property describes agreement. Consistent history must be an unchanged one for any program that satisfies the specification. The second property shows termination of global view. When the state and event of all processes are unchanged, the processes are eventually getting to close changing their output results. The third property removes trivial solutions where protocols never getting on any new view or always determine on the consistent view.

4 Group Membership Protocol in Wired Network

At this section, we address a group membership protocol that was operated upon the termination detection protocol by scattering computations. The first phase that is a diffusing phase and it works by first diffusing the "who" messages. The second phase that is a searching phase and it runs by then accumulating the id of every process that is consist of the wired networks. We represent this computation

starting processes as the start process. The third phase is a closing phase that is managed by deciding the same view and announcing it as a stable new view to all process. The three kinds of message, Who, Ack and View are used to manipulate the operations. As the first phase is diffusing computing phase, Who message is used to make a start of the group membership protocol by diffusing the Who message.

1) The first Phase: When group membership protocol is launched at a start process s, the start process makes a replying queue wl and a accepted queue rl and starts a scattering computation by forwarding a Who message to all of its immediate neighboring processes. At the starting point, the replying queue makes up of only its most close neighboring process's ids and the accepted queue has nothing.

When process i receives a Who message from the neighboring process for the first time, it immediately sends the Ack message to the start process and propagates the Who message to all its neighboring process except the process from which it first accepted an Who message. The Ack message sent by process i to the start process contains the ids of all its neighboring process that are needed for the start process to decide the stable view of the process connected with a distributed network. After that, any Who message accepted by other neighboring process will be ignored.

2) The Second Phase: Searching phase. When the start process receives the Ack message was taken out from the process j, it takes j out from the replying queue and gets j into the accepted queue and as soon as possible it detects sequentially the each process's id included in the Ack message. If there is the some process in the Ack message which has already been accepted, i.e. that means it is in the accepted queue, it is dismissed. If it is not in the accepted queue, it is inserted into the replying queue of start process. The start process will be suspends for the Ack message from one.

The replying queue is increasing and decreasing repeatedly when it was accepted based on the accepted Ack messages, however the replying queue is continually increasing by accepting the Ack messages. But the replying queue at the end could have no element and the replying queue could insert all ids of processes connected to the wired networks whenever the start process accepted the Ack messages from all other processes. Therefore the start process eventually has much information enough to decide the stable view of the group based on the replying queue. That is because the replying queue could be eventually unoccupied and it means that the start process has accepted the Ack messages from all the process.

3) The Third Phase: Once the start process has accepted Acks from all other process, it decides the stable view based on the replying queue and forwards a View message to all other process to let know the current view of the group. We show some sample running protocol as the protocol execution to explain more specific features. We address the protocol in synchronous setting even though all the behaviors of the protocol are practically asynchronous. We assume that the network shown in Figure 1(a) is asynchronous. In this shape, and for the all of the paper, thin arrows denote the route of Who message's move and dotted arrows denotes the way of route of Ack messages to the start process. As shown in Figure 1, process A is a start process that starts wla and rlb with {B,C} and {A} at each and

starts a scattering computation with forwarding out Who messages (indicated as "E" in the shape) to its immediate neighbors, viz. process B and C, shown in Figure 1(a). As indicated in Figure 1(b), process B and C in turn forward the Who message to its most close neighbors only except the start process. It sends the Ack message with close neighboring process queue to the start process A. Hence B and C also send Who messages to each other. But B and C do not acknowledge to the start process about the Who messages because process B and C have already accepted Who messages from the start process at each. The information of neighboring process is piggybacked upon the Ack message sent by all process. Upon hearing Ack messages from B and C, process A renews wla = { B,C }, rlb = { A } with the close neighboring process information piggybacked at the Ack message.

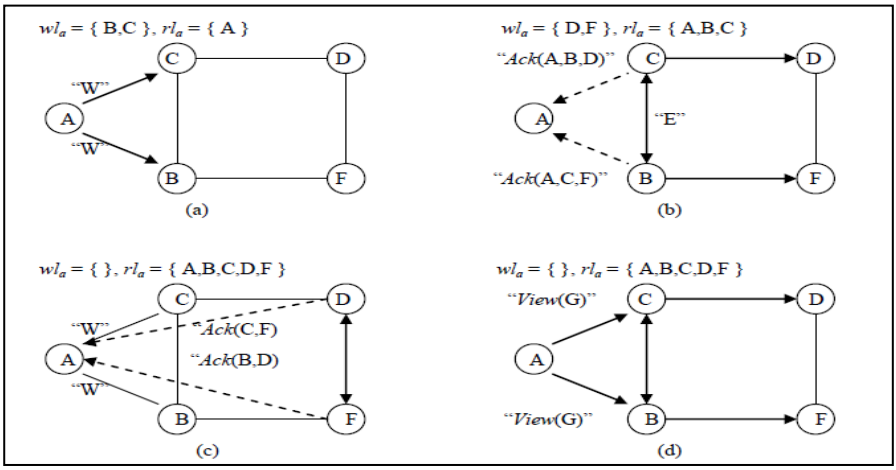


Fig. 1 An example of group membership protocol execution on the process search protocol.

The Who messages is transmitted over the arrows at the edges and the dotted arrows going parallel with the edges denotes Ack messages. In Figure 1(c), the process D and F also send the Ack messages to the starts process at the time they accepted the Who message s from the B and C one by one. Each of these Ack messages includes the ids of the neighbor. All the time, the start A accepts all acknowledgments from all of other process except itself in Figure 1(d) and then determines the stable view between the group and forwards it, that is the View message displayed in Figure 1(d).

5 Conclusion

We have addressed here the study of distributed group membership protocol for the wired networks and proved it to be correct based on the symbolic dynamics of finite state machine obtained by linear probability model. In real world, the wired network topology is actively and lively changing at random and that dynamic network changed configuration causes frequent connection and disconnection of

process over the wired network. In spite of weakness about wired networks, our group membership protocol specification guarantees the safety and progress property could be always satisfied.

As mentioned in the introduction, our main goal has been to design group membership search protocol and prove decidability of consistent view in as simple a fashion as possible without paying much attention at wired networks to the consistent membership view on every process even though complexity issues of wired networks. In particular, a more careful logical design of group membership protocol for specific classes of wired networks can be performed using the results from our design of convergence properties of consistent group membership view. This could lead to a significant improvement of our protocol from a practical environment point.

Finally, in a practical setting one may generalize the group membership protocol to more fit in some distributes systems according to network environments. It will be interesting to explore whether safety or progress could be weakened depending on distributed computing environmental factors.

References

1. Amir, Y., Moser, E., Melliar-Smith, P., Agarwal, D., Ciarfella, P.: The totem single-ring ordering and membership protocol. ACM Trans. Computer Systems. **13**(4), 311–342 (1995)
2. Anceaume, E., Charron-Bost, B., Minet, P., Toueg, S.: On the formal specification of group membership services. Technical Report Computer Science Dept. Cornell Univ. **95**(3), 1534−1559 (1995)
3. Anker, T., Chockler, G., Dolev, D., Keidar, I.: Scalable group membership services for novel applications. In: Proceedings of Workshop on Networks in Distributed Computing (DIMACS 45), pp. 23–42, August 21-23, 1998
4. Keidar, I., Sussman, J., Marzullo, K., Dolev, D.: A client server oriented protocol for virtually synchronous group membership in WANs. In: Proceedings of 20th International Conference Distributed Computing Systems, vol. 3(1), pp. 234–244, April 15-17, 2000
5. Brunekreef, J., Katoen, J., Koymans, R., Mauw, S.: Design and analysis of dynamic leader group membership protocols in broadcast networks. Distributed Computing **9**(4), 157–171 (1996)
6. Bottazi, D., Montanari, R., Rossi, G.: A self-organizing group management middleware for wired networks. Computer Communications **31**(13), 3040–3048 (2008)
7. Powell, D.: Special section on group communication. Communications of the ACM **39**(4), 50–97 (1996)
8. Pradhan, D., Krishna, P., Vaidya, N.: Recoverable environments: design and tradeoff analysis. In: Proceedings of Annual Symposium on Fault Tolerant Computing, pp. 16–25, October 12-15, 1996
9. Briesemeister, L., Hommel, G.: Localized group membership service for wired networks. In: Proceedings of International Conference on IEEE Parallel Processing Workshops, pp. 94–100, March 05-18, 2002
10. Hatzis, K., Pentaris, G., Spirakis, P., Tampakas, V., Tan, R.: Fundamental control protocols in networks. In: Proceedings of 11th ACM SPAA, pp. 251–260, October 19-21, 1999

Ontology-Driven Metamodel Validation in Cyber-Physical Systems

Kevin Lynch, Randall Ramsey, George Ball, Matt Schmit and Kyle Collins

Abstract This paper describes the development and use of an ontology to validate model integration and metamodel description in the cyber-physical systems engineering domain. A primary goal of the Defense Advanced Research Program Agency (DARPA)'s Adaptive Vehicle Make (AVM) program is to reduce product cost by reducing the design space early in the engineering life cycle, to enable a high-value focus on reasonable alternatives. The OpenMETA tool suite uses semantic constructs for model integration across engineering disciplines, and facilitates trade analysis across those disciplines (such as electrical, mechanical, thermal, fluid, and cyber). Independently, through expert interviews, the authors developed a lightweight ontology to represent important trades in the early system engineering design process, and have been using that ontology as both a communication and validation mechanism for the semantics embedded in the OpenMETA tools. The results are informing the tool development, providing an independent validation mechanism, and creating a layer of the ontology that is extremely useful to both people and machines that was not initially obvious. The ontology-based approach described in this paper facilitates the identification and encoding of the important relationships between design decisions and subsequent reasonable decision alternatives, and serves as an independent validation of the semantics and relationships encoded in the OpenMETA tool suite, both at very low cost.

Keywords Ontology · Metamodel · Cyber-physical systems

K. Lynch(✉) · R. Ramsey · G. Ball
Raytheon Engineering and Information Technology, Tucson, AZ, USA
e-mail: {Kevin_J_Lynch,Ramsey,George_Ball}@raytheon.com

M. Schmit · K. Collins
Georgia Institute of Technology Aerospace Systems Design Laboratory, Atlanta, GA, USA
e-mail: {mschmit6,kbc}@gatech.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_109

1 Introduction, Background, and Rationale

Complexity in engineered systems continues to increase, with resulting system designs requiring many years to complete. Costs have increased similarly. The goal of the DARPA Advanced Vehicle Make (AVM) program was to significantly improve the design and manufacturing of complex cyber-physical systems, increasing the speed by five times while reducing cost 1. To shorten development times to this degree, the level of abstraction needs to be raised in the design of cyber-physical systems, designs need to be composed from component model libraries to facilitate reuse, and rapid requirements and design trades have to be executed in near real-time, representing many modeling challenges 2. Understanding system model elements earlier in the life cycle enables the discovery of performance impacts, but the heterogeneity of modeling languages across multiple levels of abstraction, multiple domains, and multiple physics models make the integration at a systems engineering level very difficult 3. Multi-paradigm modeling is one domain-independent framework for addressing these integration issues [4], while orthographic software modeling provides a metaphor for integrating different development paradigms [5]. Vanderbilt University developed the OpenMETA toolset used for AVM and in particular the CyPhyML integration model language to address the heterogeneity, abstraction level, and multi-domain multi-physics nature of the problem 4. The authors' current work extends the original DARPA work by expanding the OpenMETA tool suite and approach into the air vehicle domain, taking cost more explicitly into account.[1]

Vanderbilt University and subsequently Metamorph Inc. have developed and refined the Cyber Physical Modeling Language, known as CyPhyML or CyPhy, for modeling, evaluation, and synthesis of cyber-physical systems 7. CyPhy is the model integration language that provides a semantic backplane of the OpenMETA tool suite and integration platform; semantically precise integration of domains and tools is facilitated using CyPhy, with both structural and behavioral semantics supported 8. The CyPhy language has semantic underpinnings that formalize connectedness and translations between multi-domain, multi-physics models that exist at different levels of fidelity and abstraction. The semantic underpinnings of components did not manifest themselves at a level that was usable by our subject matter experts (SMEs) in propulsion, seekers, and guidance and control, when trying to make trades early in the design process. While the OpenMETA toolset does support those kinds of trades, the authors found applying domain-specific expertise across disciplines within the toolset to be difficult because the vocabulary did not exist to map to the tools. We developed a lightweight ontology to represent the relationships between models, and recognized it can act as an independent validation of the underlying semantics in OpenMETA. The overarching goal of the work is to determine whether the approach proven in the amphibious ground vehicle domain generalizes to another domain. To make that determination, an air vehicle was chosen, with the explicit requirements of improving its flight range and speed,

[1] Work funded by DARPA under Contract D15PC0025, Distribution Statement "A"

while reducing its cost. The air vehicle diameter and length were kept constant to reduce complexity while still exercising trades. The trades incorporated aerodynamics, propulsion, kinematics, and cost models. Trade studies focused on parameters that affect figures of merit, to maximize range and speed, while minimizing cost. These trades guided the development of the ontology, narrowing the set of possible elements to the ones that would most likely be exercised in trades. The ontology approach employed by the authors is a top-down approach. Experts in air vehicle design including propulsion and navigation have been interviewed to identify the important elements in making early design trades. Relationships within and between propulsion, seeker, and airframe models are identified, as well as to a manufacturing cost model. The authors are using subject matter expertise to map these relationships in an effort to validate the link between computational models that exists in the OpenMETA tool suite. Linking these models is important for reasoning across these models computationally. Special focus is placed on the relationships between elements in different domains. These relationships can affect development time and cost in unexpected and significantly detrimental ways. Exposing these relationships early in the design cycle can reduce costly oversights that occur when incompatibilities between elements in different domains are overlooked by engineers during conceptual and preliminary design. The lightweight ontology being developed has the semantic sensor network (SSN) ontology as its base 9, and makes use of the QUDT, Owl-Time, WGS84, and SWEET ontologies, as described in 10.

2 Conclusion

Metamodels are important to achieving the goals of shortened development times, reuse of design knowledge, flexibility of abstraction and fidelity, and reduced cost for complex cyber-physical systems. This paper describes one approach for validating those metamodels involving expert interviews and the development of a lightweight ontology to focus on the important elements in the early design trade space. Example trades were investigated, and the elements involved in those trades were mapped between the lightweight ontology and the OpenMETA tool suite. While not formal or rigorous, the approach and results suggest that low-cost ontology development may indeed be useful for validation of metamodels, when subject matter expertise is available.

References

1. Eremenko, P.: Philosophical underpinnings of Adaptive Vehicle Make, DARPA-BAA-12-15, Appendix 1, December 5, 2011
2. Derler, P., Lee, E., Sangiovanni-Vincentelli, A.: Addressing modeling challenges in cyber-physical systems, Technical report no. UCB/EECS-2011-17, Electrical Engineering and Computer Sciences, University of California at Berkeley, March 4, 2011

3. Karsai, G.: Unification or integration? The challenge of semantics in heterogeneous modeling languages. In: GEMOC 2014, 2 (2014)
4. Mosterman, P.J., Vangheluwe, H.: Computer automated multi-paradigm modeling: An introduction. Simulation **80**(9), 433–450 (2004)
5. Atkinson, C., Tunjic, C., Stoll, D., Robin, J.: A prototype implementation of an orthographic software modeling environment. In: Proceedings of the 1st Workshop on View-Based, Aspect-Oriented and Orthographic Software Modelling, p. 3. ACM (2013)
6. Neema, S., Bapty, T., Scott, J.: VU-ISIS final report: meta tools extension and maturation, Institute for Software Integrated Systems, Vanderbilt University, December 2014
7. Sztipanovits, J., Koutsoukos, X., Bapty, T., Neema, S., Jackson, E.: Design tool chain for cyber-physical systems: lessons learned. In: DAC 2015, June 07–11, 2015, San Francisco, CA, USA (2015)
8. Simko, G., Levendovszky, T., Neema, S., Jackson, E., Bapty, T., Porter, J., Sztipanovits, J.: Foundation for model integration: semantic backplane. In: ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, pp. 1077–1086. American Society of Mechanical Engineers (2012)
9. Compton, M., Barnaghi, P., Bermudez, L., García-Castro, R., Corcho, O., Cox, S., Taylor, K.: The SSN ontology of the W3C semantic sensor network incubator group. Web Semantics: Science, Services and Agents on the World Wide Web **17**, 25–32 (2012)
10. Calbimonte, J., Jeung, H., Corcho, O., Aberer, K.: Enabling query technologies for the semantic sensor web. International Journal on Semantic Web and Information Systems **8**(EPFL-ARTICLE-183971), 43–63 (2012)

Developing Software in the Academic Environment

A Framework for Software Development at the University

William Phillips, Shruthi Subramani, Anusha Gorantla and Victoria Phillips

Abstract This paper proposes a framework used for real-world software development in the academic environment of Fairleigh Dickinson University (FDU). The framework thus far has been used to establish the functional baseline for the Predictive Wastewater Management System (PWMS), a participatory sensing system for monitoring and predicting problems occurring in a municipality's wastewater system for the Eastech Corporation. As software development becomes globally distributed, students need immersion in a realistic software development lifecycle (SDLC) and experience with a framework targeted toward this emerging paradigm.

Keywords Software engineering · Software configuration management · Hybrid agile software development · Model driven software development

1 Introduction

Fairleigh Dickinson University (FDU) was approached by the Eastech Company to develop a predictive participatory sensing system in order to identify potential problems in a municipality's wastewater management system using Eastech flow-cell sensors. Eastech asked FDU to deliver a system that would solve this

W. Phillips(✉) · S. Subramani · A. Gorantla
G. Haase School of Computer Sciences and Engineering,
Fairleigh Dickinson University, Teaneck, NJ, USA
e-mail: william581_phillips@fdu.edu, {sshruthi,anusha13}@student.fdu.edu

V. Phillips
Information Technologies Department, County College of Morris, Randolph, NJ, USA
e-mail: phillips.victoria@student.ccm.edu

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_110

problem. A team was formed of graduate computer science students and one faculty member. A number of issues had to be addressed before a software product with integrity could be delivered.

It was agreed that FDU would deliver a set of Requirements Analysis Documents (RAD) and a set of Unified Modeling Language (UML) [1] models to establish the functional baseline for the Predictive Wastewater Management System (PWMS) [2]. The purpose of the functional baseline is to specify and model an implementation- agnostic description of what the PWMS is to do, with functional and nonfunctional requirements partitioned into subsystems. The proposed framework (i.e., the processes, software development lifecycle (SDLC), deliverables, and tools for software development) is presented. Software configuration management (SCM) for the academic environment, derived from establishing the functional baseline for the PWMS, is described.

In related efforts, the approach is generally prescriptive in that existing methodologies are selected and applied. In this paper, important components are selected from the collection of approaches described in the literature and accepted in industry. This paper describes the process of creating a framework, customized for the academic (and perhaps, global) development environment, as work progresses on external or internal-to-FDU projects. Since the resulting framework is created from doing software work in the academic environment, it follows that the same framework would be a good starting point for other academic efforts, as opposed to an ad-hoc approach.

2 Background

Developing software in the academic environment has a number of positive and negative tradeoffs. On the positive side, there is a large pool of highly motivated talent. Students pursuing an MSCS have varying degrees of real-world work experience. Project team members are selected based on interest, work experience and classroom performance. Classroom performance is a good indicator of project performance.

High student turnover is a drawback. Students can expect to work on a project for about a year before graduating and leaving the project. In universities that offer a Ph.D. in CS, students should be placed in key positions to ensure continuity as MSCS students join and leave the project.

Risk management prescribes a disciplined approach to maintain product integrity [3][4][5]. The foundation for creating such a disciplined environment is the establishment, early on, of a plan for managing the software project, configuration and quality. Typically, the plan is captured in the Software Development Plan (SDP). The Software Configuration Management Plan (SCMP), Software Quality Assurance Plan (SQAP), and the Software Quality Evaluation Plan (SQEP) either can be separate documents or sections within the SDP. The recommendation to approach software development in a disciplined manner or not attempt it at all [5] especially pertains to the academic environment. However,

if managed optimally, the academic environment is not only an excellent setting to incubate frameworks and methodologies, but a place where tools (e.g., model evolution tracing tools) can be developed, evaluated and tested [6][7][8][9][10].

The PWMS was initially decomposed into the following subsystems:

- Data Analysis Subsystem (DAS) – A framework that is used by the Data Analyst actor to create and deploy algorithms that use flowcell data [12].
- Data Gathering Subsystem (DGS) – The PWMS subsystem that enables the Data Gatherer actor to specify from which flowcells to capture data and what data is to be captured [13].
- Data Storage Subsystem (DSS) – The PWMS subsystem responsible for storing and retrieving the water level and other environmental data from the Eastech flowcells [14].
- Access Control Subsystem (ACS) – The PWMS subsystem that defines actors for PWMS and manages access to the different PWMS subsystems [15].
- Audit and Alert Subsystem (AAS) – The PWMS subsystem that alerts the Watchman actor and allows an Auditor actor to query the history of PWMS auditable and alertable events [16].

Model driven software development begins with the creation of use cases, typically from scenarios. From use cases, the dynamic models for the PWMS system and for each of the PWMS subsystems were created [4]. Model consistency is provided by adhering to the processes, conventions and procedures used to review the RADs and models before promotion to formal internal control.

An SDP [17], SCMP [18], and SQAP [19] were created for the PWMS project. These plans are under informal internal control, meaning they are still under development and are evolving as work progresses. A final report [20] was delivered identifying the current state of the PWMS.

3 Lessons Learned from the PWMS Project

The importance of establishing software configuration management plans and procedures before placing the artifacts of the functional baseline under formal internal control cannot be overemphasized. Aiello explains why software configuration management is important throughout the software development effort [21][22]. Initially a manual paper control process is used, with automation of the process expected once it has been defined, used and tested.

Defining the review processes is a responsibility of the software quality assurance function and requires the creation of a SQAP. The first incantation of the PWMS SQAP consisted of the review criteria for the functional baseline artifacts (i.e., the RADs and UML analysis models). Later the SQAP will refer to the criteria used in the SQA gatekeeping function, typically in the form of checklists, and the SQAP will provide a description of the processes to be followed. To preserve

modifications to a document or model, a responsible engineer was assigned, and only the responsible engineer was allowed to make changes to the artifact.

Client expectations are another factor that require management. The corporate customer should realize that a software development project is not a trivial undertaking in both the academic and real-world settings. However, if appropriate software engineering controls are established and adhered to, the academic environment can be a viable environment for software development, with positive outcomes possible for all of the stakeholders.

4 Software Configuration Management of the PWMS

It is important to establish the configuration control board (CCB) as early as possible in the SDLC. Stakeholders' participation in the CCB is how their contributions are communicated and documented as the project progresses [5]. No changes to the software configuration shall be made unless authorized by the CCB and a software change request form is created.

After the allocated baseline is established, commercial off the shelf (COTS) products are procured. For example, the Data Storage Subsystem may be implemented by a COTS database management system (DBMS). Other subsystems require varying degrees of software development. Software to be developed is specified in accordance with standard object oriented (OO) development methodologies. A hybrid agile approach is used at this point. Subsets of requirements to be implemented will be identified and implemented, thereby establishing incrementally the design and product baselines with each agile run. To support software development in the academic environment, agile runs shall be scheduled such that an agile run is performed with a minimum number of interruptions. The term "agile run" is used as opposed to "agile sprint," "sprint" being a term associated with SCRUM.

Aiello points out that hybrid agile is a viable solution in developing and deploying software in the banking/financial arena [21][22]. Hybrid agility enables us to maintain the discipline enforced by following waterfall, while allowing developers to have episodes of less restraint enabling efficient development and refactoring.

When agile runs are done, a software engineer associated with the SCM function will be assigned to the team. This software engineer will run the model evolution tool and will be responsible for maintaining traceability of the requirements being implemented (i.e., horizontal traceability) and traceability as the models evolve (i.e., vertical traceability). Because of the relatively volatile academic environment, the periods of agility will be scheduled with minimal interruption.

Aiello defines a fifth SCM function, in addition to the classic four (configuration identification, change control, configuration audit and status accounting) [5], to include management and control of the operational baseline. How the organization in charge of the operational baseline (operations) is to effectively work with the development group to ensure operational baseline integrity is defined as

DevOps [21][22]. A methodology for ensuring product integrity for the operational baseline must be defined, as the multiple agile runs result in a product baseline ready to be deployed (i.e., promoted from formal internal to formal external control) and released from development, thereby creating the operational baseline.

5 Traceability of Evolving UML Models

S. Wenzel discusses SiDiff in [6]. SiDiff is a kernel providing functions that enable the tracing of changes to models that can be translated into a graphical representation. Among the models that can be traced are UML models. As shown in Fig. 1, a model evolution capture tool is required to capture changes to the design baseline's UML models (i.e., class diagrams) after an agile run. Solid boxes indicate validated phases of the SDLC as they have produced PWMS deliverables.

Developing tools based on SiDiff kernel, as opposed to an Eclipse plug-in, is the optimal approach for this framework. The SiDiff philosophy of maintaining a kernel that provides an infrastructure used to develop the tools needed for tracing the evolution of a model is required [6]. These changes may be documented by reverse engineering the UML implementation class diagrams from the changed code and then running a model evolution tool to identify the changes made during each agile run. The goal is to maintain traceability of the requirements through the software configuration. Collaboration in developing tools of this nature with any interested academic or industrial partner is enthusiastically invited.

6 Testing the Artifacts Created by the Proposed Framework

Since the framework has been developed to the point of delivering the functional baseline, testing is limited to the reviews that were performed on the RADs. Tests shall be developed based on the requirements derived from use cases and the non-functional requirements set forth in the functional baseline. These tests would be selected for the subset of requirements implemented for each agile run. Integration and unit tests will be created, run and captured as classes are created and refactored during each agile run. Regression testing shall be performed on each of the baselines as appropriate. In addition to updating the SQAP, creation of the SQEP and criteria for establishing test and evaluation plans and procedures will be established.

7 Proposal Evaluation

A proposal evaluation was performed in accordance with the criteria set forth at http://www.secf.org. The results of the proposal evaluation are shown in Fig. 2.

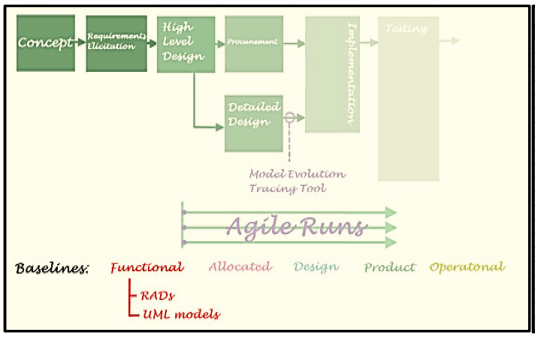


Fig. 1 PWMS SDLC

PROPOSAL EVALUATION for the paper Developing Software in the Academic Environment 1/12/2016					
Criteria	Reviewer1	Reviewer2	Reviewer3	Reviewer4	AVERAGE
Demonstration of a real need or Problem	4	5	4	5	4.5
Timeliness, urgency	5	4	5	4	4.5
Clear, tangible outcomes	5	5	5	4	4.75
Sound methodologies	5	5	5	5	5
Organizational credibility	5	5	5	4	4.75
Staffing	5	4	5	5	4.75
Participation	5	5	5	5	5
Collaboration	5	5	5	4	4.75
Innovation/Creativity/Uniqueness	5	5	5	5	5
Multicultural/Intergenerational	5	5	5	4	4.75
Evaluation plan	5	5	4	5	4.75
Dissemination	5	5	5	5	5
Replicability	5	5	5	5	5
Sustainability	5	5	5	5	5
Tokited contributions	5	5	5	5	5
Technology	5	5	5	5	5
Overall value	5	5	5	5	5
Fit with funder	5	5	4	5	4.75
Proposal clarity, organization, and completeness	5	5	5	5	5
Visual Presentation	5	5	5	5	5
TOTAL	99	98	97	95	97.25

Fig. 2 Proposal evaluation results

8 Pedagogical Outcomes

The PWMS project has provided real-world experience to a number of FDU graduate students, introducing them to the disciplined development, procedures and reviews encountered on a professional software development project. The database of software engineering artifacts resulting from the PWMS project provides a body of knowledge used for software engineering, software configuration management and software quality assurance courses at FDU. Students discover that the "soft" skills taught as part of a software engineering curriculum become difficult when they are put into practice. Students discover that working on a project requires acting as both an effective team member and as an individual contributor as circumstances dictate.

9 Future Work

It has been demonstrated that software can be developed in the academic environment, at least through the requirements elicitation phase, with establishment of the functional baseline. The functional baseline, composed of the RADs and supporting functional, dynamic and UML analysis models [4], can be used by an industrial customer to subcontract all or part of subsystems comprising a project, or the company may decide to outsource or continue software development in-house.

SCM and SQA are essential components of a successful software development effort, and their importance is expected to increase as software projects become global distributed endeavors. Standards and associated artifact descriptions for use on subsequent software projects will be created. As the demand for software engineers increases, the ability of an educational institution to supply software engineers from an environment that closely resembles the real world will benefit the school, industry and the student.

SiDiff kernel development will be researched to identify and procure the latest revision. The work of the textbook authors, C. Jones and O. Bonsignour [24], shall be used to establish the initial and ongoing metrics to be collected for the framework.

Acknowledgment Thanks to the PWMS team, The Eastech Company, Patricia Schweitzer for editing this paper and the FDU student chapter of the Association for Computing Machinery for the review and the proposal evaluation of this paper.

References

1. Fowler, M.: UML Distilled: A Brief Guide to the Standard Object Modeling Language, 3rd edn. Addison Wesley (2003)
2. Phillips, W., Tan, A.: Proposal for a predictive wastewater management system (PWMS), first phase. http://www.fdu-extlib.com/pwms/PWMSPROP.pdf. unpublished
3. Jacobson, I.: Object-Oriented Software Engineering – A Use Case Driven Approach, Addison Wesley (1992)
4. Bernd, B., Dutoit, A.: Object-Oriented Software Engineering: Using UML, Patterns and Java, 3rd edn. Prentice Hall (2010). (ISBN 0-13-606125-7)
5. Bersoff, E., Henderson, V., Siegel, S.: Software Configuration Management – An Investment in Product Integrity. Prentice Hall (1980)
6. Wenzel, S.: Unique identification of elements in evolving models: towards fine-grained traceability in model-driven engineering. URN: urn:nbn:de:hbz:467-5243 (2011)
7. Timo, K., Kelter, U., Gabriele, T.: A rule based approach to the semantic lifting of model differences in the context of model versioning. In: Proc. 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011), pp. 163–172, November 2011
8. Timo, K., Pit, P., Kelter, U., Maik, S.: Adaptability of model comparison tools. In: Proc. 27th IEEE/ACM International Conference on Automated Software Engineering (ASE 2012), Essen, Germany, pp. 306–309, September 2012
9. Timo, K., Kelter, U., Gabriele, T.: Consistency-preserving edit scripts in model versioning. In: Proc. 28th IEEE/ACM International Conference on Automated Software Engineering (ASE 2013), Palo Alto, California, pp. 191–201, November 2013
10. Timo, K., Kelter, U., Dennis, R.: Workspace updates of visual models. In: Proc. 29th IEEE/ACM International Conference on Automated Software Engineering (ASE 2014), Vasteras, Sweden (2014)
11. Phillips, W., Gunasekar, K.: System level requirements analysis document (RAD) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSSLRAD.pdf. unpublished
12. Phillips, W.: Requirements analysis document (RAD) data analysis subsystem (DAS) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSDASRAD.pdf. unpublished
13. Phillips, W., Venkannagari, S.: Requirements analysis document (RAD) data gathering subsystem (DGS) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSDGSRAD.pdf. unpublished

14. Phillips, W., Brahmarouthu, S.: Requirements analysis document (RAD) data storage subsystem (DSS) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSDSSRAD.pdf. unpublished
15. Phillips, W., Devarakonda, S.: Requirements analysis document (RAD) access control subsystem (ACS) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSACSRAD.pdf. unpublished
16. Phillips, W., Turlapati, D.T.: Requirements analysis document (RAD) audit and alert subsytem (AAS) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSAASRAD.pdf. unpublished
17. Phillips, W., Venkannagari, S., Gunasekar, K.: Software development plan (SDP) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSSDP.pdf. unpublished
18. Phillips, W., Munnuru, A.: Software configuration management plan (SCMP) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSSCMP.pdf. unpublished
19. Phillips, W., Peddi, A.: Software quality assurance plan (SQAP) for the Predictive Wastewater Management System (PWMS). http://www.fdu-extlib.com/pwms/PWMSSQAP.pdf. unpublished
20. Phillips, W.: Final Report for the Predictive Wastewater Management System (PWMS) Phase 1.0. Functional Baseline. http://www.fdu-extlib.com/pwms/PWMSFINAL.pdf. unpublished
21. Aiello, R., Sachs, L.: Configuration Management Best Practices (Practical Methods That Work in The Real World), Pearson Education, August 2010
22. Aiello, R., Sachs, L.: Agile Software Configuration Management (in-press)
23. Black and Veatch: Smart integrated infrastructure white paper. unpublished
24. Jones, C., Bonsignour, O.: The Economics of Software Quality. Addison Wesley (2012)

Automatic Reverse Engineering of Classes' Relationships

Maen Hammad, Rajaa Abu-Wandi and Haneen Aydeh

Abstract Classes are the core of object oriented systems. Any maintenance activity includes performing a code change to one or more classes. Any code change to one class may affect other classes in the project. So, developers need to be aware and fully understand the structure and the relationships between classes. This paper proposes a technique to automatically extract various types of class's relationships from source code. The proposed technique extracts relationships among classes and measures their involvements in relationships. Fan-in and Fan-out metrics are used to give developers more comprehensive picture about the current status of coupling for each class.

Keywords Reverse engineering · Software metrics · Class coupling

1 Introduction

Reverse engineering is essential to program comprehension [1]. Usually, maintenance tasks begin with program comprehension activities. Developers need to fully understand the current structure and status of the source code. They need to understand classes and how they are coupled together. Program comprehension process consumes time and effort when the code is badly written or complex. For object oriented projects, developers need mainly to understand classes and their relationships. Developers have to be fully aware of the relationships when they implement the change. This helps them in understanding the coupling between classes and how their code changes impact other classes. When a change is required to be implemented, as adding a method, maintainers have to determine which classes are related to that class in order to make suitable and compatible modifications on them. Manual inspection and tracking relationships in the

M. Hammad(✉) · R. Abu-Wandi · H. Aydeh
Department of Software Engineering, The Hashemite University, Zarqa, Jordan
e-mail: mhammd@hu.edu.jo, rajaaenad@yahoo.com, haneenaydeh@yahoo.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_111

source code consumes times and efforts of maintainers. It is also subject to human error. The solution to this problem is to develop an approach to automatically extracts various types of relationships and present them to developers. Another issue that faces designers and developers is coupling. Low coupling between classes is a design principle that should be targeted. So, developers should be aware about the degree of coupling between classes before or after implementing their changes. They will be happy if there is a tool that can be used to automatically give them the status and degree of coupling between classes. Many reverse engineering tools recover large amount of information about classes and architecture from code but they do not focus on relationships, especially associations and dependencies. There is a need for tools that recover mainly relationships between classes.

This paper presents a technique, supported by a tool, to automatically analyze source code and extracts various types of relationships; generalizations, associations and dependencies. The degree of coupling for each class is measured by utilizing the Fan-out and the Fan-in metrics. The results are presented in a way to help understanding the current status of coupling for classes.

The rest of this paper is organized as follows. Section 2 presents the related work in the area. Section 3 details the proposed technique. Finally, Section 4 concludes the paper with an overview about our future work.

2 Related Work

Ferenc et al. [2] presented a reverse engineering framework called Columbus to analyze large and legacy C++ projects. Columbus is a big system and requires large amount of operations. Rigi environment [3] is another example for large reverse engineering systems. Our technique is lightweight and is focused on relationships. Kraft et al. [4] presented a tool chain that enables experimentation and study of real C++ applications. The approach does not focus on different types of relationships. Yeh and Kuo [5] proposed a reverse engineering approach for aggregation relationships based on their primary characteristic. Favre [6] presented a generic and methodological guide meta-model driven process for reconstructing software architecture. Telea and Voinea [7] presented Integrated Reverse-engineering Environment (IRE) for C and C++. We differ in highlighting the coupling between classes based on relationships. Lanza and Ducasse [8] proposed a lightweight software visualization technique enriched with software metrics information for reverse engineering. We included Fan-in and Fan-out metrics for each class. A proposed visualization called class Blueprint is proposed by [9] to understand the internal structure of classes. The focus is not on different types of Fan-in and Fan-out of classes. Beck et al. [10] presented a node-link graph to visualize the coupling and dependencies among classes. Hammad and Rawashdeh [11] presented a framework to visualize and measure the class coupling from source code. Demeyer et al. [12] presented a hybrid approach for reverse engineering that combines both visualization and software metrics to

understand the program structure especially classes. In collusion, our work differs from the related work in the area with its flexibility. It is lightweight and does not require heavy processing. It is focused on identifying and measuring relationships directly from the source code.

3 The Proposed Technique

The aim is to automatically support developers with useful information about the current status of relationships and coupling in the system. In this paper, we identify and measure relationships on class level and package level. The steps of the proposed method are summarized as follows:

1. Java source code is transformed into the XML representation srcML.
2. For each class A, the following information is extracted by parsing the srcML representation of the class:
   (a) The name of the super class for class A.
   (b) The names of reference types used in the attributes of class A (associations).
   (c) The names of reference types used in the methods of class A (dependencies).
3. Fan-in and Fan-out values are calculated for each class based on identified relationships in Step 2.

The process starts by a specific Java source packages or complete project as an input. In the next step, the source code is transformed into srcML[13] representation. srcML[1] is a XML representation for the source code in which each code element is tagged with its syntactic information. For example, consider the following code:

public class ProfileCpp extends BaseProfile

The srcML representation of the above code portion is:

<class> <specifier>public</specifier>class <name>ProfileCpp</name> <super><extends>extends<name>BaseProfile</name></extends></super>.

srcML is parsed by a set of XPath queries. For example, the query "//class/super/extends/name/text()" is used to extract the name of a super class. For a specific class A, associations are identified by querying all reference types defined in the data fields of class A. The number of different extracted types is reported as the total number of associations that class A is involved with. After associations are extracted, all methods of class A are parsed one by one. For each method, all locally defined reference types are identified. The total number of unique locally defined types is reported as the number of dependencies that class A is involved with. The next step deals with counting and presenting the Fan-in and Fan-out measures for each class. The Fan-out value of a class A is calculated by counting the total number of different classes involved in all relationships

[1] http://www.srcml.org/

(generalization, associations and dependencies) defined in class A. The number of different classes is returned as the Fan-out value of class A. For example, if a class A has extended one super class and defined three types, the Fan-out is four resulted from one generalization and three associations. The Fan-in value of class A is calculated by parsing the srcML representation of all other classes to count how many different classes have referenced class A. All classes are checked to determine if any class has referenced class A as a super class, a data field attribute or locally in one method.

To support better comprehension for how much a class is involved in different relationships, each Fan-in or Fan-out value is categorized into three categories. The categorization is based on the type of relationships. A visual chart, that represents a Fan-in or a Fan-out value, is shown to the developer for each class. The chart is divided into three colored sectors to represent the three different types of relationships (generalizations, associations and dependencies). The size of each sector models the number of different classes involved in a specific relationship. The summation of all sectors is the Fan-in or Fan-out value. For example, Figure 1 shows graphical charts that represent the Fan-out and Fan-in measures for class freecsXmlRpcHandler from the open source project freeCS. The class has eight Fan-out relationships with other classes. These eight relationships are categorized into one generalization (red sector), two associations (yellow sector) and five dependencies (green sector). The Fan-in value for the class is one which is a dependency relationship. So, its chart will be a green circle with value one. Two charts are presented for each class; for the Fan-out metric and the Fan-in. These charts help developers to quickly understand how much a class is coupled and what relationships that are involved in the coupling. The charts give abstract view about coupling to developers. Developers can quickly determine highly coupled classes with large Fan-in and/or Fan-out values.

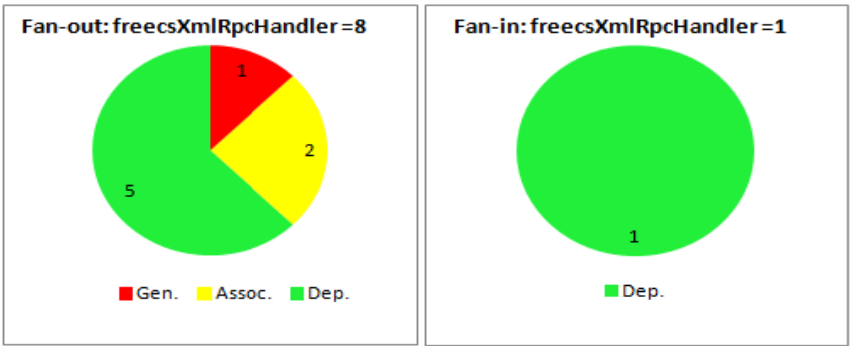


Fig. 1 Representation of Fan-out and Fan-out of class freecsXmlRpcHandler.

A preliminary working tool, **RelFinder** (Relationships Finder), has been developed to realize the technique. The tool also calculates both the Fan-in and the

Fan-out metrics for each class. The calculated values are categorized based on the type of the relationship. The tool reports useful information for developers. It reports the set of all relationships in the analyzed source code, the Fan-in and Fan-out values for each class, and data models needed to generate the visual charts for Fan-in and Fan-out.

4 Conclusions and Future Work

The paper presented a technique to automatically extract relationships between classes. All relationships are directly extracted from the source code. It also utilizes the Fan-in and Fan-out metrics to represent more comprehensive information about coupling between classes. Our future work aims to present more detailed visualizations to model coupling between methods and packages. We also plan to identify different types of coupling as control, stamp and data. We are currently working on developing a tool to save and track the evolution of relationships over time.

References

1. Canfora, G., Di Penta, M.: New frontiers of reverse engineering. In: Proc. IEEE Future of Software Engineering (FOS 2007), pp. 326–341 (2007)
2. Ferenc, R., Beszedes, A., Tarkiainen, M., Gyimothy, T.: Columbus-reverse engineering tool and schema for C++. In: Proc. IEEE International Conference on Software Maintenance (ICSM 2002), pp. 172–181, (2002)
3. Kienle, H.M., Muller, H.A.: Rigi-An environment for software reverse engineering, exploration, visualization, and redocumentation. Science of Computer Programming **75**(4), 247–263 (2010)
4. Kraft, N.A., Malloy, B.A., Power, J.F.: A tool chain for reverse engineering C++ applications. Science of Computer Programming **69**(1), 3–13 (2007)
5. Yeh, D., Kuo, W.-Y.: Reverse engineering aggregation relationship based on propagation of operations. In: Proc. IEEE Sixth European Conference on Software Maintenance and Reengineering, pp. 223–229 (2002)
6. Favre, J.-M.: Cacophony: Metamodel-driven software architecture reconstruction. In: Proc. IEEE Working Conf. on Reverse Engineering (WCRE 2004), pp. 204–213 (2004)
7. Telea, A., Voinea, L.: An interactive reverse engineering environment for large-scale C++ code. In: Proc. ACM Symposium on Software visualization, pp. 67–76 (2008)
8. Lanza, M., Ducasse, S.: Polymetric views-a lightweight visual approach to reverse engineering. IEEE Transactions on Software Engineering **29**(9), 782–795 (2003)
9. Ducasse, S., Lanza, M.: The class blueprint: visually supporting the understanding of glasses. IEEE Transactions on Software Engineering **31**(1), 75–90 (2005)
10. Beck, F., Petkov, R., Diehl, S.: Visually exploring multi-dimensional code couplings. In: Proc. International Workshop on Visualizing Software for Understanding and Analysis, pp. 1–8 (2011)

11. Hammad, M., Rawashdeh, A.: A Framework to Measure and Visualize Class Coupling. International Journal of Software Engineering and Its Applications **8**(4), 137–146 (2014)
12. Demeyer, S., Ducasse, S., Lanza, M.: A hybrid reverse engineering approach combining metrics and program visualization. In: Proc. Sixth IEEE Working Conference on Reverse Engineering (WCRE 1999), pp. 175–186 (1999)
13. Collard, M.L., Kagdi, H.H., Maletic, J.I.: An XML-based lightweight C++ fact extractor. In: Proc. IEEE International Workshop on Program Comprehension, pp. 134–143 (2003)

Developing Predictable Vehicular Distributed Embedded Systems on Multi-core

Saad Mubeen, Thomas Nolte and Kurt-Lennart Lundbäck

Abstract In this paper we address the challenges related to supporting model- and component-based development of predictable software for vehicular distributed embedded systems, utilizing multi-core platforms. We present a research plan for the development of new methods and techniques to deal with these challenges. The techniques will support various aspects such as modeling of the software architecture; supporting multiple criticality levels; verifying predictability of the system using end-to-end timing analysis; code generation; and providing a predictable run-time support on multi-core platforms by virtualizing a certified single-core real-time operating system. As a proof of concept, we will implement the newly developed techniques in a commercial tool chain (Rubus-ICE). The efficacy of the newly developed techniques and the extended tool chain will be demonstrated on the industrial case studies.

1 Introduction

A large share of customer value in modern vehicles comes from computer-controlled functionality that is realized by distributed embedded systems. With the recent advancements in the vehicular domain, the size and complexity of software in such systems has drastically increased [4]. The traditional software development techniques for these systems are no longer effective with respect to handling the complexity;

S. Mubeen(✉) · T. Nolte
Mälardalen University, Vasteras, Sweden
e-mail: saad.mubeen,thomas.nolte@mdh.se

K.-L. Lundbäck
Arcticus Systems AB, Jarfalla, Sweden
e-mail: kurt.lundback@arcticus-systems.com

The work in this paper is supported by the Swedish Foundation for Strategic Research within the project PRESS.

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_112

supporting reuse; reducing development costs, time to test and time to market. These challenges can be addressed by developing these systems using the principles of model- and component-based software engineering [7, 9, 11]. The developers of the these systems have to not only deal with their complexity but also guarantee their predictable timing behavior. This means, the timing behavior of these systems has to be verified, for instance, by using *a priori* schedulability analysis techniques such as end-to-end timing analysis [10, 12]. In addition, the execution platform should also provide predictable run-time support for these systems. There are several industrial tool chains, e.g. Rubus-ICE [2] and AUTOSAR [1] that support model-based development and predictable execution of the single-core systems. However, a lot of new challenges have emerged with the introduction of multi-core platforms for the execution of these systems [13]. A seamless tool chain for the development of vehicle software with *multiple criticality levels* and its predictable execution on partitioned single-core and multi-core platforms is missing in the state of the practice.

1.1 Research Goals and Challenges

We aim to develop techniques for model- and component-based software development of systems utilizing multi-core platforms. The techniques will support various development steps, i.e., from modeling of the software architecture to its synthesis and execution. The software in these systems can have multiple criticality levels. For instance, a high-critical function may co-exist with other low-critical functions on the same ECU. Multiple criticality levels in the vehicle software will be supported by means of virtual partitions in the core(s) of single-core as well as multi-core platforms. The main focus of this paper is on supporting predictable execution of these systems on such platforms. In this context, an end-to-end timing analysis framework will be developed to verify timing behavior of these systems. In order to provide a predictable run-time support for these systems on multi-core platforms, we aim to develop a virtualization technique that supports the reuse of a certified single-core Real-Time Operating System (RTOS) by means of a multi-core hypervisor.

Our goal is to implement the newly developed techniques in a tool chain that can be easily adopted by the vehicle industry. We plan to provide a proof-of-concept demonstrator by implementing these techniques in existing commercial tools that are actually used by the industry. The modeling and synthesis techniques will be implemented in the existing industrial model, the Rubus Component Model (RCM) and its tool suite Rubus-ICE [2]. The end-to-end timing analysis will be implemented as a plug-in for Rubus-ICE. Rubus RTOS, already certified in ISO 26262, currently supports single-core platforms. The newly developed multi-core hypervisor will reuse a separate instance of the Rubus RTOS per core. The efficacy of the extended tool suite will be demonstrated on the industrial use cases. The main research challenge can be formulated as follows. *"How to support model- and component-based development and predictable execution of software for vehicular distributed embedded systems on multi-core platforms?"*

2 Research plan and proposed contributions

We propose six phases to accomplish the research goals outlined in the paper.

1. A technique will be developed to support modeling of component-based vehicular distributed embedded systems on multi-core platforms.
2. The technique developed in the first phase will be extended to support modeling and development of systems that have more than one criticality level. This means, some parts of the software architecture (e.g., subsystems or chains of components) in these systems may have more stringent timing requirements compared to other parts. The motivation behind the need for multiple criticalities comes from the requirements on the certification of the system. The modeling of the software architecture for these systems will be implemented by means of virtual partitions in the core(s) of single-core and multi-core platforms respectively. The idea is to allocate the parts of the software architecture with different criticality levels to separate virtual partitions or cores. In this context, policies will be developed to handle dependencies and interferences among partitions.
3. In order to provide pre-runtime guarantees on the timing behavior of the system, the existing end-to-end timing analysis will be extended. Here, the focus will be on analyzing the end-to-end delays (age and reaction delays) [10] on distributed cause-and-effect chains. The analysis framework will also consider partitioned systems and multi-core platforms.
4. In order to provide a predictable run-time support for these systems on multi-core platforms, we will develop a virtualization technique that supports the reuse of certified single-core RTOS. This objective will be realized by developing a multi-core hypervisor that will instantiate the certified single-core RTOS for each core. One of the benefits for this approach is that there is no need to recertify the RTOS if the multi-core hypervisor can be certified. Moreover, virtual partitions will be supported.
5. We will provide the proof of concept by assembling a complete tool chain that will support modeling, timing analysis, synthesis and predictable execution of the systems on multi-core platforms. The modeling and synthesis techniques will be implemented in RCM and Rubus-ICE. The end-to-end timing analysis will be implemented as a plug-in for Rubus-ICE. The multi-core hypervisor will reuse a separate instance of the certified Rubus RTOS per core.
6. In this phase the validation of the newly developed techniques will be carried out. The tool chain, implementing the new techniques, will be validated on several industrial case studies provided by the industrial partners.

3 Related Work

The research community has produced a large body of research on schedulability analysis of mixed-criticality systems since the seminal paper by Vestal [14]. Burns and Davis [5] provide a survey on mixed criticality systems for both single-core and multi-core platforms. However, none of the existing works has addressed end-to-end

path delay analysis for distributed systems with multiple criticality levels on multi-core platforms. We will address this analysis in our work. The focus of the survey and related works is on scheduling and schedulability analysis. These works can be regarded as complementary to our work because schedulability analysis is one of our sub-goals in providing a model- and component-based development and predictable execution support for mixed criticality systems on partitioned single-core as well as multi-core platforms.

There are several ongoing research projects that explore prospects of multi-core platforms for embedded systems. EMC2, a large European project[1], aims to utilize multi-core technology for the execution of mixed-criticality systems. In this context, the focus of the project is to support adaptive systems in dynamic and changeable real-time environments by means of a service-oriented architecture approach. On the other hand, our focus is on the development of a seamless tool chain for model-based development and predictable execution of such systems. Moreover, we focus on a component-based architecture. Unlike EMC2, we aim to develop a virtualization technique by reusing the certified single-core RTOS to support predictable multi-core run-time support.

The MCC project[2] focuses on development of mixed-criticality systems on many-core platforms [3]. Whereas the focus of our work is on both partitioned single-core and multi-core platforms. Also, MCC does not provide the model- and component-based development support. SMARTCore [6] and other related works [8] exploit model-driven engineering for the development of multi-core systems. They focus on run-time monitoring, deployment optimization and back propagation of extra-functional properties from run-time to the system models. Whereas, our aim is to develop a seamless tool chain that supports various activities including modeling, timing analysis, synthesis and predictable execution of such systems. Also, these works do not address multiple criticality systems.

Similarly, there are other projects that address various aspects of mixed criticality and multi-core systems, e.g., RECOMP, CERTAINTY, PROXIMA, CONTREX and DREAMS. To the best of our knowledge, none of these projects seemingly address our sub-goals that include focus on the vehicular domain; model- and component-based software development; predicting the timing behavior with respect to the end-to-end path delays; and virtualization of certified single-core RTOS to support predictable multi-core run-time support.

References

1. AUTOSAR Techincal Overview, Release 4.1, Rev.2, Ver.1.1.0. http://autosar.org
2. Rubus models, methods and tools. http://www.arcticus-systems.com
3. Bate, I., Burns, A., Davis, R.: A bailout protocol for mixed criticality systems. In: 27th Euromicro Conference on Real-Time Systems (2015)

[1] http://www.artemis-emc2.eu/

[2] https://www.cs.york.ac.uk/research/research-groups/rts/mcc/

4. Broy, M., Kruger, I., Pretschner, A., Salzmann, C.: Engineering automotive software. Proceedings of the IEEE **95**(2), 356–373 (2007)
5. Burns, A., Davis, R.: Mixed criticality systems - a review. Tech. rep., Dept. of Computer Science, University of York (2015)
6. Ciccozzi, F., Corcoran, D., Seceleanu, T., Scholle, D.: Smartcore: boosting model-driven engineering of embedded systems for multicore. In: 12th International Conference on Information Technology, pp. 89–94. New Generations, April 2015
7. Crnkovic, I., Larsson, M.: Building Reliable Component-Based Software Systems. Artech House Inc., Norwood (2002)
8. Feljan, J., Ciccozzi, F., Carlson, J., Crnkovic, I.: Enhancing model-based architecture optimization with monitored system runs. In: 41st Euromicro Conference on Software Engineering and Advanced Applications, pp. 216–223, August 2015
9. Henzinger, T.A., Sifakis, J.: The embedded systems design challenge. In: 14th International Symposium on Formal Methods (2006)
10. Mubeen, S., Mäki-Turja, J., Sjödin, M.: Support for end-to-end response-time and delay analysis in the industrial tool suite: Issues, experiences and a case study. Computer Science and Information Systems **10**(1) (2013)
11. Mubeen, S., Mäki-Turja, J., Sjödin, M.: Communications-Oriented Development of Component-Based Vehicular Distributed Real-Time Embedded Systems. Journal of Systems Architecture **60**(2), 207–220 (2014)
12. Tindell, K., Clark, J.: Holistic schedulability analysis for distributed hard real-time systems. Microprocess. Microprogram. **40**, 117–134 (1994)
13. Vector Informatic GmbH: Autosar goes multi-core the safe way. In: Techncal Article, June 2014. http://vector.com
14. Vestal, S.: Preemptive scheduling of multi-criticality systems with varying degrees of execution time assurance. In: 28th IEEE International Symposium on Real-Time Systems, pp. 239–243, December 2007

Formalizing the Process of Usability Heuristics Development

Daniela Quiñones, Cristian Rusu, Silvana Roncagliolo, Virginica Rusu and César A. Collazos

Abstract Heuristic evaluation is one of the widely used methods for evaluating usability. Several authors have developed different sets of usability heuristics in order to evaluate the usability of specific applications. In this regard, it is important to know if authors use a formal methodology to develop their heuristics. This paper presents several such methodologies highlighting the importance of a formal usability heuristics development process.

Keywords Usability · Evaluation methods · Methodologies · Usability heuristics

1 Introduction

One of the widely used methods for evaluating usability is heuristic evaluation [1]. Several authors have developed different sets of usability heuristics in order to evaluate the usability of specific applications.

In this regard, it is important to know whether researchers use a formal methodology to develop their heuristics, which elements consider when formulating them, in which concepts are based for their specification and how the set of heuristics is validated.

D. Quiñones(✉) · C. Rusu · S. Roncagliolo
Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: danielacqo@gmail.com, {cristian.rusu,silvana}@ucv.cl

V. Rusu
Universidad de Playa Ancha de Ciencias de la Educación, Valparaíso, Chile
e-mail: virginica.rusu@upla.cl

C.A. Collazos
Universidad del Cauca, Popayán, Colombia
e-mail: ccollazo@unicauca.edu.co

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_113

This paper presents several methodologies that have been used to develop new sets of usability heuristics. Section 2 presents the concept of usability. Section 3 presents several methodologies to establish usability heuristics. The Section 4 presents conclusions, highlighting the importance of a formal process when developing new usability heuristics.

2 Usability

According to ISO 9241-11 standard, usability can be defined as [2]: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". The standard was updated and it is still under review [3].

3 Methodologies to Establish Usability Heuristics

Several authors have developed set of specific usability heuristics for different types of applications. Many of these sets of heuristics have been developed based on Nielsen's generic heuristics [4], and include specific heuristics to evaluate features that are not covered by Nielsen's heuristics.

Reviewing if these authors have used a formal methodology to establish their new set of heuristics is possible to determine that most of them did not use one, but:

- They use existing sets of heuristics (both generic and specific heuristics established by other authors) to generate their set including new heuristics [5, 6].
- Based on usability problems identified (by heuristic evaluation, user testing, among other methods), they generate their set of heuristics that are the inverse that problems found [7].
- They are translating guidelines and design recommendations on usability problems, and then generate their set of heuristics [8].

Besides, some authors proposed more formal methodologies to develop usability heuristics. These methodologies include phases to formulate heuristics.

On the one hand, authors in [9] present a methodology to develop a set of domain-specific usability heuristics. Their methodology has two phases. They perform a collaborative group session with usability experts to select, specify and create heuristics based on specific application in study. Then, they validate the heuristics with actual users. Finally, the results of both phases are analyzed to identify the points on which experts and users agree and disagree.

On the other hand, authors in [10] define a methodology to establish new usability heuristics, for specific applications. The methodology includes 6 stages: (1) Exploratory stage: to explore the specific applications that require new usability heuristics; (2) Descriptive stage: to re-examine the very meaning of usability and its characteristics; (3) Correlational stage: to identify the characteristics that the usability heuristics

for specific applications should have; (4) Explicative stage: to specify new heuristics; (5) validation stage: to evaluate the set of heuristics defined at STEP 4 against Nielsen's heuristics; and (6) Refinement stage: to refine the set of heuristics defined at STEP 4. Stages 1 to 6 may be applied iteratively. Specific usability checklist may also be developed, detailing usability heuristics and helping heuristic evaluations practice.

4 Conclusions

Many sets of usability heuristics have been developed for applications with aims and specific features, elements that generic heuristics (Nielsen) not considered, so it does not fully evaluate their usability.

However, it is possible to note that when creating new sets of heuristics most authors do not use a formal methodology. Mainly they follow an intuitive process.

If authors did not use a formal methodology to develop their heuristics, and are based only on informal processes, it is possible to establish a set of heuristics with some shortcomings as ambiguous and unclear definitions; difficult to understand or too large; sets of heuristics difficult to use or that not effectively evaluate the usability of the application to which they are aimed. This implies that the set should be refined n-times until it accomplishes its main objective: to evaluate the usability.

The process of developing usability heuristics would be facilitated if using a formal methodology. The methodology should clarify the relevant concepts to consider, how to specify heuristics properly and how to validate them. It would be certainly help to establish sets of heuristics that evaluate usability effectively and efficiently.

Acknowledgments Daniela Quiñones has been granted the "INF-PUCV" Graduate Scholarship.

References

1. Nielsen, J.: Usability engineering, 1st edn. Academic Press, USA (1993)
2. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDT's) – Part 11: Guidance on usability. International Organization for Standardization, Geneva (1998)
3. Bevan, N., Carter, J., Harker, S.: ISO 9241-11 revised: what have we learnt about usability since 1998?. In: Kurosu, M. (ed.) Human-Computer Interaction: Design and Evaluation, Lecture Notes in Computer Science. pp. 143–151. Springer (2015)
4. Nielsen, J.: Ten Usability Heuristics (1995). http://www.useit.com/papers/heuristic/heuristic_list.html
5. Alsumait, A.: Usability Heuristics Evaluation for Child Elearning Applications. Journal of Software **5**(6), 654–661 (2010)

6. Andreu-Vall, M., Marcos, M.: Evaluación de sitios web multilingües: metodología y herramienta heurística. El Profesional de la Información **21**, 254–260 (2012)
7. Pinelle, D., Wong, N., Stach, T.: Heuristic evaluation for games: usability principles for video game design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1453–1462 (2008)
8. Al-Razgan, M., Al-Khalifa, H., Al-Shahrani, M.: Heuristics for evaluating the usability of mobile launchers for elderly people. In: Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience, vol. 8517, pp. 415–424 (2014)
9. Lechner, B., Fruhling, A., Petter S., Siy, H.: The chicken and the pig: user involvement in developing usability heuristics. In Proceedings of the Nineteenth Americas Conference on Information Systems (2013)
10. Rusu, C., Roncagliolo, S., Rusu V., Collazos, C.: A methodology to establish usability heuristics. In: Proceedings ACHI2011: The Fourth International Conference on Advances in Computer-Human Interactions, pp. 59–62 (2011)

Analysis of a Training Platform for the Digital Battlefield, Based on Semiotics and Simulation

Cristian Barría, Cristian Rusu, Claudio Cubillos, César Collazos and Miguel Palma

Abstract In the development of a system, the perception of usability is vital, moreover in this context, the role of designers is a fundamental pillar between the interaction of different interfaces with the users. This is how semiotics is appreciated as a factor in the expression of ideas at the moment of building a proposal regarding the implementation of a digital battlefield simulator for training prototype, taking into account other avant-garde experiences in this area.

Keywords Semiotics · Usability · Digital battlefield · Meta-comunication

1 Introduction

When developing a software as a tool, the interaction that should exist between the user and it, needs to be specially considered from the perspective of usability [1].

As proof of the concept´s relevance, usability has been made part of the set of norms regarding quality and management established by the International Standard Organization (ISO), which has defined it as: "The degree with which a product can be used by specific users with the purpose of achieving goals with effectiveness, efficiency and satisfaction in a context of specific usage" [2].

C. Barría · C. Rusu · C. Cubillos
Pontificia Universidad Católica de Chile, Valparaíso, Chile
e-mail: {cristian.barria,cristian.rusu,claudio.cubillos}@ucv.cl

C. Collazos
Universidad del Cauca Popayán, Cauca, Colombia
e-mail: ccollazo@unicauca.edu.co

M. Palma(✉)
Universidad Tecnológica de Chile, Santiago, Chile
e-mail: miguel.palma06@inacapmail.cl

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_114

This document represents a literary review of the semiotics and simulation, and describes the development of a training system for the personnel which will be or is employed in a cyber defense context. Section 2 presents semiotics in the development of a simulator. Section 3 explains the analysis of a digital battlefield simulator. Section 4 presents conclusions and future work.

2 Semiotics in the Development of a Simulator for Battlefield Training

The results of each project must be as unique as its context [3]. Within the development of a software, certain aspects related to the meaningful learning process of the trainer must be privileged, and the following components must be considered:

1) Framework
2) Architecture
3) Design

3 Analysis of a Digital Battlefield Simulator for Training

The simulation software, CyberCIEGE, is analyzed from the perspective of the use of semiotics. This is a tool which teaches concepts of security and of networks.

An individual analysis about meta-linguistics symbols, static symbols and dynamic symbols was made, from which a meta-communication message was constructed for each one. Finally, the global quality of the meta-communication was evaluated.

Metalinguistics

Metalinguistics is present through the selection of a sign or image which provides information about itself. The linguistic function relates semiotics at the moment in which a symbol provides you with information when lighting up.

Static Signs

Static signs are managed in the contextual menu through the unfolding of the same, allowing the user to have a clear vision regarding the distinct possibilities of interaction, in front of different situations that he will have to confront.

Dynamic Signs

The dynamic signs that are exposed in the superior area, in which the user interacts because, for example, he can control the different actions that occur, which gives them an immediate vision of the situation that is happening inside the organization. The following is an identified dynamic symbol.

Global Quality of the Meta-Communication

Given the revision of the meta-linguistic signs mentioned, it was concluded that the ones that make reference to the interface's language are static symbols which have no causal or temporal relation, and finally, dynamic symbols which have a relation with temporal and causal aspects that arise from the interaction with the web site. In the following paragraphs, some of the most representative elements are mentioned, which were identified in the phase of analysis developed in the previous points:

1) *Meta-linguistic Signs:*
   When observing an element, this presents a description which details, in a way that is not only composed by the image presented by the user, but also indicates its functionality and other precedents of the same.
2) *Static Symbols:*
   The designer offers a superior and lateral bar in which the principal situations of the system are displayed, and a lateral bar which allows the user to share, compare and save information regarding the status of their means.
3) *Dynamic Symbols:*
   When positioning over an image, this gets dark and a text in its inferior part appears with a brief description.

4 Conclusions

Semiotics plays an important role in the development of a simulator, because the user is the one that will be confronted with the tool, so the designer must think as the user might when developing a simulator, because the interface must proportionate him with guidance and information. Unless this happens, the tool will not fit its meta-communicative purpose, as the different symbols will not be understood equally by the different subjects. A possibility of a future work is to develop lab tests in such a way that the evaluation of the different signs of interaction within semiotics is possible, in simulation.

References

1. Garrido, M., Lavín, C., Rodríguez, N.: Medición de usabilidad de trámites públicos en línea en Chile: Un caso de estudio en Gobierno Electrónico, vol. 21(1), TECSI FEA USP, Santiago (2014)
2. International Standard ISO 9241, Ergonomic requirements for office work with visual display terminals. Ginebra: ISO/IEC (1998)
3. Cummings, M.: User Experience.UX Design, Recuperado el 04 de 11 de 2014, de User Experience.UX DEsign (2010). http://uxdesign.com/ux-defined
4. Eco, U.: Tratado de Semiótica general, Lumen S.A., Barcelona (2000)
5. Scolari, C.: Hacer clic: hacia una sociosemiótica de las interacciones digitales, p. 30. Gedisa, Barcelona (2004)

6. Garrett, J.: The elements of user experience. New Riders, Estados Unidos (2011)
7. Ausubel, D.: Teoría el aprendizaje significativo. Trillas, México (1983)
8. Esquembre, F.: Creación de simulaciones interactivas en JAVA: aplicaciones a la enseñanza de la física, p. 10. Pearson Educación, Madrid (2004)
9. Magariños, J.: Archivo de Semiótica: Manual de estudios semióticos, p. 50 (2007)
10. Sieckenius de Souza, C., Junqueira, S., Olivera, R.: A semiotic engineering approach to user interface design. Elsevier, Rio de Janeiro (2001)

Usability Heuristics and Design Recommendations for Driving Simulators

Aníbal Campos, Cristian Rusu, Silvana Roncagliolo, Fabiola Sanz, Raúl Gálvez and Daniela Quiñones

Abstract User eXperience (UX) is one of the most important aspects when developing a simulator. It ensures that the final product is not just functional, but users find a value in what they are getting. The paper presents a set of heuristic to assess usability in driving simulators. It also defines a set of design recommendations focused on UX.

Keywords Driving simulator · User experience · Usability · Usability heuristics · Design recommendations

1 Introduction

Driving simulators are objects of learning that attempt to model part of a reply of the reality phenomena. They are interactive environments, which allow users to modify parameters and to see how the system reacts. Simulators are used in many disciplines.

The ISO standard 9241-11 defines usability as "the degree to which a product can be used by specific users to accomplish certain specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [1]. The ISO 9241-210 standard defines User eXperience (UX) as the result of the perceptions and answers of a person for the use and early use of a product, system or service [2]. UX extends the usability concept. The heuristic evaluation is probably the most widely used usability inspection method [3].

The paper presents a set of usability heuristics for driving simulators and a set of design recommendations. Section 2 presents the proposed heuristics. Section 3 briefly references the UX – oriented design recommendations. Finally, Section 4 presents the conclusions and future work.

A. Campos · C. Rusu · S. Roncagliolo · F. Sanz · R. Gálvez · D. Quiñones(✉)
Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
e-mail: anibal.campos.baez@gmail.com, faby.sanz.b@gmail.com,
raul.galvez.cabeza@gmail.com, danielacqo@gmail.com, {cristian.rusu,silvana}@ucv.cl

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_115

2 A Set of Usability Heuristics for Driving Simulators

The set of usability heuristics for driving simulators (HS) was developed based on iterative methodology [4]. Three iterations have been done. The set of heuristic is presented below. Heuristics are presented in an abbreviated form: (ID), Name and Definition.

(HS1) *Virtual environment*: The driving simulator should have a virtual environment that provides the user the possibility to drive in several real-life situations.
(HS2) *Visibility of the vehicle*: The driving simulator should give the user control over the vehicle components and enable set the visual effects.
(HS3) *Consistency and standard*: The driving simulator must be consistent in the use of language and visual elements used in accordance with reality.
(HS4) *Cabin of control*: The driving simulator must have a dashboard of the vehicle easy to understand by the user.
(HS5) *Simplicity*: The driving simulator must not overload the user with information.
(HS6) *Challenges and rewards*: The driving simulator must explain which is the purpose or goal of the given lessons.
(HS7) *Level of difficulty*: The simulator must have different levels of difficulty, allowing implement several real-life situations.
(HS8) *User memory load*: The user should not have to remember information of a part of dialog one part to another.
(HS9) *Camera control*: The user must be able to see the vehicle from different angles, allowing the user to select the desired view.
(HS10) *Vehicle control*: The driving simulator must offer the option to use diverse controls that allow the user to interact with the virtual environment, as well as select the type of vehicle transmission.
(HS11) *User control and freedom*: The driving simulator must provide emergency exits, undo and redo.
(HS12) *Flexibility and efficiency of use*: The driving simulator adapts to different work styles and also lets the user customize the type of vehicle.
(HS13) *Learning and continuity*: The driving simulator provides lessons in line with current traffic regulations and records the results.
(HS14) *Prevention of errors*: The driving simulator must prevent the user commits critical errors that may cause an infraction.
(HS15) *Error Recovery*: The driving simulator must clearly indicate to the user how to recover of a mistake.
(HS16) *Help and Documentation*: The driving simulator should provide help information, easy to find for the user.

The development of the HS usability heuristics was based on three case studies: *SimuDrive*, *City Car Driving*, and *DriverTest*. Several experiments were performed.

The refined set of HS heuristics was compared against Nielsen's usability heuristics [5]. Three experts evaluated *DriverTest* using HS heuristics (experimental group); a total of 50 usability issues were identified. Other three experts evaluated *DriverTest* using Nielsen's heuristics (control group); a total of 25 usability issues were identified. The HS heuristics allowed finding additional problems that were not considered when using Nielsen's heuristics, associated with the specific characteristics of the driving simulators, and directly related to the learning process.

3 Design Recommendations for Driving Simulators

Some basic features that a driving simulator should have are described by Navarrete et al. [6]. Based on these features, the Peter Morville's UX model [7], and the set of usability heuristics previously defined, the following design recommendations were proposed. They are presented in an abbreviated form: (ID) and Name.

(RS01) *Learning rules of the road*. (RS02) *Real challenges*. (RS03) *Virtual environment*. (RS04) *Exit Control*. (RS05) *Flexibility of use*. (RS06) *Easy to remember*. (RS07) *Vehicle control*. (RS08) *Vehicle status*. (RS09) *View Control*. (RS10) *Control cabin element*. (RS11) *Consistent language*. (RS12) *Simplicity*. (RS13) *Control of difficulty*. (RS14) *Error prevention*. (RS15) *Error recovery*. (RS16) *User Help*.

4 Conclusions

A set of usability heuristics for driving simulators was developed and validated through three case studies, and based on expert evaluators' feedback. Experimental validation shows that the heuristics we proposed provides better results than generic usability heuristics. We also proposed a set of UX – oriented design recommendations for driving simulators.

As future work, we intend to validate the design recommendations through the implementation of functional prototypes, which will then be tested through usability inspections and UX testing.

References

1. ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDT's) – Part 11: Guidance on usability. International Organization for Standardization, Geneva (1998)
2. ISO 9241-210. Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. International Organization for Standardization (2010)
3. Nielsen, J.: Usability engineering. Academic Press, San Diego (1993)

4. Rusu, C., Roncagliolo, S., Collazos, C.: A methodology to establish usability heuristics. In: The Fourth International Conference on Advances in Computer-Human Interactions (ACHI 2011), pp. 59-62. IARIA (2011)
5. Nielsen, J.: Ten Usability Heuristics (2005). http://www.useit.com/papers/heuristic/heuristic_list.html (accessed August 04, 2015)
6. Navarrete, R., Urquiza, M., Hernandez, V., Madrid, M.: Generalidades sobre el entrenamiento de conductores y el desarrollo de simuladores de manejo. Sanfandila, Ciudad de México (2004)
7. Morville, P.: User Experience Basics (2014). http://www.usability.gov/what-and-why/user-experience.html (accessed April 16, 2015)

Model for Describing Bioprinting Projects in STL

Luiz Angelo Valota Francisco and Luis Carlos Trevelin

Abstract The bioprinting is an essential element of biomanufacturing process and is an area of emerging and promising multidisciplinary research, which proposes the "manufacture" of tissues and organs by means of techniques and rapid prototyping capabilities, such as additive manufacturing. A model and a framework for the definition of projects that can be used in bioprinting living tissues and organs is set.

Keywords Additive manufacturing · Bioprinting · Biomodel · BioCAD · Bioprinting · Modeling · Stereolithography · 3D printing

1 Introduction

Biomanufacturing is directly related to the process of manufacturing biological products [1]. Currently there are many studies focused on biomanufacturing of living tissue. According to Nakamura, Nisiyama and Henmi (2008), the engineering of some fine and simple tissues was already established and clinically applied, as is the case of skin and cornea engineering. But there are still some open questions regarding the manufacture of thick and complex tissues composed of different kinds of cells and microstructure [2]. According to Khatiwala et al. (2012), 3D cellular bioprinting has the potential to expand the application of cell-based therapies as a tool to generate tissues and organs of complex functions and structures used in transplants [3]. This work proposes a computational model to describe biosystems biomanufacturing projects focused on living tissues and organs.

L.A.V. Francisco(✉)
Federal Institute of Education, Science, and Technology of São Paulo, São Paulo, Brazil
e-mail: lavfrancisco@ifsp.edu.br

L.C. Trevelin
Federal University of São Carlos (UFSCar), D.C, São Carlos, SP, Brazil
e-mail: trevelin@dc.ufscar.br

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_116

2 Biomanufacturing

The living tissues and organs bioprinting is not a simple task. According to Mironov, Kasyanov and Markwald (2011), biomanufacturing a human organ will demand the use of a series of automated and integrated robotic devices or a biomanufacturing line [4]. The phases of a biomanufacturing process are shown in Figure 1.

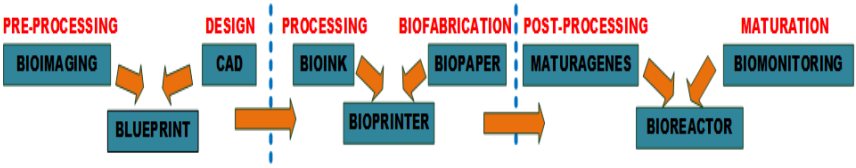


Fig. 1 Biomanufacturing phases [4].

3 Cad Modelling in Bioprinting Processes

Internal and external geometries from the structure to be built can be obtained from a 3D drawing software and be described by mathematical equations. These geometries can also come from clinical data from MRI or CT. The CAD file describes the geometry and dimensions of the piece to be built. The STL file lists the triangle coordinates that are used to form the projected 3D surface [5].

Advances in computer-aided technologies and their applicability in tissue engineering created a new field called BioCAD. This field integrates digital drawing, image processing, projection, modelling, simulation and manufacture of biological tissue and organ replacements [6].

The Visualization ToolKit (VTK) is a free and open source system. It is object-oriented and multi-platform and is used in 3D computer graphics, scientific visualization and image processing to represent data like images, polygons, meshes and volumes. [7].

According to Amorim et al. (2011), the Centro de Tecnologia da Informação Renato Archer – CTI developed the InVesalius software to create physical models that aid surgeries. [8].

4 Computational Model for Bioprinting Projects

We use the biomanufacturing phases defined by Mironov, Kasyanov and Markwald (2011) as a template for our model, focused on the first phase. [4].

4.1 Framework for Computational Bioprint Modelling

The proposed framework uses VTK 5.8.0 for data representation and Java for user interaction and database access. Our framework was based on the bioprinting model proposed by Mironov [4]. It uses, in its first version, a STL file as input, which comes from the InVesalius software. The bioprinting project creation using our framework was developed to allow a specialist to describe the details of the bioprinting process.

5 Assessment

The experiment began with reading a STL file using our framework. The chosen object was a cube with 27.95 cm2 area in each face. This object was divided in slices of thickness 0.05 cm, creating 559 slices. For each slice, we can add two types of parameters: textual and visual. Textual parameters are descriptive information regarding the object, and visual parameters are primary geometric figures – squares, circles and rectangles. In this experiment, we created visual parameters described by a rectangle. These parameters were added to 55 slices, which amount to approximately 10% of the whole volume.

6 Conclusion and Future Work

Our proposed model was conceived based on these phases in order to bioprint living tissues [4].

We also developed a framework to assess our model. This assessment was made by an experiment of printing a STL file, formatted and parameterized to its printing.

The main contribution of this work is the bioprinting model that can be used as a template in projects to be printed. Even though our work proposed a viable option to create living tissue bioprinting projects, there is still a long way to go, both in hardware and software wise.

References

1. Nakamura, M., et al.: Computer-Assisted Biofabrication: The challenges on manufacturing 3-D biological tissues for tissue and organ engineering. Symposium on VLSI Technology Digest of Technical Papers, pp. 2–5, June 2011
2. Nakamura, M., Nishiyama, Y, Henmi, C.: 3D Micro-fabrication by Inkjet 3D biofabrication for 3D tissue engineering IEEE, November 2008
3. Khatiwala, C., et al.: 3D cell bioprinting for regenerative medicine research and therapies. Gene Therapy and Regulation, vol. 07(01), p. 1230004, dez. 2012

4. Mironov, V., Kasyanov, V., Markwald, R.R.: Organ printing: from bioprinter to organ biofabrication line. Current Opinion in Biotechnology, vol. 22(5), pp. 667–673, out. 2011
5. Wang, C.-S., Chang, T.-R.: Re-triangulation in STL meshes for rapid prototyping and manufacture. The International Journal of Advanced Manufacturing Technology **37**(7–8), 770–781 (2007)
6. Sun, W.: BioCAD in tissue science and engineering IEEE, ago. 2009
7. Júlio de Mesquita, F.: Reconstrução e geração de malhas em estruturas biomecânicas tridimensionais para análise por elementos finitos. Universidade Estadual Paulista, Bauru (2008)
8. Amorim, P.H.J., Moraes, T.F., Azevedo, F.S., Silva, J.V.L.: InVesalius: Software Livre de Imagens Médicas. XXXI Congresso da Sociedade Brasileira de Computação, pp. 1735-1740 (2011)

The Fractal Nature of Mars Topography Analyzed via the Wavelet Leaders Method

Adrien Deliège, Thomas Kleyntssens and Samuel Nicolay

Abstract This work studies the scaling properties of Mars topography based on Mars Orbiter Laser Altimeter (MOLA) data through the wavelet leaders method (WLM). This approach shows a scale break at $\approx$ 15 km. At small scales, these topographic profiles display a monofractal behavior while a multifractal nature is observed at large scales. The scaling exponents are greater at small scales. They also seem to be influenced by latitude and may indicate a slight anisotropy in topography.

1 Introduction

Previous works about the scaling properties of Mars topography revealed two distinct scaling regimes (at small scales and at large scales) while the scale break varies from one paper to another. These works do not analyze the whole surface of Mars. The power-law exponents, or Hurst exponents, associated with these scaling laws appear to differ according to the region considered but some common features can still be noted. Besides, the mono- or multifractality of Mars topography has been studied, although it may depend on the definition of fractality and of the tools used. Table 1 summarizes the main previous results.

These studies are all based on along-track measurements, which implies that the 2D part of the topographic field has not been taken into account. In this paper, we perform our analysis on the MOLA data, using the 128 pix/deg map (more details can be found in [4] and on http://pds-geosciences.wustl.edu).

The aim of this paper is twofold: perform a complete study of the surface roughness of Mars while taking both longitudinal and latitudinal topographic profiles into account and show that the WLM may be a suitable candidate for the study of scaling properties of planetray surfaces. The WLM is the only tool that is theoretically justified [5] and

A. Deliège(✉) · T. Kleyntssens · S. Nicolay
Institute of Mathematics, University of Liège, Liège, Belgium
e-mail: {adrien.deliege,tkleyntssens,s.nicolay}@ulg.ac.be

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_117

Table 1 Some previous results

Methods	small scales	large scales
power spectral density (PSD) [1]	$H \approx 1.2$ (< 10 km)	$H \approx 0.2 - 0.5$
variance of a wavelet transform [2]	$H \approx 1.25$ (< 24 km)	$H \approx 0.5$
statistical moments [3]	$H \approx 0.76$ (< 10 km)	$H \approx 0.52$

that has already been used to study the bifractality of a signal [6] (for details, see e.g. [5, 7, 8]).

In this work, we use the D3 Daubechies wavelet. Let $d_{\lambda_{j,k}}$ be the wavelet leader a function f associated to the cube $\lambda_{j,k} = k/2^j + [0, 1/2^{j+1})$. One sets

$$S(j,q) = 2^{-j} \sum_{\lambda} d_{\lambda}^{q} \quad \text{and} \quad \eta(q) = \liminf_{j \to +\infty} \frac{\log S(j,q)}{\log 2^{-j}}.$$

If λ contains a point with Hurst exponent H, then $d_\lambda \sim 2^{-Hj}$ and $\eta(q) = Hq$.

2 Results

For a fixed value of q (ranges from -2 to 2 in order to limit the influence of anomalously large coefficients in the computations), $\eta(q)$ is obtained by performing linear regressions of $j \mapsto \log_2 S(j,q)$. Fig. 1 reveals a scale break at $\approx$ 15 kilometers. We thus consider two scaling regimes: the first one at the small scales ($<$ 15km) and the second one at the large scales ($>$ 60km); the scales in between represent the transition from one regime to the other.

If η is "linear enough", i.e. the associated linear correlation coefficient c is greater than 0.98, then the signal is said monofractal at the scales used to build η and the slope of η corresponds to the Hurst exponent, which characterizes the irregularity at these scales; otherwise it is multifractal at those scales. In this case, the slope of the best-fit (in the least square sense) linear regression of η gives a scaling exponent which does not fully represent the irregularity of the signal; other notions have then to be used (see e.g. [3, 5]) but are beyond the scope of this work. Table 2 summarizes our results.

At small scales, our results are in agreement with previous studies for the longitudinal signals (e.g. [1, 2]). For the latitudinal signals, the drop in the proportion of monofractal signals may be explain by the crustal dichotomy of Mars, the presence of polar caps, among others. Moreover, if we only keep latitudes between $80°S$ and $80°N$, then more than 96.7% of the signals are monofractal. The influence of latitude can be clearly seen in Fig. 1. The difference of results between the longitudinal and latitudinal signals may indicate a slight anisotropy of the surface roughness at small scales, as mentioned in [9].

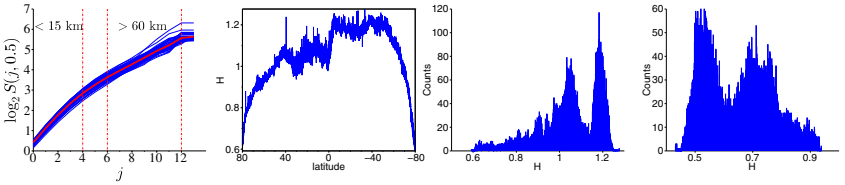


Fig. 1 Left: Blue lines: $\log_2 S(j, 0.5)$ versus j for several longitudinal bands. Red line: the mean of these functions. The scale j corresponds to $0.463 * 2^{j+1}$ kilometers (1 pixel ≈ 0.463 km). Scales 0 to 4 are used for the first scaling regime and scales 6 to 12 delimit the second scaling regime. Middle Left: Exponent H as a function of latitude at small scales. Middle Right: The corresponding histogram of the distribution of the exponents H. Right: The histogram of the distribution of H at large scales.

Table 2 Results obtained with the WLM for the longitudes (l) and latitudes (L)

scales	monofractality (l)	mean H (l)	monofractality (L)	mean H (L)
small	99.7%	1.15 ± 0.06	92.1%	1.05 ± 0.13
large	91.7%	0.78 ± 0.087	63.2%	0.65 ± 0.11

At large scales, it turns out that the longitudinal and the latitudinal analyses mostly display a multifractal behavior, as in [3]. A large percentage of latitudinal bands still has to be considered monofractal; however, there is a clear difference compared to the small scales case (Fig. 1).

3 Conclusion

This work confirms that the WLM is well-suited for studying the irregularity of planetary bodies. Since the WLM can be adapted to 2D signals, we will examine the scaling properties of Mars topography using the 2D version of the WLM.

References

1. Aharonson, O., Zuber, M., Rothman, D.: Statistics of Mars' Topography from the Mars Orbiter Laser Altimeter: Slopes, Correlations, and Physical Models. Journal of Geophysical Research **106**(E10), 23723–23735 (2001)
2. Malamud, B., Turcotte, D.: Wavelet analyses of Mars polar topography. Journal of Geophysical Research **106**(E8), 17497–17504 (2001)
3. Landais, F., Schmidt, F., Lovejoy, S.: Universal multifractal Martian topography. Nonlinear Processes in Geophysics Discussions **2**, 1007–1031 (2015)
4. Smith, D., Neumann, G., Arvidson, R.E., Guinness, E.A., Slavney, S.: Mars global surveyor laser altimeter mission experiment gridded data record, mgs-m-mola-5-megdr-l3-v1.0. Technical report, NASA Planetary Data System (2003)
5. Jaffard, S.: Wavelet techniques in multifractal analysis. Proceedings of Symposia in Pure Mathematics **72**, 91–152 (2004)

6. Nicolay, S., Brodie of Brodie, E.B., Touchon, M., Audit, B., d'Aubenton Carafa, Y., Thermes, C., Arneodo, A.: Bifractality of human DNA strand-asymmetry profiles results from transcription. Phys. Rev. E **75**, 032902, March 2007
7. Daubechies, I.: Ten lectures on Wavelets. SIAM (1992)
8. Meyer, Y., Salinger, D.: Wavelets and Operators, vol. 1. Cambridge University Press (1995)
9. Alvarez-Ramirez, J., Rodriguez, E., Cervantes, I., Echeverria, J.: Scaling properties of image textures: A detrending fluctuation analysis approach. Physica A **361**, 677–698 (2006)

Privacy Enhancement in E-Mail Clients

Prabhat Kumar, Jyoti Prakash Singh, Rajni Kant Raman and Rohit Raj

Abstract Privacy on top of security is a desired feature in web mail clients and requires more attention from research community. This paper describes a system which aims at providing improved privacy features in web mail clients so that a user can open emails even in presence of other people without disclosing anything about the subject or the sender of the mail or both as per the privacy requirement. The work has been implemented by tweaking the source code of squirrel mail open source.

Keywords WSN privacy · Context privacy · Random walk · Dynamic routing · Safety period

1 Introduction

Privacy is a multidimensional concept which encompasses attitude and behavior of people [3]. Privacy [1] concerns exist wherever uniquely identifiable data related to a person or persons are collected and stored, in digital form or otherwise. Internet privacy [2] is the ability to determine what information one reveals or withholds about oneself over the Internet, who has access to such information, and for what purposes one's information may or may not be used. For example, email users would be concerned if their email was being accessed by third parties for privacy reasons. They may even not like the idea of others reading their emails even in their presence. Our aim is to enhance the existing privacy features in open source web mail clients. This paper identifies possible privecy vulnerabilities in web mail clients and proposes a novel approach to enhance the privecy features being provided in e-mail clients.

P. Kumar · J.P. Singh(✉) · R.K. Raman · R. Raj
Department of Computer Science and Engineering, National Institute of Technology Patna, Patna, Bihar, India
e-mail: {prabhat,jps,rohit124197}@nitp.ac.in, rajnikant7008@gmail.com

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_118

2 Proposed Solution for Privacy Issues in E-mail System

The privacy of web mail clients can be enhanced by concealing the information visible in the inbox such as the subject and the sender of the email. The proposed solution provides three levels of privacy. The user can hide a) only the sender of the mail, b) only the subject of the mail, or c) both the sender as well as the subject of the mail.

2.1 Initial Stages of Development

The displayed sender or the subject information can be concealed by replacing the desired string with its bitwise complement for each charecter. This will serve the purpose but has certain limitations like:

- The number of characters in the original and encoded string remains same and thus the pattern may become vulnerable with time.
- The user may want to remember who the sender of the mail is but with the characters being replaced by characters in the ASCII set this may be difficult for humans.
- Lastly the implementation requires bit level operations which may demand heavy use of resources.

2.2 Using Inbuild Encryption Function

The crypt() function in PHP can hide the sender and the subject using one or more encryption algorithms such as DES, Blowfish or MD5. PHP checks what arguments are available and what algorithms to use when it is installed.

3 Methodology

For implementation, we used SquirrelMail [4] which is a standards-based webmail package written in PHP. It includes built-in pure PHP support for the IMAP and SMTP protocols, and all pages render in pure HTML 4.0 or maximum compatibility across browsers. We have also used hMailServer [5] which is a free, open source, e-mail server for Microsoft Windows. The necessary changes were made in the source code and new files were incorporated to provide the additional features. Figure 1 shows the overall system design.

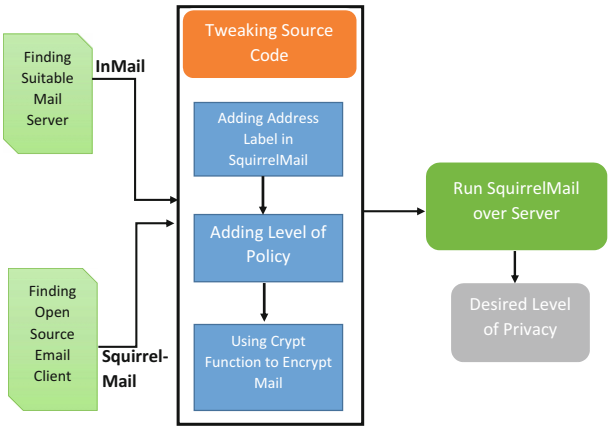


Fig. 1 System Design

Fig. 2 a)User inbox without privacy b)User inbox with privacy

3.1 Blending Privacy with Security

Encryption fuction may provide different levels of privacy using in-built funtion of PHP. Our purpose was to scan the subject of each incoming message and to filter messages such that it will be encrypted if the user has added it to his list.

3.2 Level of Privacy

An additional label in the address home page can be implemented for deciding differnt levels of privacy. a) A value of 1 implies that the user wants to hide the information only about the sender email address of the email. b) A value of 2 implies that the user wants to hide the information only about the subject of the email. c) A value of 3 implies that user wants to encrypt both the sender as well as the subject of the email. d) Any other value in this field will be treated as normal display of mail.

4 Results and Conclusions

The functionalities were successfully incorporated in an open source web mail client. Figure 2 (a) and (b) depicts user view of the SquirrelMail inbox prior and after the implementation of the proposed work respectivily. The subject and the sender of certain mails are encrypted while others are displayed normally. Thus, an enhanced level of privacy can be achieve as per the requirements of the users.

References

1. Srivastava, A., Geetha kumari, G.: Measuring privacy leaks in online social networks. In: International Conference on Advances in Computing, Communications and Informatics (ICACCI) (2013)
2. Patel, V., Juric, R.: Internet users and online privacy: a study assessing whether Internet users' privacy is adequately protected. In: 23rd International Conference on Information Technology Interfaces (2001)
3. Li, Y.: A multi-level model of individual information privacy beliefs. Electron. Commer. Rec. Appl. **13**, 32–44 (2014)
4. SquirrelMail. http://squirrelmail.org/, http://en.wikipedia.org/wiki/SquirrelMail
5. hMailServer. https://www.hmailserver.com/

Author Index